Ordering through learning in two-dimensional Ising spins

Pranay Bimal Sampat,[1,2,*] Ananya Verma ,[1,†] Riya Gupta ,[1,‡] and Shradha Mishra[1,§]
[1]*Department of Physics, Indian Institute of Technology (BHU), Varanasi 221005, India*
[2]*Department of Physics, Brown University, Providence, Rhode Island 02912, USA*

(Received 29 December 2021; revised 22 September 2022; accepted 7 November 2022; published 22 November 2022)

We study two-dimensional Ising spins, evolving through reinforcement learning using their state, action, and reward. The state of a spin is defined by whether it is in the majority or minority with its nearest neighbors. The spin updates its state using an ϵ-greedy algorithm. The parameter ϵ plays a role equivalent to the temperature in the Ising model. We find a phase transition from long-ranged ordered to a disordered state as we tune ϵ from small to large values. In analogy with the phase transition in the Ising model, we calculate the critical ϵ and the three critical exponents β, γ, ν of magnetization, susceptibility, and correlation length, respectively. A hyperscaling relation $d\nu = 2\beta + \gamma$ is obtained between the three exponents. The system is studied for different learning rates. The exponents approach the exact values for a two-dimensional Ising model for lower learning rates.

DOI: 10.1103/PhysRevE.106.054149

I. INTRODUCTION

The two-dimensional Ising model [1] is a prototype model which has been used to study various magnetic systems. The model has proven to be helpful in understanding the basic features of magnetic materials. Nonmagnetic systems such as the lattice gas model for a liquid gas phase transition [2,3] can also be mapped to the Ising model. Several studies have investigated the Ising model theoretically and numerically [4–7]. Prior numerical investigations of the model have predominantly been done using Maxwell-Boltzmann statistics and the Metropolis-Hastings algorithm. However, the system can be studied alternatively: where a spin learns from its experience and also from its neighbors. Although in previous studies, the Ising model with memory is explored with some time-dependent coupling (kernel) in the Hamiltonian [8,9], these models can be solved only for specific kernels. This paper aims to investigate a many-body spin system where the spin acts as an agent that learns from its experience.

In the current paper, we develop an algorithm for the two-dimensional Ising spin system using a reinforcement learning (RL) approach [10–12]. RL is a branch of machine learning [13–15]. It is based on choosing suitable actions to maximize reward (appreciation) in a particular problem where the agent learns from its experience [16,17]. Previously, RL has been studied in many areas, such as game theory [18], operations research [19], information theory [20], statistics, etc. In recent years, an RL framework has been used to model interacting systems as well [16,17,21,22]. Our study uses RL to study two-dimensional Ising spins. Here, the particles take action by considering the state of neighbors. To do so, we have used Q learning [15], which is a basic form of reinforcement learning where an agent uses Q values (also called action values) to optimize its actions iteratively. The objective is to optimize a value function suited to a given environment.

We are introducing a way to study the many-particle interacting system, where we do not require a Hamiltonian, hence this can be applied to systems where the Hamiltonian is not known. The dynamics of the spins evolve through their states, actions (whether or not to change their orientation), and the reward associated with each action. The reward associated with the actions incentivizes the spin to align with the majority within its four nearest neighbors on the square lattice. At each step, spins learn from their previous actions and update the reward associated with their actions. This is done using the ϵ-greedy algorithm [23], either to flip or retain its spin at each step. ϵ plays a role analogous to that of temperature in models which employ the Metropolis-Hastings algorithm to simulate the Ising model [5]. The observables we are calculating are the measure of ordering and fluctuations present in the system, hence they can be compared with the physical observables such as magnetization, susceptibility, and the Binder cumulant. We find a transition from long-range ordered (finite magnetization) to a disordered state (zero magnetization) by tuning ϵ from small to large values. We studied systems of different sizes and learning rates, performed finite-size scaling [24], and extracted the critical exponents β, γ, and ν.

The rest of the paper is divided in the following manner. Section II gives the details of the algorithm we introduced to study the system using reinforcement learning. In Sec. III we discuss our main results and calculation of various observables. Section IV discusses the finite-size scaling, and finally, we conclude our study in Sec. V.

II. LEARNING MODEL FOR TWO-DIMENSIONAL ISING SPIN

We consider a two-dimensional system of Ising spins ($S = \pm 1$) on a square lattice of size $L \times L$ with a periodic

*pranayb.sampat.phy16@itbhu.ac.in
†ananyaverma.phy18@itbhu.ac.in
‡riyagupta.phy18@itbhu.ac.in
§smishra.phy@iitbhu.ac.in

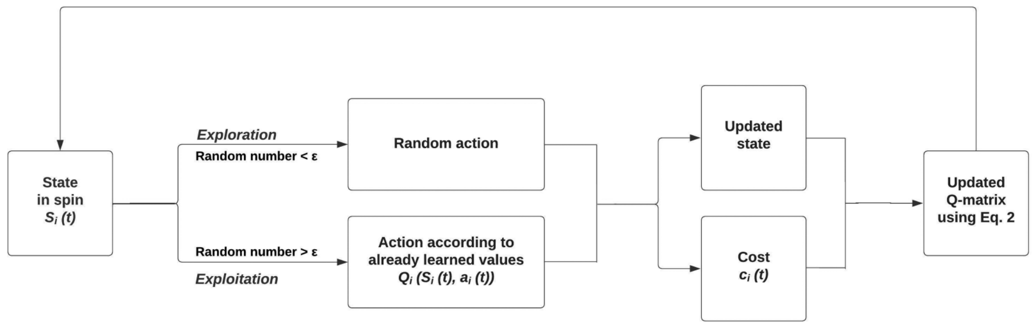

FIG. 1. Schematic block diagram of the model. The agent here is the spin with its own $\mathcal{Q}$ matrix. The state of the spin $[\mathcal{S}_i(t)]$ is defined by whether it is in the majority or minority. The spin can be in either of the states. The action is taken according to the ϵ-greedy algorithm, that balances both exploration and exploitation. During exploitation, the spin's action is chosen as the action with a minimum $\mathcal{Q}$ value, whereas in exploration, the spin can either be flipped or left unchanged without any bias. Based on the action, an updated state and cost are provided. The $\mathcal{Q}$ matrix is updated according to Eq. (2). The spin uses the updated state and $\mathcal{Q}$ matrix and goes through the same process in a loop until a steady state is reached.

boundary condition in both directions. Previously, the Ising model is exactly solved in one and two dimensions or numerically studied using the Metropolis-Hasting algorithm [5]. In the Metropolis-Hastings algorithm, the temperature enters through the Boltzmann probability distribution $P \propto \exp(-\beta E)$, where E is the local energy of the state and $\beta = \frac{1}{K_B T}$ is the inverse temperature [2,3]. The ratio of interaction strength and temperature controls the degree of spin fluctuation. The system shows a transition from long-ranged ordered to disordered state in two dimensions on increasing temperature or decreasing interaction strength.

In this study, at each time step, we select a spin and define its state as $\mathcal{S}_i(t) = +1$ if the spin is in the majority of its nearest neighbors (aligned with two or more nearest-neighbor spins) and $\mathcal{S}_i(t) = -1$ if it is in the minority (aligned with less than two nearest-neighbor spins). Each spin has its own $\mathcal{Q}$ matrix corresponding to its states and actions $(\mathcal{Q}_i[\mathcal{S}_i(t), a_i(t)])$. The actions $a_i(t)$ of a spin are of two types: exploration (choosing a random action) and exploitation (choosing actions based on already learned $\mathcal{Q}$ values).

To prevent the action from always taking the same path and possibly overfitting, we will introduce another parameter called ϵ to handle this during training. Instead of just selecting the optimal learned $\mathcal{Q}$-value action, the spin sometimes explores the action space further. A higher ϵ value results in the spin taking actions with a greater cost on average. The action $a_i(t)$ is taken using the ϵ-greedy algorithm [23] for the parameter ϵ as given below:

$$a_i(t) = \begin{cases} \text{argmin } \mathcal{Q}_i[\mathcal{S}_i(t), a_i(t)], & \text{probability } 1-\epsilon, \\ \text{random action}, & \text{probability } \epsilon. \end{cases} \tag{1}$$

Thus, for large ϵ values, the system performs more random exploration. We further define the cost function $c_i(t) = 0$ or 1, if the chosen action leads the spin to its majority or minority, respectively. The $\mathcal{Q}$ matrix is updated iteratively with the following equation as given in Ref. [15],

$$\mathcal{Q}_i[\mathcal{S}_i(t), a_i(t)] \leftarrow \mathcal{Q}_i[\mathcal{S}_i(t), a_i(t)] + \alpha[c_i(t) - \mathcal{Q}_i[\mathcal{S}_i(t), a_i(t)]], \tag{2}$$

where α is the learning rate. We studied the system for different learning rates $\alpha = 0.0001$, 0.001, 0.05, and 0.1. The parameters ϵ and α control the exploration and exploitation, respectively. A function of (ϵ, α) will control the amount of fluctuations present in the spin degrees of freedom. For each realization, we fixed α and varied ϵ as a temperature-like parameter in MCMC. The details of the RL scheme are explained in Fig. 1. The scheme given in Fig. 1 is repeated for all the spins one by one, and a time step is defined as an update of all the spins at once. The total run time of the simulation is $t = 1.2 \times 10^7$, and averaging is performed after a simulation step of $\tau = 1.18 \times 10^7$. The system is studied for various sizes from $L = 16$ to 96. A total of 80 different realizations are produced, and averaging is performed to improve the quality of the data. Each realization was initiated with a random distribution of up and down spins on the two-dimensional square lattice, and all the values of the $\mathcal{Q}$ values were set as 0. The system is studied for various values of randomness parameter $\epsilon \in (0, 1)$.

III. RESULTS

We define the magnetization order parameter $M(\epsilon) = \langle \frac{|\sum_{i=1}^{N} S_i(t)|}{N} \rangle$, for different ϵ, where $\langle \cdots \rangle$ implies the average value over time in the steady state and across multiple realizations. We show the plots for only two values of $\alpha = 0.0001$ and 0.05. The results (data) for other values of α are given in Table I.

In Figs. 2(a) and 2(b), we plot the $M(\epsilon)$ vs ϵ for different system sizes L for $\alpha = 0.0001$ and 0.05, respectively. For all L, $M(\epsilon)$ remains close to 1 (majority of spins ordered in the same direction) for small values of ϵ and approaches zero (random arrangement of spins on the lattice) as ϵ is increased. Thus, there is a phase transition from ordered $M \simeq 1$ to disordered state as a function of ϵ. We further calculate the fluctuations in M, or the susceptibility, defined as $\chi(\epsilon) = \langle M(\epsilon)^2 \rangle - \langle M(\epsilon) \rangle^2$, for different L, where $\langle \cdots \rangle$ have the same meaning as defined previously. The plot of $\chi(\epsilon)$ vs ϵ is shown in Figs. 2(c) and 2(d), for different system sizes L. We determine $\epsilon_c(L)$ by calculating the location of the maximum of susceptibility $\chi_{\max}(L)$. $\epsilon_c(L)$ decreases to

TABLE I. Obtained values of ϵ_c, critical exponents ν, γ, β, LHS and RHS of hyperscaling relation among the three exponents for different values of learning rate, α.

α	ϵ_c	ν	γ	β	$d\nu$	$2\beta+\gamma$
0.0001	0.1692 ± 0.001	0.9894 ± 0.00042	1.7253 ± 0.01679	0.1127 ± 0.00456	1.9788 ± 0.00084	1.9507 ± 0.02592
0.001	0.1689 ± 0.001	0.9499 ± 0.00167	1.6562 ± 0.02982	0.1003 ± 0.00456	1.8998 ± 0.00343	1.8568 ± 0.06334
0.05	0.1718 ± 0.001	0.6958 ± 0.00153	1.2117 ± 0.00058	0.1167 ± 0.01437	1.3916 ± 0.00306	1.4451 ± 0.0310
0.1	0.17095 ± 0.001	0.7934 ± 0.00315	1.3739 ± 0.00494	0.1193 ± 0.01227	1.5869 ± 0.00632	1.6125 ± 0.02949

lower ϵ for larger L. We also calculated $\chi_{\max}(L)$ for different system sizes and found that it increases with increasing L. To further characterize the phase transition, we calculate the fourth-order moment of the mean magnetization $M(\epsilon)$, the Binder cumulant (BC), defined as $U(\epsilon) = 1 - \frac{\langle M(\epsilon)^4\rangle}{3\langle M(\epsilon)^2\rangle^2}$. In Figs. 2(e) and 2(f), we show the variation of $U(\epsilon)$ vs ϵ for different L. $U(\epsilon)$ remains close to $2/3$ in the ordered state for small ϵ and smoothly decays to zero in the disordered phase for large ϵ. We then performed a finite-size analysis of the data to characterize the critical behavior in the thermodynamic limit, from the finite-size (L) data shown in Figs. 2(a) and 2(b), Figs. 2(c) and 2(d), and Figs. 2(e) and 2(f).

IV. FINITE-SIZE ANALYSIS

We use a finite-size scaling (FSS) analysis to understand the dependence of the three quantities $M(\epsilon)$, $\chi(\epsilon)$, and $U(\epsilon)$ on the system size $L \times L$. We assume that the FSS forms for these quantities are the same as those for a system of a two-dimensional Ising model in previous studies [5],

$$M(\epsilon, L) = L^{-\beta/\nu}\mathcal{M}(\epsilon - \epsilon_c)L^{1/\nu}, \tag{3a}$$

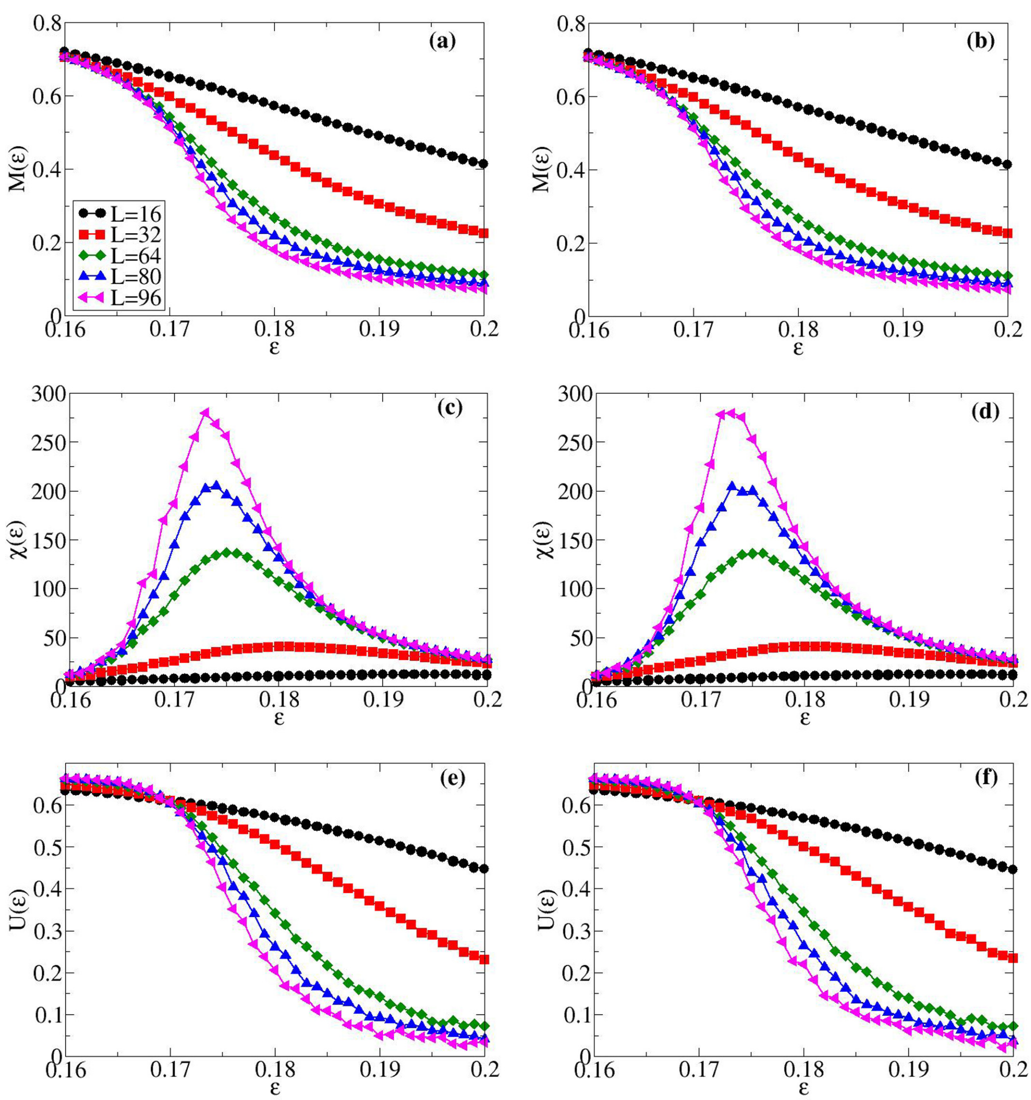

FIG. 2. Variation of (a) and (b) M, (c) and (d) χ, and (e) and (f) U vs ϵ for $\alpha = 0.0001$ and $\alpha = 0.05$, respectively.

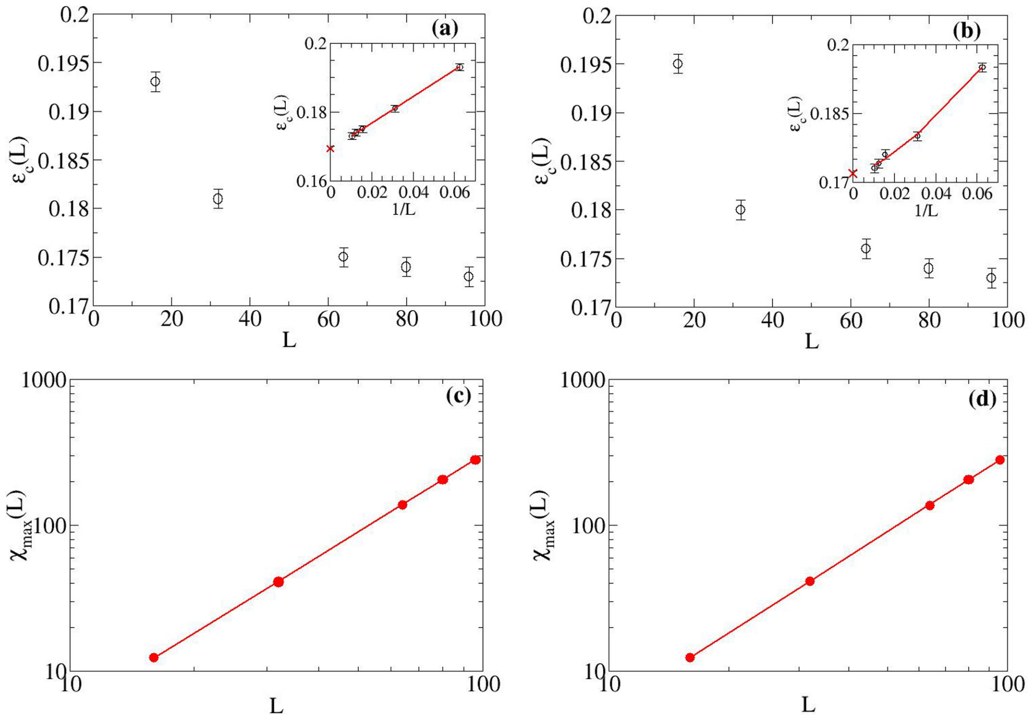

FIG. 3. Variations of (a), (b) $\epsilon_c(L)$ vs L, with insets showing the best fit to Eq. (4) and ϵ_c (marked as a red cross on the y axis). [$\epsilon_c = 0.1692 \pm 0.001$, $\nu = 0.9894 \pm 0.000\,42$] and [$\epsilon_c = 0.1718 \pm 0.001$, $\nu = 0.6958 \pm 0.001\,53$] for $\alpha = 0.0001$ and $\alpha = 0.05$, respectively. (c), (d) $\chi_{\max}(L)$ vs L on a log-log scale. The lines in each are the best fits to the form $\sim L^{\gamma/\nu}$, where $\gamma = 1.7253 \pm 0.016\,79$ and $1.2117 \pm 0.000\,58$ for $\alpha = 0.0001$ and 0.05, respectively.

$$\chi(\epsilon, L) = L^{\gamma/\nu}\chi(\epsilon - \epsilon_c)L^{1/\nu}, \quad (3b)$$

$$U(\epsilon, L) = \mathcal{U}(\epsilon - \epsilon_c)L^{1/\nu}, \quad (3c)$$

where $\mathcal{M}$, χ, and $\mathcal{U}$ are the scaling functions for the mean magnetization, susceptibility, and the Binder cumulant, respectively. The exponents ν, γ, and β are the exponents for the correlation length, susceptibility, and magnetization, respectively, for the standard Ising model [24]. We are going to describe below the FSS analysis of our results within the numerical accuracy of our data using the RL model. We plot $\epsilon_c(L)$, the location of the maximum of susceptibility $\chi_{\max}(L)$, as a function of system size in Figs. 3(a) and 3(b) for $\alpha = 0.0001$ and 0.05, respectively. Further, we assume that the pseudocritical point for different system size L is

$$\epsilon_c(L) = \epsilon_c + c_1 L^{-1/\nu}, \quad (4)$$

where $\epsilon_c \equiv \epsilon_c(L \to \infty)$. In the insets of plots in Figs. 3(a) and 3(b), the lines shows the fit of $\epsilon_c(L)$ with respect to $1/L$ and three parameters ϵ_c, c_1, and ν are obtained from the fitting. The values of these three parameters we found are [$\epsilon_c = 0.1692 \pm 0.001$, $c_1 = 0.3925$, $\nu = 0.9894 \pm 0.000\,42$] and [$\epsilon_c = 0.1718 \pm 0.001$, $c_1 = 1.2430$, $\nu = 0.6958 \pm 0.001\,53$] for $\alpha = 0.0001$ and $\alpha = 0.05$, respectively. We further plot the $\chi_{\max}(L)$ versus system size L in Figs. 3(c) and 3(d) on a log-log scale. In a finite system of size L, $\chi_{\max}(L) \sim L^{-z}$, where $z = \gamma/\nu$. This way, we extracted the exponent γ. The power-law fit to the plot gives $z = 1.7438 \pm 0.0164$ and $z = 1.7417 \pm 0.0047$ for $\alpha = 0.0001$ and $\alpha = 0.05$, respectively. Using the values of ν from Figs. 3(a) and 3(b), the exponent $\gamma = 1.7253 \pm 0.016\,79$ and 1.212 ± 0.0006 for $\alpha = 0.0001$ and 0.05, respectively. We further calculate the β exponent for the magnetization. At the critical ϵ_c, we calculate the root mean square of magnetization $M_{\rm rms} = \sqrt{M(\epsilon_c)^2} \sim L^{-\beta/\nu}$ [24]. The power-law fit provides the exponent β/ν. We found that $\beta = 0.1127 \pm 0.0046$ and $0.1167 \pm 0.014\,37$ for the two values of the learning rate $\alpha = 0.0001$ and 0.05, respectively. Table I shows the values of the three exponents from our calculation for different α values and the exact values of these exponents for the standard two-dimensional Ising model are in given in Table II. The exponents obtained for $\alpha = 0.05$ differ from the previous study of the two-dimensional Ising model theory [25–28]. But as we lower the learning rate α, the exponents converge to the exact values for the two-dimensional Ising model. Hence, a slower learning rate converges the RL model to the two-dimensional Ising model. Altogether, our RL method is able to predict the phase transition similar as from the other methods and shows similar phase transition characteristics.

We further check the hyperscaling relation between the three exponents. The hyperscaling relation among the three exponents is $d\nu = 2\beta + \gamma$, which is obtained from the scaling hypothesis of the underlying scaling functions, and implies

TABLE II. Values of critical exponents ν, γ, β, obtained from two-dimensional Ising model theory [25–28].

Critical exponents	Two-dimensional Ising model theory [25–28]
ν	1
β	0.125
γ	1.75

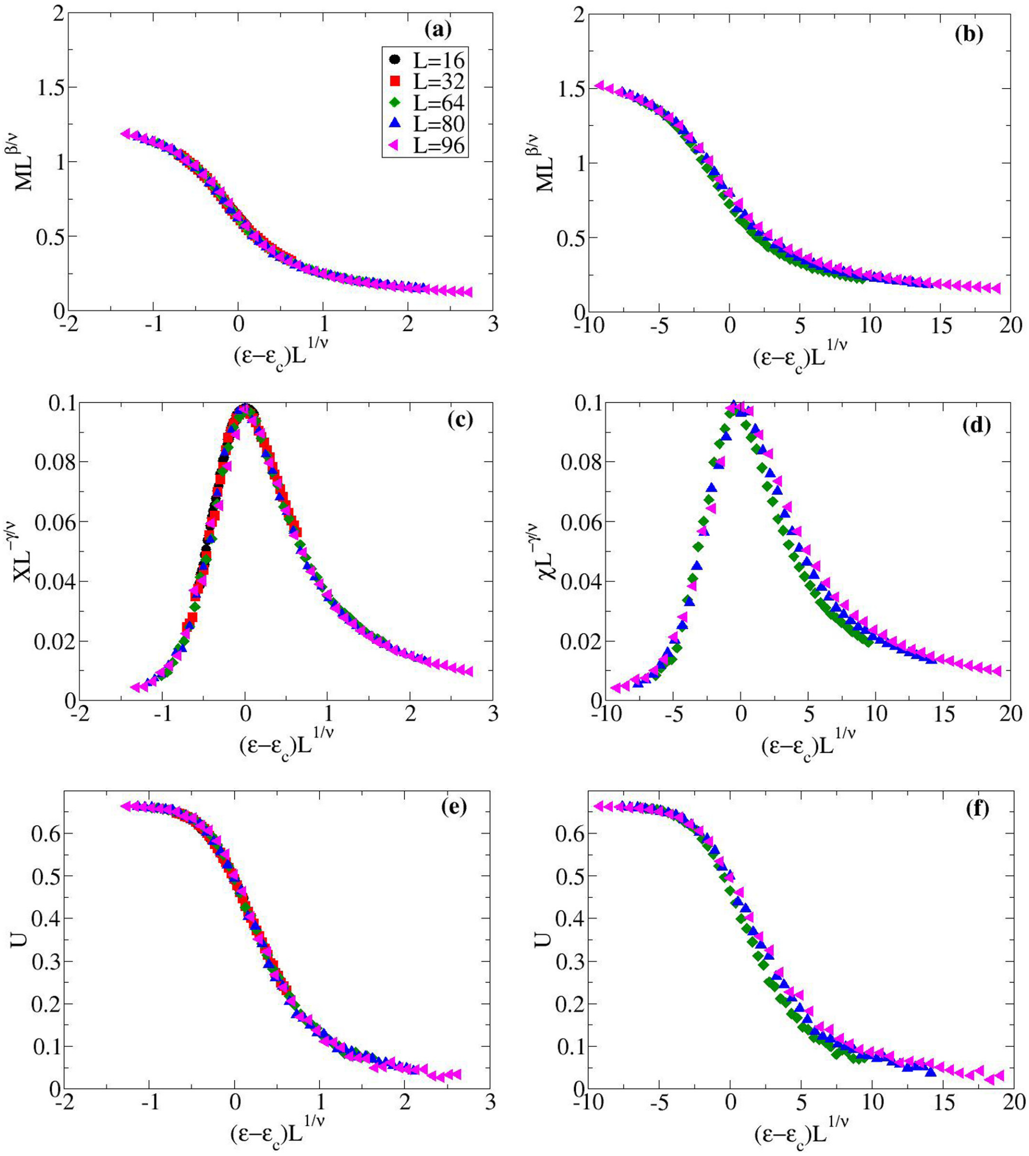

FIG. 4. Scaling collapse of (a) and (b) $M(\epsilon)$, (c) and (d) $\chi(\epsilon)$, and (e) and (f) $U(\epsilon)$, by plotting according to Eqs. (3a)–(3c) for $\alpha = 0.0001$ and $\alpha = 0.05$, respectively.

that all the observables can be calculated by the partition function of the system. $d = 2$ is the dimensionality of space. We then substitute the values of the exponents we found in our RL study. The left-hand side of the relation $d\nu$ and right-hand side of the relation $2\beta + \gamma$ are reported in the sixth and seventh columns of Table I. We also observe that the accuracy of the hyperscaling relation improves on lowering the learning rate α. Hence again model converges to the two-dimensional Ising model for smaller α.

We also show the plots of the scaling collapse of data for magnetization M, susceptibility χ, and Binder cumulant U using the values of the three exponents found in our study for $\alpha = 0.0001$ and 0.05. In Figs. 4(a) and 4(b), we plot the $ML^{\beta/\nu}$ vs $(\epsilon - \epsilon_c)L^{1/\nu}$ using Eq. (3a), and find a good collapse of data. We further plot the scaled susceptibility $\chi L^{-\gamma/\nu}$ vs $[\epsilon - \epsilon_c(L)]L^{1/\nu}$ in Figs. 4(c) and 4(d) using Eq. (3b) and find a good collapse of data for different system sizes. Similarly, we plot the U vs scaled ϵ, $(\epsilon - \epsilon_c)L^{1/\nu}$, in Figs. 4(e) and 4(f) using Eq. (3c) and again find a good scaling collapse of data. The scaling collapse of data also improves by lowering the learning rate. We also find the scaling collapse for other α values (data not shown).

V. DISCUSSION

In this paper, we used an RL framework to investigate Ising spins in two dimensions with a periodic boundary condition in both directions. Each spin can have two characteristic states and it can switch between these states using an action selected using the ϵ-greedy algorithm. We find that if the spins get rewarded for staying in the majority, then the system shows a phase transition from a disordered to ordered state on decreasing ϵ. Hence, ϵ plays a role similar to the temperature in the Metropolis Monte Carlo for a two-dimensional Ising model.

We further characterize the phase transition by calculating the critical point and different critical exponents for magnetization, susceptibility, and correlation length β, γ, and ν, respectively. Data show a scaling collapse for scaled magnetization, susceptibility, and the Binder cumulant. We observe the exponents match with the exact exponents for a two-dimensional Ising model for a lower learning rate α. Hence our RL model converges to the two-dimensional Ising model for lower learning rates.

Our current study provides a reinforcement learning approach to study the spin system. It can be further used for other many-particle interacting systems, where the form of the Hamiltonian is not known. The rudimentary nature of this approach makes it adaptable to study many-particle complex systems [16,21,29,30].

ACKNOWLEDGMENTS

The support and the resources provided by PARAM Shivay Facility under the National Supercomputing Mission, Government of India at the Indian Institute of Technology, Varanasi are gratefully acknowledged by all authors. S.M. thanks DST-SERB India, ECR/2017/000659 and CRG/2021/006945 for financial support. S.M., A.V., and P.S. also thank the Centre for Computing and Information Services at IIT (BHU), Varanasi.

[1] E. Ising, Z. Phys. **31**, 253 (1925).
[2] R. K. Pathria, *Statistical Mechanics* (Elsevier, Amsterdam, 1972).
[3] K. Huang, *Statistical Mechanics*, 2nd ed. (Wiley, New York, 1987).
[4] L. Onsager, Phys. Rev. **65**, 117 (1944).
[5] K. Binder, Monte Carlo investigations of phase transitions and critical phenomena, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green, Vol. 17b (Academic, New York, 1976).
[6] K. Binder and D. Stauffer, *Applications of the Monte Carlo Method in Statistical Physics* (Springer, Berlin, 1984), pp. 1–36.
[7] K. Binder, Rep. Prog. Phys. **60**, 487 (1997).
[8] S. Galam, Physica A **238**, 66 (1997).
[9] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).
[10] R. S. Sutton, A. G. Barto, and F. Bach, *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA, 1998).
[11] L. Panait and S. Luke, Auton. Agent. Multi-Agent Syst. **11**, 387 (2005).
[12] L. Busoniu, R. Babuska. and B. De Schutter, IEEE Trans. Syst., Man., Cybern. C **38**, 156 (2008).
[13] M. Kubat, *An Introduction to Machine Learning* (Springer, Basingstoke, U.K., 2015).
[14] D. E. Goldberg and J. H. Holland, Mach. Learn. **3**, 95 (1988).
[15] C. J. C. H. Watkins and P. Dayan, Mach. Learn. **8**, 279 (1992).
[16] M. Durve, F. Peruani, and A. Celani, Phys. Rev. E **102**, 012601 (2020).
[17] C. Wang, J. Wang, and X. Zhang, in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (IEEE, New York, 2018), pp. 1228–1232.
[18] M. Zinkevich, A. Greenwald, and M. L. Littman, in *Proceedings of the 18th International Conference on Neural Information Processing Systems (NIPS'05)* (ACM, New York, 2005), pp. 1641–1648.
[19] T. Gabel and M. Riedmiller, in *Proceedings of the 2007 IEEE Symposium on Approximate Dynamic Programming and Reinforcement Learning (ADPRL 2007)* (IEEE, New York, 2007).
[20] I. M. de Abril and R. Kanai, in *2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil* (IEEE, 2018), pp. 1–6.
[21] M. Gerhard, A. Jayaram, A. Fischer, and T. Speck, Phys. Rev. E **104**, 054614 (2021).
[22] D. Krongauz and T. Lazebnik, bioRxiv:2022.06.09.495429.
[23] A. dos Santos Mignon and R. L. de Azevedo da Rocha, Procedia Comput. Sci. **109C**, 1146 (2017).
[24] J. Cardy, *Finite-Size Scaling* (Elsevier/North-Holland, Amsterdam, 1988).
[25] M. Fisher, J. Math. Phys. **5**, 944 (1964).
[26] C. Domb, Adv. Phys. **9**, 245 (1960).
[27] L. Kadanoff, Nuovo Cimento B **44**, 276 (1966).
[28] M. Fisher, *Phase Transitions and Critical Phenomena* (International Atomic Energy Agency (IAEA), Vienna, 1969).
[29] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. **75**, 1226 (1995).
[30] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).