

Curvature-driven shifts of the Potts transition on spherical Fibonacci graphs: A graph-convolutional transfer-learning study

Zheng Zhou(周政), Xu-Yang Hou(侯旭阳), and Hao Guo(郭昊)
Citation: Chin. Phys. B, 2026, 35 (**7**): 070505. DOI: 10.1088/1674-1056/ae6637
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Thermodynamics of classical one-dimensional generalized nonlinear Klein-Gordon lattice model

Hu-Wei Jia(贾虎伟) and Ning-Hua Tong(同宁华)
Chin. Phys. B, 2025, 34 (**8**): 080501. DOI: 10.1088/1674-1056/add50f

Impacts of bus holding strategy on the performance and pollutant emissions of a two-lane mixed traffic system

Yanfeng Qiao(乔延峰), Ronghan Yao(姚荣涵), Baofeng Pan(潘宝峰), and Yu Xue(薛郁)
Chin. Phys. B, 2024, 33 (**11**): 110502. DOI: 10.1088/1674-1056/ad7e98

Thermodynamics of charged AdS black hole surrounded by quintessence in restricted phase space

Siyu Jian(简思雨), Siying Long(龙思颖), Juhua Chen(陈菊华), and Yongjiu Wang(王永久)
Chin. Phys. B, 2024, 33 (**10**): 100504. DOI: 10.1088/1674-1056/ad624f

Transport property of inhomogeneous strained graphene

Bing-Lan Wu(吴冰兰), Qiang Wei(魏强), Zhi-Qiang Zhang(张智强), and Hua Jiang(江华)
Chin. Phys. B, 2021, 30 (**3**): 030504. DOI: 10.1088/1674-1056/abe3e3

Dynamic phase transition of ferroelectric nanotube described by a spin-1/2 transverse Ising model

Chundong Wang(王春栋), Ying Wu(吴瑛), Yulin Cao(曹喻霖), and Xinying Xue(薛新英)
Chin. Phys. B, 2021, 30 (**2**): 020504. DOI: 10.1088/1674-1056/abbbfd

---

Curvature-driven shifts of the Potts transition on spherical Fibonacci graphs: A graph-convolutional transfer-learning study

Zheng Zhou(周政)[1,†], Xu-Yang Hou(侯旭阳)[1,†], and Hao Guo(郭昊)[1,2,‡]

[1] *School of Physics, Southeast University, Jiulonghu Campus, Nanjing* 211189, *China*
[2] *Hefei National Laboratory, Hefei* 230088, *China*

(Received 13 February 2026; revised manuscript received 22 April 2026; accepted manuscript online 29 April 2026)

We investigate the ferromagnetic q-state Potts model on spherical Fibonacci graphs. These graphs are constructed by embedding quasi-uniform sites on a sphere and defining interactions via a chord-distance cutoff chosen so as to yield a network approximating four-neighbor connectivity. By combining Swendsen–Wang cluster Monte Carlo simulations with graph convolutional networks (GCNs), which operate directly on the adjacency structure and node spins, we develop a unified phase-classification framework applicable to both regular planar lattices and curved, irregular spherical graphs. Benchmarks on planar lattices demonstrate an efficient transfer strategy: after a fixed binarization of Potts spins into an effective Ising variable, a single GCN pre-trained on the Ising model can localize the transition region for different q values without retraining. Applying this strategy to spherical graphs, we find that curvature- and defect-induced connectivity irregularities induce only modest shifts in the inferred transition temperatures relative to their planar counterparts. Further analysis shows that the curvature-induced shift of the critical temperature is most pronounced at small q and diminishes rapidly as q increases. This trend is consistent with the physical picture that, in two dimensions, the Potts model undergoes a transition from a continuous phase transition to a weakly first-order one for $q > 4$, accompanied by a pronounced reduction in the correlation length.

Keywords: q-state Potts model, spherical fibonacci graphs, graph convolutional networks (GCNs), critical temperature

PACS: 05.50.+q, 64.60.Cn, 02.70.Uu, 07.05.Mh

DOI: 10.1088/1674-1056/ae6637 **CSTR:** 32038.14.CPB.ae6637

1. Introduction

In recent years, machine learning has developed into a powerful tool for identifying phase transitions directly from raw configurations, serving as an important complement to traditional thermodynamic approaches. Early pioneering studies demonstrated that neural networks can automatically detect qualitative changes in configurations near phase transitions without explicit knowledge of order parameters.[1–5] Building on these insights, subsequent works systematically extended this idea to a variety of classical and quantum models, establishing the applicability and stability of machine-learning methods across diverse physical systems.[6–9] These approaches were further generalized to systems in different spatial dimensions, demonstrating strong robustness and generalization with respect to dimensionality.[10–12] In addition, a series of studies investigated the robustness of machine-learning models against noise and finite-size effects, showing that reliable phase identification remains possible even under non-ideal data conditions.[13–15] Beyond qualitative classification, further research explored the quantitative extraction of critical points and effective critical exponents using machine-learning techniques, thereby elevating these methods to quantitative tools for analyzing phase transition behavior.[16–20]

Most early successes employed convolutional neural networks (CNNs), which assume image-like data on regular grids.[16,21,22] However, for curved surfaces and irregular lattices, the data are not arranged on a translationally invariant grid; instead, the notion of "nearest neighbors" is encoded by an interaction graph. This mismatch motivates the use of graph-based learning models. Graph convolutional networks operate directly on a graph adjacency structure and therefore provide a natural framework for treating both regular and irregular lattices within a common representation.[23–25]

Phase transitions on curved or irregular lattices also provide a setting to probe how geometry and connectivity shape collective phenomena beyond the standard Euclidean paradigm. In many contexts, such as spin degrees of freedom constrained on curved manifolds, or interaction networks without a regular embedding, the system is not a translationally invariant Bravais lattice, and "nearest neighbors" are not uniquely defined by a single bond length.[26–28] A natural approach is to treat the system as a statistical model on a graph, where geometric information enters primarily through connectivity.

Among discrete spin models, the ferromagnetic q-state

[†] These authors contributed equally to this work.
[‡] Corresponding author. E-mail: guohao.ph@seu.edu.cn
© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Potts model is a canonical extension of the Ising model, exhibiting rich critical behavior on two-dimensional lattices. On an infinite planar square lattice, the critical temperature is exactly known, $T_c^{2D}(q)/J = 1/\ln(1+\sqrt{q})$, with a continuous transition for $q \leq 4$ and a first-order transition for $q > 4$.[29] These exact planar results provide a well-established benchmark, making the Potts model an ideal testbed for assessing how curvature and irregular coordination modify critical properties on a spherical graph. The precise knowledge of $T_c^{2D}(q)$ offers a clear reference against which shifts induced by non-Euclidean geometry can be quantified.

A practical challenge is to discretize the sphere uniformly enough for a meaningful comparison with planar lattices while keeping the construction simple for arbitrary N. We adopt the spherical Fibonacci lattice[30–32] which provides quasi-uniform coverage through a closed-form construction. Any local interaction graph built on such a point set inevitably contains coordination defects. To emulate a "square-lattice-like" local environment, we define the interaction graph by a chord-distance cutoff r_c, chosen to maximize the number of four-neighbor nodes. This yields predominantly four-neighbor spherical graphs with a controlled fraction of three- and five-neighbor sites. This construction extends our recent spherical Fibonacci lattice Ising study[25] and provides a consistent baseline for generalizing to $q > 2$.

To sample equilibrium Potts configurations efficiently across temperatures, we use the Swendsen–Wang cluster Monte Carlo algorithm.[33–37] This method updates collective Fortuin–Kasteleyn (FK) clusters, mitigating critical slowing down near the transition. Beyond improving sampling efficiency, the FK representation offers an intuitive visualization of connectivity and domain formation on irregular graphs. For phase classification, we employ GCNs: for a fixed interaction graph, each configuration is represented by node spins and the adjacency matrix, and the GCN outputs confidence scores for the ordered and disordered phases. We adopt an "extreme-temperature" supervised training protocol, where configurations from very low and very high temperatures are labeled as ordered and disordered, respectively. The trained network is then evaluated across the full temperature range, and the transition region is identified where the two output confidences become comparable. To avoid training a separate network for every q, we introduce a simple binarization rule that maps the q-state Potts spins into an effective Ising variable. This allows a GCN pre-trained on the Ising model to be applied directly to binarized Potts configurations for different q, preserving the large-scale domain geometry relevant for phase classification. We first validate this transfer strategy on the planar square lattice using the exact $T_c^{2D}(q)$ as a benchmark, and then apply the same framework to spherical Potts models on Fibonacci graphs.

The main goals of this work are threefold:

(i) to construct and characterize nearly four-neighbor spherical Fibonacci graphs suitable for Potts simulations;

(ii) to develop a unified GCN-based phase classification framework applicable to both planar and spherical lattices, including an efficient transfer strategy across q based on binarization;

(iii) to quantify how curvature and local connectivity affect Potts critical temperatures on the sphere relative to planar counterparts.

By combining controlled graph construction, Swendsen–Wang sampling, and graph-based learning, we systematically compare $T_c(q)$ across geometry (plane vs. sphere), spin multiplicity, and local connectivity (tuned via r_c).

2. Spherical fibonacci lattice

We discretize a spherical surface of radius R by placing N sites using the spherical Fibonacci construction,[31] which provides a quasi-uniform point set. The Cartesian coordinates of the i-th site ($i = 1, 2, \ldots, N$) are

$$x_i = \sqrt{R^2 - z_i^2}\cos\theta_i, \quad y_i = \sqrt{R^2 - z_i^2}\sin\theta_i,$$
$$z_i = R\left(\frac{2i-1}{N} - 1\right), \quad \theta_i = 2\pi i\varphi, \tag{1}$$

where $\varphi = (\sqrt{5}-1)/2$ is the inverse golden ratio. Figure 1 shows the resulting Fibonacci sites for $N = 1000$ in two orthographic views (front and top).

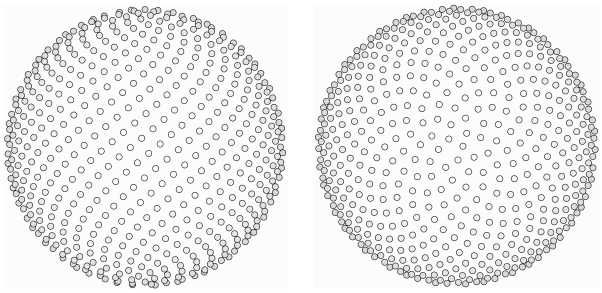


Fig. 1. Spherical Fibonacci sites for $N = 1000$ ($R = 10$), shown in front (left) and top (right) orthographic views.

To define short-range interactions on this irregular point set, we construct an interaction graph by introducing a cutoff radius r_c: two sites i and j are connected if their chord distance

$$d_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j| \tag{2}$$

satisfies $d_{ij} < r_c$. For a fixed spherical radius R and number of sites N, the cutoff radius r_c is chosen numerically by scanning candidate values and maximizing the number of four-coordinated sites in the resulting interaction graph. More explicitly, for each trial cutoff we compute the coordination number $k_i(r_c)$ of every site and evaluate

$$N_4(r_c) = \sum_i \delta_{k_i(r_c),4}, \tag{3}$$

then select the cutoff that maximizes $N_4(r_c)$. This criterion is motivated by the fact that the planar square lattice has coordination number four, so maximizing the number of four-neighbor sites yields the closest spherical analogue of a square-lattice-like local environment. For example, at $N = 1000$ and $R = 10$, this procedure gives $r_c/R = 0.1298$. For the larger graph used in the main q-dependent comparison, namely $N = 16384$ and $R = 10$, the same procedure yields $r_c \approx 0.3211$, corresponding to $r_c/R \approx 0.0321$.

In numerical simulations, once the sphere radius R and the total number of sites N are fixed, we typically choose an appropriate cutoff radius (r_c) (or equivalently the dimensionless ratio r_c/R) such that the resulting interaction graph is approximately four-neighbor. As a concrete example, for $N = 1000$ and $R = 10$, choosing $r_c/R = 0.1298$ (i.e., $r_c = 1.298$) yields a system in which the vast majority of sites are four-neighbor. The coordination statistics are $(N_3, N_4, N_5) = (76, 850, 74)$, corresponding to an average coordination number $\bar{z} = (3N_3 + 4N_4 + 5N_5)/N \simeq 3.998$. Under this construction, the lattice contains only a small fraction of three- and five-neighbor sites as unavoidable coordination defects, while remaining a close approximation to a four-neighbor lattice. Figure 2 shows the resulting connectivity from the same two views as in Fig. 1.

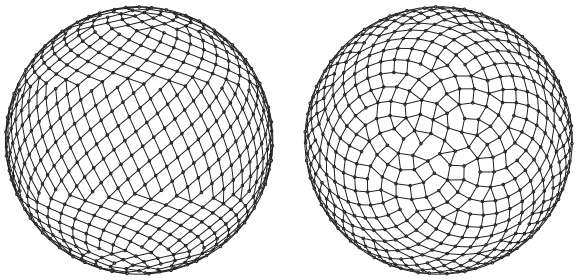


Fig. 2. Interaction graph constructed from the Fibonacci sites in Fig. 1 using $r_c/R = 0.1298$: edges connect pairs with $d_{ij} < r_c$ (front and top orthographic views).

3. The q-state Potts model on spherical Fibonacci lattice

3.1. Spin configurations

The q-state Potts model is a classical spin model in which each site i carries a discrete spin variable $\sigma_i \in \{0, 1, \ldots, q-1\}$. The Hamiltonian is defined as

$$H = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i, \sigma_j}, \tag{4}$$

where $\langle i, j \rangle$ runs over all interacting pairs with $d_{ij} < r_c$, $J > 0$ is the coupling constant, and δ is the Kronecker delta. The case $q = 2$ reduces to the Ising model. We measure the temperature T in units of J/k_B, and the corresponding Boltzmann weight is given by

$$P(\{\sigma_i\}) \propto \exp\left(\frac{J}{T} \sum_{\langle i,j \rangle} \delta_{\sigma_i, \sigma_j} \right). \tag{5}$$

For convenience, natural units are adopted in the following, with the Boltzmann constant set to $k_B = 1$. On the infinite planar square lattice, the critical temperature of the q-state Potts model is exactly $T_c^{2D}(q)/J = 1/\ln(1 + \sqrt{q})$,[29] with the transition being second-order for $q \leq 4$ and first-order for $q > 4$. In this work, we investigate the phase transition properties of the q-state Potts model on the spherical Fibonacci lattice.

We generate equilibrium configurations using the Swendsen–Wang cluster Monte Carlo algorithm.[33–37] Within the FK representation, each nearest-neighbor edge $\langle i, j \rangle$ is activated with probability $p_{add} = 1 - e^{-J/T}$ if $\sigma_i = \sigma_j$, and remains inactive if $\sigma_i \neq \sigma_j$. These activated edges connect like-spin sites into FK clusters. Each cluster is then collectively reassigned a new spin value, drawn uniformly from $\{0, 1, \ldots, q-1\}$. This cluster flip preserves detailed balance and significantly suppresses critical slowing down near T_c.

In numerical simulations, for each set of parameters (R, N, r_c) and q, Swendsen–Wang cluster simulations are performed over a common temperature range $T/J \in [0.01, 4.0]$, with a step of $\Delta T/J = 0.01$. At each temperature, we start from a random initial configuration, perform $N_{eq} = 2000$ equilibration sweeps, and then record $N_{samp} = 100$ configurations separated by an interval of 40 sweeps. The saved configurations are treated as approximately uncorrelated samples.

Both the equilibrium spin configurations and the corresponding activated FK bonds are recorded; the latter are used to visualize cluster connectivity (Fig. 3). These same configurations serve as input for the subsequent machine-learning analysis. Figure 3 illustrates the FK cluster geometry for $q = 3$ at $T/J = 1.0$, obtained on a spherical Fibonacci lattice with parameters $N = 1000$, $R = 10$, and $r_c/R = 0.1298$: the left panel displays the spin configuration, while the right panel overlays the activated FK bonds from one Swendsen–Wang update. Because bonds are activated exclusively between spins of the same value, the bond network directly outlines the FK clusters and provides a visualization of the correlation structure.

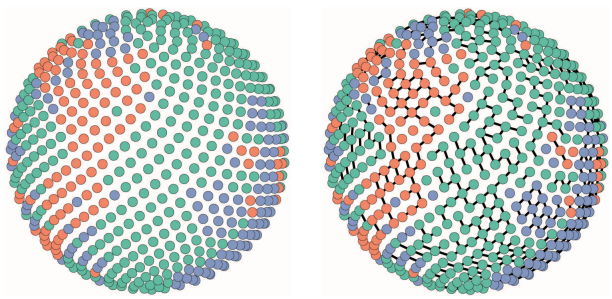


Fig. 3. Bond activation in the Swendsen–Wang algorithm ($N = 1000$, $R = 10$, and $r_c/R = 0.1298$). Left panel: A representative spin configuration at $T/J = 1.0$ for $q = 3$. Right panel: the same configuration overlaid with the activated FK bonds, which connect equal-spin neighbors into FK clusters.

3.2. Planar benchmarks and binarization-based GCN estimator

Before applying GCNs to the spherical Potts model, we first validate the method on the two-dimensional square lattice. A square lattice can be represented as a graph in which lattice sites are nodes and nearest-neighbor couplings are edges. We use the same GCN architecture as in our spherical Ising analysis:[25] the inputs are the spin configuration and the adjacency matrix, and the outputs are two components $(p_{\rm o}, p_{\rm d})$ representing the classification confidence for the ordered and disordered phases. Compared with convolutional neural networks applied to planar images, GCNs operate directly on graph structures and treat regular and irregular lattices within a unified framework.[23] Here, the term "unified framework" refers to the use of a common graph-based representation together with the same class of neural-network architecture for both regular and irregular lattices, rather than implying any universal superiority of GCN models over alternative machine-learning approaches.

Direct training on planar $q = 3$

We first train a GCN directly on the two-dimensional (2D) $q = 3$ Potts model on a square lattice of linear size $L = 128$. Using the Swendsen–Wang algorithm, we generate equilibrium configurations at extremely low temperatures $T/J \in [0.01, 0.05]$ (labeled as ordered with $(p_{\rm o}, p_{\rm d}) = (1, 0)$) and extremely high temperatures $T/J \in [100, 104]$ (labeled as disordered with $(p_{\rm o}, p_{\rm d}) = (0, 1)$). As node features, we use the raw Potts spin values $\sigma_i \in \{0, 1, 2\}$ encoded as a scalar per node. After training converges, we freeze the network and perform inference across $T/J \in [0.01, 4.0]$ to obtain $p_{\rm o}(T)$ and $p_{\rm d}(T)$.

Figure 4 displays the confidence curves obtained for $q = 3$. A monotonic decrease in $p_{\rm o}(T)$ and a concomitant increase in $p_{\rm d}(T)$ are observed with increasing temperature. At the resolution of our simulation ($\Delta T = 0.01J$), the curves intersect at $T/J \simeq 0.99$, in close agreement with the exact critical value $T_{\rm c}^{\rm 2D}(3)/J \approx 0.9950$. This benchmark validates that the GCN, trained exclusively on extreme-temperature configurations, can correctly identify the phase transition, thereby capturing the underlying change in domain structure.

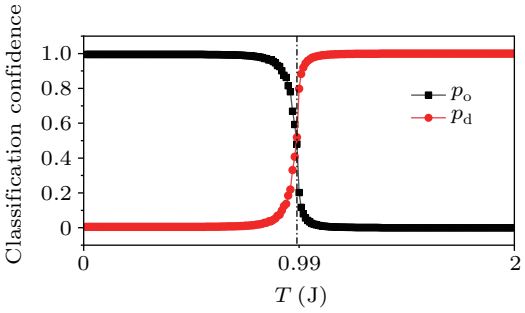


Fig. 4. GCN confidence for the 2D $q = 3$ Potts model ($L = 128$).

Binarization and transfer learning

Training a separate network for each q is computationally inefficient. Recent CNN studies suggest that, after an appropriate binarization of Potts spins into two groups, a network trained on the Ising model can transfer to the q-state Potts model without losing critical information.[38] Motivated by this, we adopt a binarization strategy in the GCN setting. For a given q, we partition the q spin states into two sets and define an effective Ising variable $s_i = \pm 1$ as

$$s_i = \begin{cases} -1, & \sigma_i \in \{0, 1, \ldots, \lfloor (q-1)/2 \rfloor\}, \\ +1, & \sigma_i \in \{\lfloor (q-1)/2 \rfloor + 1, \ldots, q-1\}. \end{cases} \tag{6}$$

For example, for $q = 3$, we choose $s_i = -1$ for $\sigma_i = 0$ and $s_i = +1$ for $\sigma_i = 1, 2$. This mapping coarsens the q-valued Potts field into a binary one while leaving the interaction graph unchanged. At the cluster level, it approximately preserves the large-scale geometry of domains associated with the chosen partition, specifically the interfaces separating the two groups, while discarding distinctions between Potts states within the same group.

This physical picture is consistent with the work of Fukushima and Sakai, who showed that an Ising-trained CNN can accurately detect the phase transition of the q-state Potts model after an analogous binarization procedure.[38] In particular, they demonstrated that under this transformation, the transition temperature and transition type remain invariant in the thermodynamic limit, while important geometric information of spin clusters is still preserved in the binarized configurations. By further analyzing the hidden-layer representations using principal component analysis, they found that away from criticality the network output mainly correlates with the internal energy in the high-temperature regime and with the magnetization (and possibly also the internal energy) in the low-temperature regime, whereas near the transition point more global features associated with the complexity of spin clusters become relevant. These results provide additional support for using binarization as an efficient transfer-learning bridge between the Ising and Potts models.

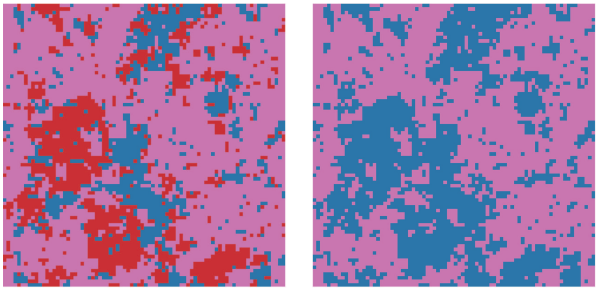


Fig. 5. Planar Potts configurations ($q = 3$) and their binarized versions at $T/J = 1.0$. Left panel: Potts spins $\{\sigma_i\}$; right panel: binarized Ising spins $\{s_i = \pm 1\}$ (see Eq. (6)).

Keeping the network architecture fixed, we recompute the confidence curves on the $q = 3$ dataset using the binarized s_i as inputs. As shown in Fig. 6, the crossing remains

at $T_c/J \approx 0.99$, indicating that binarization does not destroy the critical information carried by Potts cluster geometry. Together with the above physical arguments, this agreement suggests that the large-scale morphological information relevant to transition detection survives the binarization, even though the full microscopic q-state information is compressed. For a visual demonstration, Fig. 5 compares the original Potts configurations and their binarized counterparts at $T/J = 1.0$ for $q = 3$: large-scale cluster contours remain essentially unchanged.

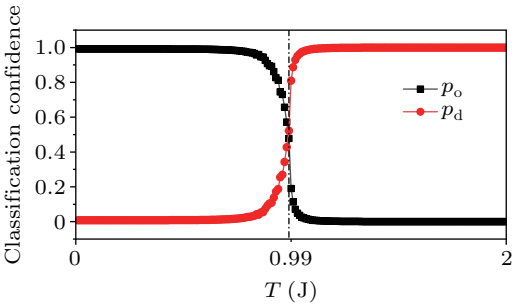


Fig. 6. GCN confidence curves for the 2D $q = 3$ Potts model ($L = 128$) after binarization, plotted versus temperature.

Based on these observations, we adopt a more economical strategy: we train a single GCN on the 2D Ising model and then apply it directly to binarized q-state Potts configurations. Specifically, we train on the $L = 128$ Ising model using configurations at $T/J \in [0.01, 0.05]$ (ordered) and $T/J \in [100, 104]$ (disordered), with Ising spins $s_i = \pm 1$ as node features. After training, all parameters are frozen. We then generate equilibrium configurations for planar Potts models with $q \in \{2, \ldots, 8, 10, 15\}$ over $T/J \in [0.5, 1.5]$, binarize each configuration, and feed it into this Ising-pretrained GCN to obtain $p_o(T)$ and $p_d(T)$.

Figure 7 presents the GCN confidence curves, $p_o(T)$ and $p_d(T)$, for the two-dimensional Potts model with $q \in \{2, \ldots, 8, 10, 15\}$. As temperature increases, $p_o(T)$ decreases monotonically, while $p_d(T)$ increases correspondingly. The crossing point where $p_o(T) \simeq p_d(T)$ shifts to lower temperatures as q grows – a trend that aligns with the exact decrease of $T_c^{2D}(q)$. Moreover, for $q > 4$, the transition region becomes noticeably sharper, reflecting the expected first-order character of the phase transition.

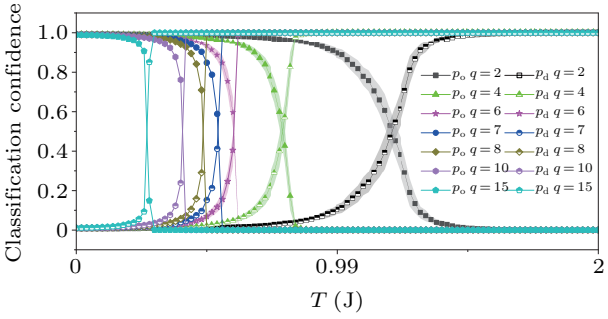


Fig. 7. The confidence curves $p_o(T)$ and $p_d(T)$ are obtained by applying an Ising-pretrained GCN to binarized 2D Potts configurations for $q \in \{2, \ldots, 8, 10, 15\}$ ($L = 128$) across the temperature range $T/J \in [0.5, 1.5]$.

Bootstrap estimation of T_c and uncertainties

To extract quantitative critical temperatures $T_c(q)$ from the confidence curves, one could fit the crossing point $p_o(T) = p_d(T)$ within a narrow window. However, for first-order transitions, the confidence curves can change extremely rapidly near the crossing, making conventional error propagation unstable. We therefore employ bootstrap resampling: [39–41] at each temperature we resample the classification outcomes with replacement to generate multiple pseudo-datasets, extract the crossing point for each, and collect the bootstrap samples $\{T_c^{(b)}(q)\}$. The mean and standard deviation of this set are then reported as the final estimate of $T_c(q)$ and its statistical uncertainty, respectively. Further implementation details are provided in Appendix C.

Using this bootstrap procedure, we obtain the planar baseline critical temperatures as $T_c^{(2D)}(q)/J = \{1.1038(8), 0.9877(18), 0.8958(4), 0.8439(5), 0.8016(1), 0.7715(2), 0.7429(1), 0.7034(1), 0.6354(1)\}$ for $q \in \{2, 3, \ldots, 8, 10, 15\}$ (uncertainties refer to the last two digits). Figure 8 compares these values with the exact results $T_c^{2D}(q)/J = 1/\ln(1+\sqrt{q})$, showing excellent agreement. The quoted uncertainties reflect only statistical fluctuations arising from finite sampling. The residual $\sim 1\%$ deviation is attributed to a small systematic bias inherent to the GCN-based estimator, which remains stable against variations in network hyperparameters (see Appendix).

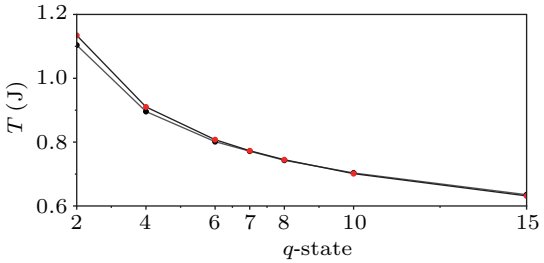


Fig. 8. $T_c(q)$ for the 2D Potts model ($L = 128$). Black: values obtained via an Ising-pretrained GCN with bootstrap – Red: exact results $T_c^{2D}(q)/J = 1/\ln(1+\sqrt{q})$.

The planar benchmarks establish that (i) GCNs can locate critical points of regular 2D Potts models accurately, and (ii) after binarization an Ising-pretrained GCN transfers across different q without retraining while retaining critical information. We therefore apply the same framework to the spherical Potts model to quantify how curvature and spherical topology shift T_c relative to planar baselines.

3.3. Spherical GCN analysis and geometric effects

Spherical transfer across q via an Ising-pretrained GCN. On the sphere, each configuration is treated as a scalar field defined on the spherical Fibonacci lattice, where nodes correspond to lattice sites, spins serve as node features, and edges connect site pairs separated by a chord distance smaller

than the cutoff r_c. We first train the GCN exclusively on the spherical Ising model ($q = 2$) with parameters $R = 10$, $N = 1000$, and $r_c/R = 0.1298$. Configurations from the low-temperature regime $T/J \in [0.01, 0.05]$ are labeled as ordered, while those from the high-temperature regime $T/J \in [100, 104]$ are labeled as disordered. Once trained, the network parameters are frozen. This frozen GCN is then applied directly to binarized spherical Potts configurations for $q > 2$.

Typical spherical Potts configurations for $q = 3$ at $T/J = 1.0$ and their binarized counterparts are displayed in Fig. 9. The large-scale cluster boundaries and characteristic length scales are preserved after binarization, whereas the original Potts states are compressed into two effective categories. This visual consistency under binarization provides a geometric justification for transferring the spherical Ising-pretrained GCN to multi-state Potts models on the sphere.

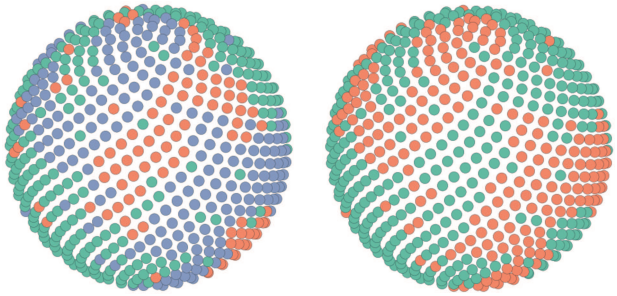


Fig. 9. Spherical Potts configurations and their binarized versions ($q = 3$; original – left, binarized – right) at $T/J = 1.0$.

Benchmark of transfer at $q = 3$. To validate the transfer strategy on the spherical graph, we first perform a benchmark at $q = 3$. The upper panel of Fig. 10 displays confidence curves obtained from a GCN trained directly on spherical $q = 3$ configurations, which give a crossing at $T_c/J \approx 0.98$. The lower panel illustrates our primary transfer approach: spherical $q = 3$ configurations are first binarized and then fed into the frozen GCN that was pretrained on the spherical Ising model. This procedure yields a crossing at $T_c/J \approx 1.01$. These two estimates agree within a few percent, confirming the validity of the transfer strategy based on binarization.

After completing the above validation, we apply the same spherical–Ising-pretrained GCN to the spherical Potts model with $q \in \{2, \ldots, 8, 10, 15\}$ on a fixed graph. To enable a direct comparison with the planar results for $T_c(q)$ of the two-dimensional Potts model at $L = 128$, the geometry is fixed as follows: $N = 128 \times 128 = 16384$, $R = 10$, and $r_c/R = 0.0321$. The coordination statistics for this geometry are $(N_3, N_4, N_5) = (210, 15658, 516)$, yielding an average coordination $\bar{z} \simeq 4.018$. Thus approximately 95.6% of nodes are four-coordinated, while the remaining $\sim 4.4\%$ are three- or five-coordinated defects.

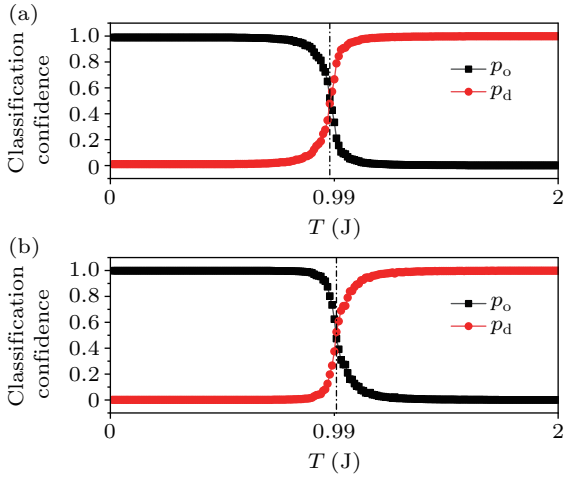


Fig. 10. Spherical $q = 3$ Potts GCN confidence curves: direct training – top, vs transfer from spherical-Ising pretraining on binarized data – bottom.

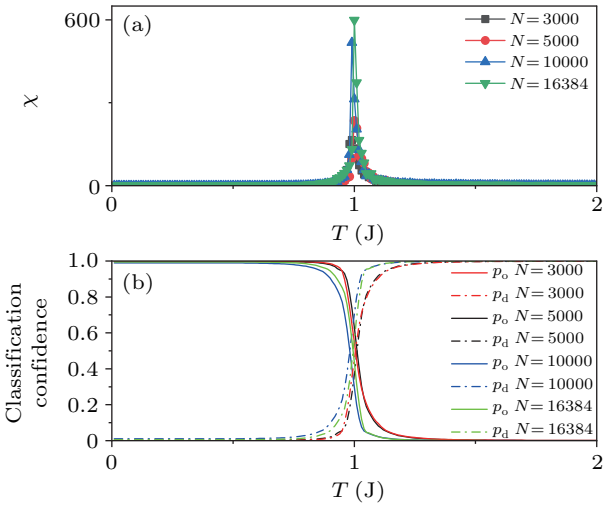


Fig. 11. Consistency check for the spherical $q = 3$ Potts model at different discretization sizes N. Magnetic susceptibility $\chi(T)$ – top – is shown for spherical Fibonacci graphs with $N = 3000$, 5000, 10000, and 16384, while the corresponding GCN confidence curves – bottom – are presented for the same set of graphs. The close agreement between susceptibility peak positions and GCN crossings indicates that the inferred transition temperature is robust against discretization refinement on the sphere.

Consistency check from susceptibility at different N. To provide a consistency check of the GCN-based identification using a conventional thermodynamic diagnostic, we further examine the magnetic susceptibility for the spherical $q = 3$ Potts model on a fixed sphere with different discretization sizes $N = 3000$, 5000, 10000, and 16384. The upper panel of Fig. 11 shows the corresponding susceptibility curves, while the lower panel presents the GCN confidence curves obtained for the same set of graphs. The transition temperatures extracted from the susceptibility peaks are $T_c^{(\chi)}/J = 0.99$, 1.00, 0.99, and 1.00, whereas the GCN crossings give $T_c^{(\mathrm{GCN})}/J = 1.00$, 1.01, 0.99, and 1.00, respectively. These two estimators agree within $0.01J$ for all considered N. Meanwhile, both the susceptibility peak positions and the GCN crossings display only a very weak dependence on N. This indicates that, for a sphere of fixed radius, increasing

the number of Fibonacci sites mainly refines the discretization of the curved surface rather than induces an appreciable shift of the transition temperature. This consistency check further supports the interpretation that the temperatures inferred from the GCN are not network-specific artifacts, but are quantitatively consistent with a standard thermodynamic diagnostic criterion.

Figure 12 (top) shows the resulting GCN confidence curves $p_{\rm o}(T)$ and $p_{\rm d}(T)$. The crossing temperature $p_{\rm o}(T) \simeq p_{\rm d}(T)$ shifts to lower temperatures as q increases, and the transition region becomes markedly sharper for $q > 4$, consistent with a first-order transition. Bootstrap analysis gives $T_{\rm c}^{\rm sph}(q)/J = \{1.1273(6), 0.9976(10), 0.9072(3), 0.8519(3), 0.8095(3), 0.7734(1), 0.7441(1), 0.7042(1), 0.6349(1)\}$ for $q \in \{2,\ldots,8,10,15\}$. For $q = 2$, the extracted $T_{\rm c}$ is consistent with our earlier spherical Fibonacci-lattice Ising study.[25]

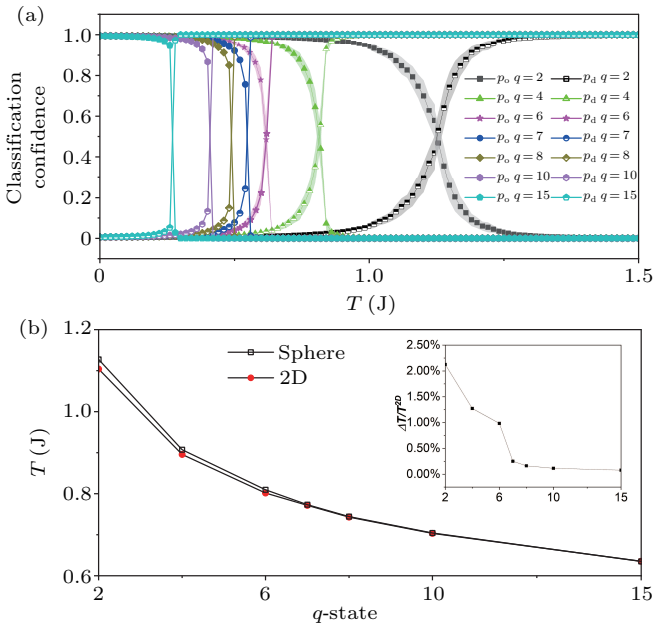


Fig. 12. Spherical Potts results compared with planar baselines. GCN confidence curves $p_{\rm o}(T)$ and $p_{\rm d}(T)$ – top – for binarized $q \in \{2,\ldots,8,10,15\}$ on the fixed spherical Fibonacci graph ($N = 128\times128 = 16384$, $R = 10$, $r_{\rm c}/R = 0.0321$). Critical temperatures versus q – bottom – show spherical $T_{\rm c}^{\rm sph}(q)$ (black) compared with planar $T_{\rm c}^{\rm 2D}(q)$ from the $L = 128$ square lattice (red). The inset displays the relative difference ΔT between spherical and planar critical temperatures, normalized by the planar $T_{\rm c}^{\rm 2D}$.

In contrast to the square lattice, where finite-size effects are often substantial, our earlier spherical Ising study indicated that increasing the number of Fibonacci sites N on a fixed sphere produces only a weak change in the inferred transition temperature under the same graph-construction protocol.[25] The additional $q = 3$ comparison shown in Fig. 11 is fully consistent with this picture: both the susceptibility-based and GCN-based estimates remain nearly unchanged over the range $N = 3000$–16384. Based on this, we forgo an explicit finite-size analysis for the spherical Potts models here. Instead, we adopt $N = 16384$ as a representative spherical discretization and use it to quantify the geometric shift relative to planar benchmarks.

Comparing the spherical estimates at $N = 16384$ with the planar baselines, $T_{\rm c}^{\rm (2D)}(q)/J = \{1.1038, 0.9877, 0.8958, 0.8439, 0.8016, 0.7715, 0.7429, 0.7034, 0.6354\}$ for $q \in \{2,\ldots,8,10,15\}$, we obtain relative shifts $\frac{T_{\rm c}^{\rm sph}-T_{\rm c}^{\rm (2D)}}{T_{\rm c}^{\rm (2D)}} \approx \{2.13\%, 1.00\%, 1.27\%, 0.95\%, 0.98\%, 0.25\%, 0.16\%, 0.11\%, -0.08\%\}$ for the same sequence of q. The spherical-planar difference therefore decreases rapidly with q: for $q < 6$, the difference is approximately 1%, whereas for $q \geq 7$ it drops sharply below 0.25%.

This behavior can be naturally understood in terms of the weakly first-order transition mechanism of the two-dimensional Potts model for $q > 4$. It is well established that $q = 4$ marks the boundary between continuous and first-order phase transitions. For $q > 4$, the transition is first order in the strict sense; however, when q is only slightly larger than 4 (typically $q = 5, 6$), the first-order character is extremely weak and accompanied by pronounced pseudocritical behavior. In this regime, the latent heat is very small and the correlation length near the transition becomes anomalously large, so that observables such as the specific heat and susceptibility exhibit an apparent near-divergent behavior characteristic of a continuous transition on finite length scales.[42–44] From a renormalization group perspective, this weakly first-order behavior originates from fixed-point annihilation and the associated walking dynamics, which generate a broad quasi-critical (crossover) regime at finite scales, while the system ultimately flows to a first-order transition at sufficiently large scales.[45]

Within this broad quasi-critical regime, weak geometric perturbations introduced by the spherical Fibonacci graph, including global curvature and a small fraction of coordination defects, can more readily modify fluctuations and cluster geometry. As a result, the apparent critical temperature $T_{\rm c}$ extracted from the classifier crossing is shifted at the percent level. As q increases, the first-order character strengthens and the correlation length decreases rapidly. The associated crossover window thus becomes significantly narrower, effectively reducing the sensitivity of the extracted $T_{\rm c}$ to geometric perturbations. This provides a natural explanation for why the spherical–planar difference drops rapidly to the 10^{-3} level for $q \geq 7$.

To provide a direct thermodynamic signature of the change in transition character for $q > 4$, we further examine the energy probability distributions near the transition region for $q = 4, 5, 6$, and 7. These calculations are performed on the same spherical Fibonacci graph used in the q-dependent comparison above, with system parameters fixed at $N = 16384$, $R = 10$, and $r_{\rm c}/R = 0.0321$. For each value of q, the energy histograms are obtained from Swendsen–Wang cluster Monte

Carlo simulations performed in the vicinity of the corresponding transition region. To reduce possible initialization bias and better resolve phase coexistence, simulations are carried out starting from both ordered (cold) and disordered (hot) initial states. After discarding the equilibration stage, the energy per spin is sampled from the subsequent Monte Carlo configurations to construct the probability distributions. The resulting histograms are shown in Fig. 13. For $q = 4$, a well-developed double-peak structure is not resolved within the present system size, although the distribution already exhibits noticeable broadening over an intermediate energy interval. This indicates that, for the present spherical graph and discretization level, the borderline $q = 4$ case remains strongly rounded and does not yet display a clearly separated bimodal structure. For $q = 5$, an incipient bimodal structure becomes visible, consistent with the expected weakly first-order behavior. For $q = 6$ and $q = 7$, a clear double-peak structure is observed, providing direct thermodynamic evidence of phase coexistence near the transition region and demonstrating that the first-order character becomes progressively stronger as q increases. These results further support the physical picture that the transition evolves from continuous behavior near $q = 4$ to increasingly first-order behavior for larger q, and are consistent with the sharpening of the GCN confidence crossings observed as q increases.

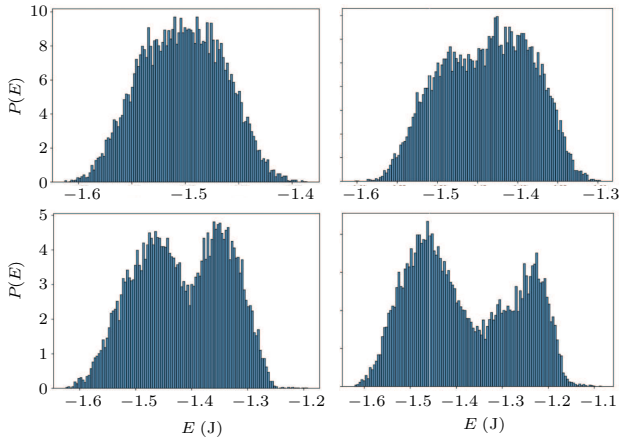


Fig. 13. Energy probability distributions near the transition region for the spherical Potts model on a fixed spherical Fibonacci graph with $N = 16384$, $R = 10$, and $r_c/R = 0.0321$, for $q = 4, 5, 6$, and 7. A bimodal structure becomes progressively more pronounced as q increases, indicating increasingly clear two-state coexistence and strengthening first-order character.

Effect of cutoff and average coordination at fixed N. We investigate how the local connectivity influences T_c by varying the cutoff r_c while keeping $R = 10$ and $N = 1000$ fixed. The resulting coordination distributions are listed in Table 1 for three choices: $r_c/R = 0.1298$, 0.1500, and 0.1700. For $r_c/R = 0.1298$, the graph is predominantly four-neighbor with an average coordination $\bar{z} \simeq 3.998$; for $r_c/R = 0.1500$, five- and six-neighbor sites dominate ($\bar{z} \simeq 5.71$); and for $r_c/R = 0.1700$, six- to eight-neighbor sites dominate ($\bar{z} \simeq 6.31$). Thus, tuning r_c allows interpolation between a sparse, nearly four-neighbor network and a densely connected graph on the same spherical surface.

Table 1. Coordination statistics for spherical Fibonacci graphs ($R = 10$, $N = 1000$) at different cutoffs r_c/R.

r_c/R	N_3 (w_3)	N_4 (w_4)	N_5 (w_5)	N_6 (w_6)	N_7 (w_7)	N_8 (w_8)
0.1298	76 (7.6%)	850 (85%)	74 (7.4%)	–	–	–
0.1500	–	26 (2.6%)	240 (24%)	734 (73.4%)	–	–
0.1700	–	–	8 (0.8%)	722 (72.2%)	224 (22.4%)	46 (4.6%)

The corresponding $T_c^{\mathrm{sph}}(q; r_c)$ values are listed in Table 2. For every q, T_c^{sph} increases significantly with r_c. Taking $q = 3$ as an example,

$$T_c^{\mathrm{sph}}(3; r_c/R = 0.1298, 0.1500, 0.1700)/J \tag{7}$$

$$\simeq (1.0085, 1.5254, 1.7058), \tag{8}$$

indicates an increase by nearly a factor of two as the connectivity changes from sparse to dense. This trend holds for all q: as the average coordination $\bar{z}$ increases from ~ 4.0 to ~ 6.3, T_c^{sph} is systematically shifted upward, while the monotonic decrease of T_c with q is preserved. These findings indicate that variations in local connectivity primarily rescale the overall magnitude of T_c, whereas the spin multiplicity q governs the transition order and the relative q-dependence of the critical temperature.

Table 2. Spherical Potts critical temperatures $T_c^{\mathrm{sph}}(q; r_c)$ at $N = 1000$ for different r_c/R.

q	$r_c/R = 0.1298$	$r_c/R = 0.1500$	$r_c/R = 0.1700$
2	1.1564(16)	1.7544(33)	1.9630(47)
3	1.0085(15)	1.5254(33)	1.7058(32)
4	0.9081(8)	1.3665(15)	1.5191(20)
5	0.8496(8)	1.2722(26)	1.4178(11)
6	0.8048(2)	1.2042(4)	1.3296(12)

This observation also helps clarify the respective roles of local graph structure and global topology in the present problem. Although the sphere has a nonzero Euler characteristic, the Potts transition considered here is not driven by topological defects and is therefore not directly constrained by global topology. This situation differs from the spherical XY model, where spins form a tangent vector field on S^2 and the Poincaré–Hopf theorem enforces a nonzero net topological charge, so that topology directly enters the transition physics (see Ref. [23]). By contrast, in the present Potts case, the influence of topology is indirect and mainly manifests through graph-structural constraints such as unavoidable coordination defects and fluctuations in the coordination number on the sphere. Accordingly, the observed spherical–planar

shifts of T_c are primarily associated with graph-theoretic quantities such as the coordination-defect density and the average coordination number $\bar{z}$, rather than global topological constraints themselves.

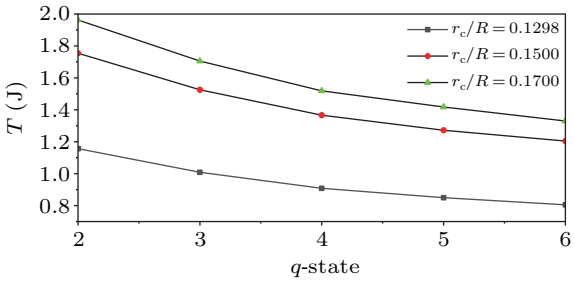


Fig. 14. Critical temperatures $T_c^{\mathrm{sph}}(q; r_c)$ versus q for spherical Potts models on a fixed graph ($R = 10$, $N = 1000$) with three different cutoff ratios: $r_c/R = 0.1298$, 0.1500, and 0.1700. Increasing r_c enhances local connectivity and systematically shifts T_c upward, while the monotonic decrease of T_c with q is preserved.

In summary, these results demonstrate that a spherical-Ising-pretrained GCN reliably transfers to binarized spherical Potts models, producing critical temperatures that closely match planar benchmarks. For locally four-neighbor graphs, the critical temperatures on the spherical Fibonacci lattice already show good agreement with planar results, with curvature and coordination defects introducing only percent-level shifts. The rapid reduction of the spherical–planar difference with increasing q can be naturally understood in terms of the weakly first-order nature of the Potts transition for $q > 4$: when q is only slightly above 4, the broad quasi-critical regime enhances sensitivity to weak geometric perturbations, whereas at larger q the strengthened first-order character significantly narrows the crossover window, effectively reducing the sensitivity of the extracted T_c to geometric effects.

At fixed N, varying r_c primarily rescales the overall transition temperature for all q, while leaving the functional form of $T_c(q)$ largely unchanged. Together, these observations support a clear separation between geometric influences (curvature and connectivity) and the intrinsic role of spin multiplicity q, all captured within a unified GCN-based framework.

4. Conclusion

We have studied the ferromagnetic q-state Potts model on spherical Fibonacci graphs, quantifying how curvature and irregular coordination modify phase transitions relative to planar benchmarks. The spherical graphs are constructed using a chord-distance cutoff that maximizes local four-coordination, resulting in predominantly four-neighbor networks with a controlled fraction of coordination defects. Equilibrium configurations are generated via the Swendsen–Wang cluster algorithm, with Fortuin–Kasteleyn bonds recorded to visualize connectivity.

Critical temperatures are identified using graph convolutional networks (GCNs) that operate directly on the graph adjacency matrix and node spins. After validating the method on planar lattices, we introduce a transfer strategy based on spin binarization, which allows an Ising-pretrained GCN to be applied across different q without retraining. Bootstrap resampling provides robust T_c estimates even near first-order transitions. For the spherical $q = 3$ case, an additional comparison with susceptibility peaks at different N shows agreement with the GCN-based transition temperatures within $0.01J$, further supporting robustness against discretization refinement on the sphere.

For nearly four-neighbor spherical graphs, the curvature-induced shift in T_c is most pronounced at small q (about 1–2% for $q < 7$) and decreases to about 0.2% or less for $q \geq 7$, consistent with the weakly first-order character of the Potts transition for $q > 4$. This implies enhanced geometric sensitivity at small q and rapid suppression of such effects as q increases. At fixed N, increasing the cutoff r_c (and thus the average coordination $\bar{z}$) primarily shifts T_c upward for all q, while preserving the monotonic q-dependence and the order of the transition.

These results highlight a clear separation between geometric effects, namely curvature and connectivity, and the role of spin multiplicity q. More broadly, this work demonstrates that GCNs provide a unified framework for phase classification on both regular and curved graphs, using the same input representation of adjacency and spins. The combination of spherical Fibonacci discretization, cluster Monte Carlo sampling, and binarization-based transfer learning offers a scalable approach for studying multi-state models on non-Euclidean geometries. This framework can be extended to other curved manifolds, alternative graph constructions, or higher-q regimes where first-order behavior becomes more pronounced. Code and graph generators are available upon reasonable request.

Appendix A: Swendsen–Wang cluster update for the spherical Potts model

We generate equilibrium configurations for the ferromagnetic q-state Potts model on spherical Fibonacci graphs using the Swendsen–Wang cluster Monte Carlo algorithm. This section outlines one SW sweep on an irregular interaction graph.

Interaction graph The system is defined on a graph $G = (V, E)$ embedded on a sphere of radius R, where $|V| = N$ sites are placed at Fibonacci coordinates $\{\boldsymbol{r}_i\}_{i=1}^{N}$. Two sites i and j are connected by an edge if their chord distance

$$d_{ij} \equiv |\boldsymbol{r}_i - \boldsymbol{r}_j|, \tag{A1}$$

satisfies $d_{ij} < r_c$. We denote interacting pairs by $\langle i, j\rangle \in E$.

The Potts Hamiltonian on this graph is

$$H(\{\sigma\}) = -J\sum_{\langle i,j\rangle}\delta_{\sigma_i,\sigma_j},\ \ \sigma_i \in \{0,1,\ldots,q-1\},\ \ J>0, \quad \text{(A2)}$$

with inverse temperature $\beta \equiv 1/T$ (setting $k_{\mathrm{B}} = 1$).

Fortuin–Kasteleyn bond activation The SW algorithm employs auxiliary bond variables in the FK representation. Given a spin configuration $\{\sigma_i\}$, a bond is activated on edge $\langle i,j\rangle$ only if the two spins are equal

$$b_{ij} = \begin{cases} 1, & \text{with probability } p_{\mathrm{add}} = 1-\mathrm{e}^{-\beta J}, \text{ if } \sigma_i = \sigma_j, \\ 0, & \text{if } \sigma_i \neq \sigma_j. \end{cases} \quad \text{(A3)}$$

Thus, bonds appear exclusively within regions of like spins.

Cluster identification and recoloring (One SW sweep) Let $E_{\mathrm{act}} = \{\langle i,j\rangle \in E : b_{ij} = 1\}$ be the set of activated bonds. The subgraph (V, E_{act}) decomposes into connected components (FK clusters) $\{C_\alpha\}$.

A single SW sweep proceeds as follows:

(i) **Bond construction** Sample all b_{ij} according to Eq. (A3).

(ii) **Cluster recoloring** For each FK cluster C_α, choose a new Potts state $\sigma'_\alpha \in \{0,1,\ldots,q-1\}$ uniformly at random and assign

$$\sigma_i \leftarrow \sigma'_\alpha,\ \ \forall\ i \in C_\alpha. \quad \text{(A4)}$$

By reassigning entire clusters collectively, the update performs nonlocal moves that efficiently decorrelate long-wavelength fluctuations near criticality.

Detailed balance and efficiency The SW update can be viewed as alternating conditional sampling in the joint FK spin–bond ensemble: sample bonds conditioned on spins, then sample spins conditioned on bonds. This construction ensures the Markov chain satisfies detailed balance and has the correct Boltzmann distribution as its stationary distribution. Compared to local single-spin updates, cluster flips substantially mitigate critical slowing down by updating strongly correlated regions as a unit.

Appendix B: GCN architecture and training details

We treat each spin configuration as a graph signal defined on the interaction graph. The graph convolutional network (GCN) takes the graph connectivity and node spins as input and outputs two confidence scores for the ordered and disordered phases.

Interaction graph and matrices For a spherical Fibonacci point set $\{\boldsymbol{r}_i\}_{i=1}^N$ on a sphere of radius R, we construct an unweighted, undirected interaction graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1,\ldots,N\}$. Two vertices i and j are connected if their chord distance $d_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ is smaller than a cutoff radius r_{c}. The adjacency matrix $\boldsymbol{A} \in \{0,1\}^{N\times N}$ is defined by

$$A_{ij} = \begin{cases} 1, & i \neq j \text{ and } d_{ij} < r_{\mathrm{c}}, \\ 0, & \text{otherwise}, \end{cases} \quad \text{(B1)}$$

and the degree matrix is $\boldsymbol{D} = \mathrm{diag}(d_1,\ldots,d_N)$ with $d_i = \sum_j A_{ij}$. We use the random-walk normalized Laplacian

$$\boldsymbol{L}^{\mathrm{rw}} = \boldsymbol{D}^{-1}(\boldsymbol{D} - \boldsymbol{A}). \quad \text{(B2)}$$

Only $\boldsymbol{A}$ (hence $\boldsymbol{D}$ and $\boldsymbol{L}^{\mathrm{rw}}$) is used by the GCN; the explicit coordinates $\{\boldsymbol{r}_i\}$ are used only to construct the connectivity.

Node features For each Monte Carlo configuration, we form a node-feature matrix $\boldsymbol{X} \in \mathbb{R}^{N\times F}$. For the Ising model, we use $F = 1$ and $X_i = s_i \in \{-1,+1\}$. For the Potts model, we either use the raw Potts labels $X_i = \sigma_i$ (for direct training at a fixed q) or the binarized variable $X_i = s_i$ defined below (for transfer learning across q).

Graph convolution layer A single graph-convolution layer is implemented as

$$\boldsymbol{H} = \mathrm{ReLU}(\boldsymbol{L}^{\mathrm{rw}}\boldsymbol{X}\,\boldsymbol{W}_{\mathrm{h}} + \boldsymbol{b}_{\mathrm{h}}), \quad \text{(B3)}$$

where $\boldsymbol{W}_{\mathrm{h}} \in \mathbb{R}^{F\times F_{\mathrm{h}}}$ is the trainable weight matrix, $\boldsymbol{b}_{\mathrm{h}} \in \mathbb{R}^{F_{\mathrm{h}}}$ is a trainable bias (broadcast to all nodes), $\mathrm{ReLU}(x) = \max(x,0)$ is applied elementwise, and $\boldsymbol{H} \in \mathbb{R}^{N\times F_{\mathrm{h}}}$ is the hidden node embedding.

Graph-level readout and confidence output To classify an entire configuration, we aggregate node embeddings by global average pooling

$$\boldsymbol{h}_{\mathcal{G}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{H}_i \in \mathbb{R}^{F_{\mathrm{h}}}, \quad \text{(B4)}$$

followed by a fully connected layer producing two logits,

$$\boldsymbol{z} = \boldsymbol{W}_{\mathrm{o}}\,\boldsymbol{h}_{\mathcal{G}} + \boldsymbol{b}_{\mathrm{o}} \in \mathbb{R}^2, \quad \text{(B5)}$$

and a softmax output

$$(p_{\mathrm{o}}, p_{\mathrm{d}}) = \mathrm{softmax}(\boldsymbol{z}),\ \ \mathrm{softmax}(z_k) = \frac{\mathrm{e}^{z_k}}{\sum_{\ell=1}^{2}\mathrm{e}^{z_\ell}}. \quad \text{(B6)}$$

Here p_{o} and p_{d} are interpreted as the confidences for the ordered and disordered phases, respectively.

Supervised training at extreme temperatures The network is trained in a supervised manner using configurations generated at extreme temperatures: low-T configurations are labeled as ordered $(1,0)$, and high-T configurations as disordered $(0,1)$. Denoting the ground-truth label by $\boldsymbol{y} \in \{(1,0),(0,1)\}$, we minimize the cross-entropy loss

$$\mathcal{L} = -\sum_{k\in\{o,d\}} y_k \ln p_k. \quad \text{(B7)}$$

After training converges, the network parameters are frozen. Inference is then performed on configurations sampled across the target temperature range, producing the confidence curves $p_{\mathrm{o}}(T)$ and $p_{\mathrm{d}}(T)$.

Binarization (Potts → Ising) for transfer across q To use a single Ising-pretrained classifier for different q, we map the Potts spin $\sigma_i \in \{0,1,\ldots,q-1\}$ to a binary variable $s_i = \pm 1$ via

$$s_i = \begin{cases} -1, & \sigma_i \in \{0,1,\ldots,\lfloor (q-1)/2 \rfloor\}, \\ +1, & \sigma_i \in \{\lfloor (q-1)/2 \rfloor + 1,\ldots,q-1\}. \end{cases} \tag{B8}$$

This rule is unambiguous for both even and odd q. In the transfer setting, we apply Eq. (B8) to each Potts configuration to obtain $\{s_i\}$, and feed $\boldsymbol{X} = (s_1,\ldots,s_N)^{\mathrm{T}}$ into the frozen Ising-pretrained GCN on the same interaction graph.

Implementation details and Hyperparameters In our implementation, the interaction graph is fixed for each dataset, and the corresponding adjacency matrix is read in advance. The graph operator used in the network is the random-walk normalized Laplacian

$$\boldsymbol{L}^{\mathrm{rw}} = \boldsymbol{D}^{-1}(\boldsymbol{D} - \boldsymbol{A}), \tag{B9}$$

which is constructed directly from the adjacency matrix. The graph-convolution step acts on the node-feature matrix through this operator, followed by a ReLU activation. The resulting hidden representation is then passed to a fully connected readout layer that outputs two logits corresponding to the ordered and disordered phases, and a softmax function is applied to obtain the classification confidences $(p_{\mathrm{o}}, p_{\mathrm{d}})$.

For the direct-training example of the planar $q = 3$ Potts model, the raw Potts spin values $\sigma_i \in \{0,1,2\}$ are used directly as node features. The training samples are generated from equilibrium configurations in two extreme temperature windows: low-temperature configurations are labeled as ordered and high-temperature configurations as disordered. After training converges, the network parameters are fixed and the trained model is applied to configurations across the full temperature range to obtain the confidence curves $p_{\mathrm{o}}(T)$ and $p_{\mathrm{d}}(T)$.

The network is trained using the cross-entropy loss function and the stochastic gradient descent (SGD) optimizer with learning rate 0.01. In the implementation used here, the batch size used in training is 100, and the validation step is carried out separately to monitor convergence and select the model with the lowest validation loss. The total number of training epochs is 5000. These settings are kept fixed throughout the training procedure unless otherwise stated.

Dataset organization For the spherical Potts classifications discussed in this work, it is useful to distinguish two types of datasets. The first is the supervised-learning dataset, consisting of training, validation, and supervised test sets, which is used for parameter optimization, model selection, and classification-performance checking. The second is the full-temperature inference dataset, which is used only after training to evaluate the confidence curves over the entire temperature range and to extract the characteristic temperature T_{c} from the crossing of the ordered and disordered confidences. For the representative spherical Potts cases considered here, the supervised dataset contains 1000 training samples, 1000 validation samples, and 100 supervised test samples. In addition, the full-temperature inference dataset covers the range $T/J \in [0.01, 4.00]$ with step $\Delta T/J = 0.01$, and each temperature point contains 100 configurations, so that the total number of inference samples is

$$N_{\mathrm{infer}} = 400 \times 100 = 40000. \tag{B10}$$

This all-temperature inference dataset is not part of the supervised test split; rather, it is used exclusively to construct the averaged confidence curves and to determine the crossing temperature.

Convergence of the GCN estimator The physically relevant quantity in the present machine-learning analysis is not the loss function itself, but the characteristic temperature T_{c} extracted from the crossing of the ordered and disordered confidence curves. For this reason, the convergence test should be formulated at the level of T_{c} rather than solely at the level of the validation loss.

For a given checkpoint during training, we apply the network to the full-temperature inference dataset and compute the averaged confidence curves $p_{\mathrm{o}}(T)$ and $p_{\mathrm{d}}(T)$ at each temperature. The characteristic temperature is then defined by the first crossing of these two curves. In practice, we locate the first interval where $p_{\mathrm{o}}(T) - p_{\mathrm{d}}(T)$ changes sign and determine T_{c} by linear interpolation within that interval. This is precisely the same criterion as the one used in the main text.

Since the machine-learning analyses throughout the paper employ the same or closely related network structure, training protocol, and model-selection criterion, it is unnecessary to show the full training history for every confidence curve presented in the main text. Instead, we select representative cases to demonstrate the convergence of the estimator. Here we consider the spherical Potts models with $q = 3$ and $q = 5$, which serve as representative examples on the continuous-transition side and the weakly first-order side, respectively.

The corresponding results are shown in Fig. B1. In both cases, the validation loss decreases steadily during training, while the extracted characteristic temperature T_{c} may exhibit small fluctuations at early stages but rapidly converges and remains stable once the validation loss reaches a sufficiently low level. This behavior indicates that the GCN-based estimate of the transition temperature is stable after training convergence.

These representative convergence tests further demonstrate that, within the unified GCN framework adopted here, the characteristic temperature extracted from the confidence-curve crossing is a stable and reproducible estimator. Accordingly, the confidence-curve results reported in the main text

are supported by explicit numerical convergence in addition to their physical consistency with thermodynamic analysis.

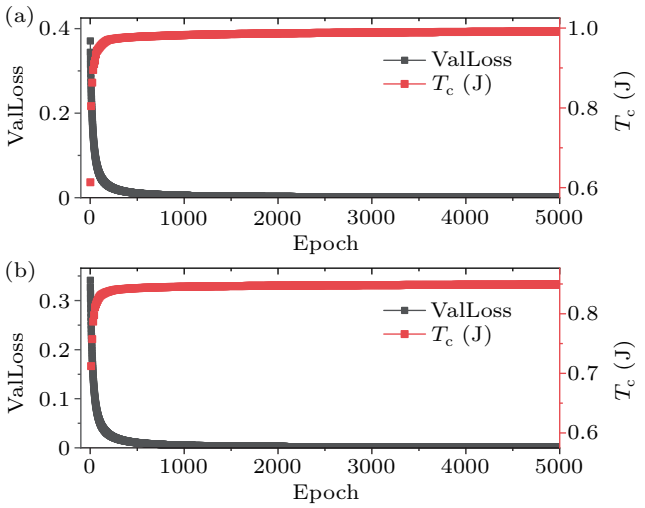


Fig. B1. Representative convergence tests of the GCN estimator for the spherical Potts model. The upper panel shows the case $q = 3$, and the lower panel shows the case $q = 5$. In each panel, the horizontal axis is the training epoch, the left vertical axis gives the validation loss, and the right vertical axis gives the characteristic temperature T_c extracted from the confidence-curve crossing. Once the validation loss reaches a sufficiently low range, the inferred T_c rapidly converges and remains stable.

Estimating T_c from confidence curves For each q, we identify the transition region as the temperature where $p_o(T) \simeq p_d(T)$. On a discrete temperature grid, this corresponds to the interval where $p_o(T) - p_d(T)$ changes sign. The critical temperature T_c is then extracted by interpolation or the bootstrap procedure described in the main text.

Appendix C: Bootstrap estimation of T_c and systematic bias discussion

For each temperature T_i, the trained GCN outputs a two-component logit vector $\boldsymbol{o} = (o_o, o_d)$ per configuration. We convert logits to classification confidences via the softmax function

$$p_o = \frac{e^{o_o}}{e^{o_o} + e^{o_d}}, \quad p_d = \frac{e^{o_d}}{e^{o_o} + e^{o_d}}, \quad (p_o + p_d = 1). \tag{C1}$$

The curves $p_o(T)$ and $p_d(T)$ represent the ordered- and disordered-phase confidences, respectively. Since $p_o + p_d = 1$, the criterion $p_o = p_d$ is equivalent to $p_o = 0.5$.

Temperature grouping and sample mean Data are grouped by temperature: for each T_i, we have N_{samp} equilibrium configurations ($N_{\text{samp}} = 100$ in our implementation). Let $\{p_{o,i}^{(k)}\}_{k=1}^{N_{\text{samp}}}$ denote the ordered-confidence values predicted at T_i. The empirical mean confidence is

$$\widehat{p}_o(T_i) = \frac{1}{N_{\text{samp}}} \sum_{k=1}^{N_{\text{samp}}} p_{o,i}^{(k)}. \tag{C2}$$

Bootstrap resampling To estimate the statistical uncertainty of T_c arising from finite sampling at each temperature, we apply nonparametric bootstrap independently at every T_i. For the b-th bootstrap replica, we resample with replacement

$$\{p_{o,i}^{(k)}\}_{k=1}^{N_{\text{samp}}} \longrightarrow \{p_{o,i}^{*(k)}\}_{k=1}^{N_{\text{samp}}}, \tag{C3}$$

and compute the resampled mean

$$\widehat{p}_o^{(b)}(T_i) = \frac{1}{N_{\text{samp}}} \sum_{k=1}^{N_{\text{samp}}} p_{o,i}^{*(k)}. \tag{C4}$$

Repeating for all i yields one bootstrap realization of the confidence curve $\widehat{p}_o^{(b)}(T)$.

Extracting T_c from each bootstrap curve For each bootstrap curve, we locate the transition as the first temperature interval where $\widehat{p}_o^{(b)}(T)$ crosses 0.5. We define $\Delta_i^{(b)} = \widehat{p}_o^{(b)}(T_i) - 0.5$.. A crossing occurs between T_i and T_{i+1} if $\Delta_i^{(b)} \Delta_{i+1}^{(b)} < 0$. If multiple crossings appear due to noise, we take the lowest-temperature one (the first sign change) to define $T_c^{(b)}$.

Given a crossing bracket (T_i, T_{i+1}), we estimate $T_c^{(b)}$ by linear interpolation

$$T_c^{(b)} = T_i + \frac{0.5 - \widehat{p}_o^{(b)}(T_i)}{\widehat{p}_o^{(b)}(T_{i+1}) - \widehat{p}_o^{(b)}(T_i)} (T_{i+1} - T_i). \tag{C5}$$

Bootstrap replicas showing no crossing over the scanned temperature window are discarded.

Final estimate and uncertainty Collecting all valid bootstrap estimates $\{T_c^{(b)}\}_{b=1}^{N_{\text{boot}}}$ (with $N_{\text{boot}} = 2000$), we report

$$\overline{T_c} = \frac{1}{N_{\text{eff}}} \sum_{b=1}^{N_{\text{eff}}} T_c^{(b)}, \quad \delta T_c = \sqrt{\frac{1}{N_{\text{eff}} - 1} \sum_{b=1}^{N_{\text{eff}}} \left(T_c^{(b)} - \overline{T_c}\right)^2}, \tag{C6}$$

where N_{eff} is the number of bootstrap replicas exhibiting a crossing. The uncertainty δT_c quantifies statistical fluctuations due to finite configuration sampling at each temperature; it does not include systematic effects from the finite temperature step ΔT, finite-size corrections, or model/architecture dependence.

Acknowledgments

Project supported by the Innovation Program for Quantum Science and Technology-National Science and Technology Major Project (Grant No. 2021ZD0301904), the National Natural Science Foundation of China (Grant No. 12447216), and the National Natural Science Foundation of China (Grant No. 12405008).

References

[1] Carleo G and Troyer M 2017 *Science* **355** 602
[2] Carrasquilla J and Melko R G 2017 *Nat. Phys.* **13** 431
[3] Deng D L, Li X and Das Sarma S 2017 *Phys. Rev. X* **7** 021021
[4] Zhang P, Shen H and Zhai H 2018 *Phys. Rev. Lett.* **120** 066401
[5] Ch'ng K, Carrasquilla J, Melko R G and Khatami E 2017 *Phys. Rev. X* **7** 031038
[6] Li C D, Tan D R and Jiang F J 2018 *Ann. Phys.* **391** 312

[7] Kim D and Kim D H 2018 *Phys. Rev. E* **98** 022138
[8] Wang L 2016 *Phys. Rev. B* **94** 195105
[9] Li Z, Luo M and Wan X 2019 *Phys. Rev. B* **99** 075418
[10] Bachtis D, Aarts G and Lucini B 2020 *Phys. Rev. E* **102** 053306
[11] Canabarro A, Fanchini F F, Malvezzi A L, Pereira R and Chaves R 2019 *Phys. Rev. B* **100** 045129
[12] D'Angelo F and Böttcher L 2020 *Phys. Rev. Res.* **2** 023266
[13] Wetzel S J 2017 *Phys. Rev. E* **96** 022140
[14] Mills K and Tamblyn I 2018 *Phys. Rev. E* **97** 032119
[15] Morningstar A and Melko R G 2018 *J. Mach. Learn. Res.* **18** 1
[16] Tanaka A and Tomiya A 2017 *J. Phys. Soc. Jpn.* **86** 063001
[17] van Nieuwenburg E P L, Liu Y H and Huber S D 2017 *Nat. Phys.* **13** 435
[18] Alexandrou C, Athenodorou A, Chrysostomou C and Paul S 2020 *Eur. Phys. J. B* **93** 113
[19] Corte I, Acevedo S, Arlego M and Lamas C A 2021 *Comput. Mater. Sci.* **198** 110702
[20] Yau H M and Su N 2022 *SciPost Phys. Core* **5** 032
[21] Beach M J S, Golubeva A and Melko R G 2018 *Phys. Rev. B* **97** 045207
[22] Zhang W, Liu J and Wei T C 2019 *Phys. Rev. E* **99** 032142
[23] Song C H, Gao Q C, Hou X Y, Wang X, Zhou Z, He Y, Guo H and Chien C C 2022 *Phys. Rev. Res.* **4** 023005
[24] Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, Wang L, Li C and Sun M 2020 *AI Open* **1** 57
[25] Zhou Z, Song C H, Hou X Y and Guo H 2025 *New J. Phys.* **27** 124601
[26] Kleman M 1989 *Adv. Phys.* **38** 605
[27] Janke W and Villanova R 2002 *Phys. Rev. B* **66** 134208
[28] Giomi L 2009 *Unordinary order: A Theoretical, Computational and Experimental Investigation of Crystalline Order in Curved Space* (Syracuse University)
[29] Wu F 1982 *Rev. Mod. Phys.* **54** 235
[30] Gonzalez A 2010 *Math. Geosci.* **42** 49
[31] Hannay J H and Nye J F 2004 *J. Phys. A: Math. Gen.* **37** 11591
[32] Keinert B, Innmann M, Saenger M and Stamminger M 2015 *ACM Trans. Graph.* **34** 1
[33] Edwards R and Sokal A 1988 *Phys. Rev. D* **38** 2009
[34] Galanis A, Stefankovic D and Vigoda E 2019 *Random Struct. Algorithms* **54** 82
[35] Barbu A and Zhu S C 2005 *IEEE Trans. Pattern Anal. Mach. Intell.* **27** 1239
[36] Blanca A and Gheissari R 2023 *Proc. IEEE 64th Annu. Symp. Found. Comput. Sci.* (Santa Cruz, CA) p. 2006
[37] Ivaneyko D, Ilnytskyi J, Berche B and Holovatch Y 2005 *Condens. Matter Phys.* **8** 149
[38] Fukushima K and Sakai K 2021 *Prog. Theor. Exp. Phys.* **2021** 043
[39] Efron B 1992 *Breakthroughs in Statistics: Methodology and Distribution* (New York: Springer) pp. 569
[40] Kawano H and Higuchi T 1995 *Geophys. Res. Lett.* **22** 307
[41] Lee C Y 2022 *Size Scaling in Ising Model* (Ph.D. thesis)
[42] Binder K 1981 *J. Stat. Phys.* **24** 69
[43] Buddenoir E and Wallon S 1993 *J. Phys. A: Math. Gen.* **26** 3045
[44] Iino S, Morita S, Kawashima N and Sandvik A W 2019 *J. Phys. Soc. Jpn.* **88** 034006
[45] Gorbenko V, Rychkov S and Zan B 2018 *SciPost Phys.* **5** 050