

# Energy-Weighted Site Percolation in Two Dimensions

Sayan Sircar\* and Kabir Ramola†

Tata Institute of Fundamental Research, Hyderabad - 500046, India

We study a generalization of two-dimensional site percolation by assigning an energy cost  $\varepsilon$  to bonds between nearest-neighbor occupied sites. This leads to a competition between entropy-driven cluster growth and energetic suppression (or enhancement) of connectivity. Varying  $\varepsilon$  continuously interpolates between dense ferromagnetic-like clusters, ordinary classical percolation, and a dilute regime of minimally connected isolated clusters. Using Monte Carlo simulations and real-space renormalization-group (RG) methods, we show that bond energy shifts the percolation threshold smoothly. We define an energy-weighted correlation length that remains finite at the classical site occupation threshold ( $p_c(\varepsilon = 0)$ ) and shrinks with increasing  $\varepsilon$ , capturing the energetic suppression of large-scale connectivity. The cluster size distribution exhibits an energy-dependent cutoff that drives the transition from percolation-like clusters to isolated clusters. A real-space RG with Kadanoff block recursions reveals a systematic evolution of the correlation-length exponent  $\nu$  from  $\nu = 1/2$  (dense clusters) to  $\nu = 4/3$  (classical percolation), approaching  $\nu = 1$  (minimally connected isolated clusters), in agreement with Coulomb-gas predictions for loop models where bond energy renormalizes loop fugacity. For large values of  $\varepsilon$  (isotropic case), the suppression of nearest-neighbor bonds results in the emergence of antiferromagnetic sub-lattice ordering at high densities. Additionally, anisotropic bond energies lead to directionally selective cluster growth. Finally, we also discuss a lattice gas RG approach and scenarios where bond energy is renormalized across different scales.

## I. INTRODUCTION

Percolation theory provides a fundamental framework for understanding geometric phase transitions in disordered systems [1–3]. In its classical formulation, each site or bond of a lattice is independently occupied with probability  $p$ , and large-scale connectivity emerges when  $p$  exceeds a critical threshold  $p_c$  [1]. Near this threshold, the system exhibits universal scaling behavior governed by a diverging correlation length, fractal cluster geometry, and well-characterized critical exponents. These features place percolation within a broad universality class shared across problems ranging from transport in porous media to epidemic spreading, polymer networks, and the onset of rigidity in disordered solids [4–7]. Beyond the traditional connectivity-based framework, several generalized percolation models incorporate additional geometric, algebraic, and topological constraints. These include rigidity percolation [8–11], bootstrap percolation [12–15], quantum percolation [16–18], among other generalizations [19], where global order does not depend only on local occupancy. Standard or classical percolation models treat all occupied nearest-neighbor connections as equivalent. However, many physical systems exhibit additional energetic constraints or interactions that bias cluster formation. Examples include polymer networks with bending or stretching energies, resistor networks with weighted conductances [1, 4, 20–22], and random-cluster models where cluster weights depend on geometric or topological features [23, 24]. In such systems, connectivity is governed not only by quenched disorder but also

by energetic weights that modify the probability of realizing particular cluster configurations. Understanding how such energetic factors modify percolation transitions and critical scaling remains an important question in statistical mechanics, network science, and the physics of disordered materials.

Percolation can also be defined on correlated, scale-invariant backgrounds, rather than solely on uncorrelated lattices. A notable example of this is found in critical configurations of the two-dimensional  $O(n)$  loop model [25], where fractal loops partition the lattice into extensive regions. By introducing probabilistic connections between neighboring regions, a percolation transition can occur within this critical geometric framework. Despite the strong correlations created by the loop structure, this transition has been demonstrated to fall within the universality class of Kasteleyn–Fortuin (FK) random-cluster models, which are linked to the Potts universality class [25]. This indicates that universal percolative behavior persists even in systems where connectivity arises from an interacting framework.

In this work, we introduce and analyze an energy-weighted site-percolation (EWSP) model, which is defined by the site occupation probability. In this model, each bond between nearest-neighbor occupied sites incurs an energy cost denoted by  $\varepsilon$ . A cluster consisting of  $T$  such bonds then contributes a Boltzmann weight of  $e^{-\varepsilon T}$ . This results in a grand canonical ensemble of clusters, where the combinatorial multiplicity of a cluster with a given number of occupied sites competes with an energetic penalty that is proportional to its connectivity. When  $\varepsilon = 0$ , the model simplifies to classical site percolation. As  $\varepsilon$  approaches infinity, the model resembles an isolated cluster phase, where clusters with the fewest bonds between nearest-neighbor occupied sites dominate. Therefore, the tuning parameter  $\varepsilon$  allows for a continu-

\* sircars@tifrh.res.in

† kramola@tifrh.res.in

ous interpolation between classical percolation universality and a regime with strongly suppressed or enhanced clusters. The introduction of the energy cost  $\varepsilon$  significantly alters both the cluster-size distribution and the correlation length. Consequently, the correlation length remains finite even at  $p = p_c(\varepsilon = 0)$ . This behavior parallels the suppression of FK clusters in the  $q$ -state Potts model [26–28] and admits a natural interpretation within the Coulomb-gas formulation, where tuning  $\varepsilon$  corresponds to modifying the loop fugacity and thereby the thermal scaling field [29–31]. Energy-weighted site percolation thus provides a tunable extension of classical percolation in which energetic constraints reshape connectivity, critical scaling, and emergent spatial order.

To quantitatively characterize this model, we employ real-space RG methods based on Kadanoff block scaling [32]. These allow the site occupation probability  $p$  to renormalize under coarse-graining. Using different spanning-rules, we construct explicit RG recursion relations and compute fixed points and correlation-length exponents. Our results demonstrate a smooth change of the exponent  $\nu$  with increasing  $\varepsilon$ , consistent with Coulomb-gas predictions in which the thermal exponent interpolates between classical percolation ( $\nu = 4/3$ ) and the dilute cluster limit ( $\nu = 1$ ). Furthermore we formulate a lattice-gas representation defined by site fugacity  $e^\mu = \frac{p}{1-p}$  and bond fugacity  $e^{J-\varepsilon}$ . This framework produces exact recursion relations for block-renormalized parameters, allowing us to determine the RG flow. In our analysis using lattice gas RG, we explore two scenarios: one where the bond energy is tunable and does not depend on the length scale, and another where the bond energy cost is dependent on the length scale and undergoes renormalization. Overall, the EWSP model introduced here provides a tunable extension of classical percolation with rich crossover behavior and nontrivial modifications to critical scaling. Unlike traditional interacting percolation models [33, 34], which generate correlations through spin interactions or dynamic occupation rules, the EWSP model incorporates an explicit energetic cost for nearest-neighbor occupied bonds. Consequently, connectivity is influenced by this energetic factor, creating a tunable competition between entropy and bond energetics.

The paper is organized as follows: In Section II, we introduce the model of EWSP on a square lattice, the partition function, and we define the associated parameters that will be used in subsequent sections. Section III discusses our Monte Carlo simulation of the lattice gas in the grand canonical ensemble. It examines the cluster configurations for both isotropic and anisotropic bond energy cases. In the isotropic case, we illustrate the emergence of antiferromagnetic sub-lattice ordering in site occupancy, while in the anisotropic case, we focus on strip-like ordering. In Section IV, we compute real-space RG recursion relations and determine the critical site occupation probability, and the correlation length critical exponent (both isotropic and anisotropic). In Section V, we use results

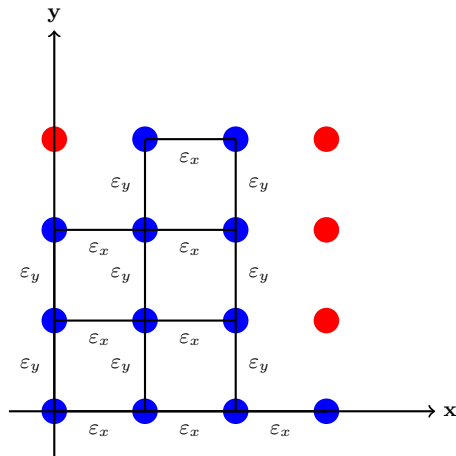


FIG. 1. A  $4 \times 4$  square lattice with 12 occupied (blue) and 4 unoccupied (red) sites. Bonds between nearest-neighbor occupied sites are shown in black and have an energy cost  $\varepsilon_x$  for bonds along horizontal  $\mathbf{x}$  direction, and energy cost  $\varepsilon_y$  for bonds along vertical  $\mathbf{y}$  direction. The connected occupied sites form a cluster. In this diagram, there is only one cluster containing 12 occupied sites and 16 bonds connecting nearest-neighbor occupied sites.

from the Coulomb gas to explain different limits of energy cost, denoted as  $\varepsilon$ , and the corresponding changes in the critical exponent associated with these limits. In Section VI, we derive the expression for the correlation length related to energy-weighted percolation in the canonical ensemble and discuss cluster size statistics. Finally, in Section VII, we derive the RG recursion relation using the lattice gas analogy expressed in terms of site fugacity and bond fugacity.

## II. THE MODEL

We propose an energy-weighted generalization of site percolation on a two-dimensional square lattice of linear dimension  $L$ , with  $N = L^2$  total sites, as shown in Fig. 1. Each lattice site  $i$  is associated with a binary occupation variable  $n_i \in \{0, 1\}$ , where  $n_i = 1$  denotes an occupied site and  $n_i = 0$  an unoccupied site. In the absence of energetic constraints, sites are independently occupied with probability  $p$ , recovering ordinary site percolation. Undirected bonds connect occupied nearest-neighbor sites. A connected component of occupied sites defines a cluster  $C$ . For a given cluster  $C$ , we denote by  $s(C)$  the number of occupied sites and by  $T(C)$  the total number of nearest-neighbor bonds between occupied sites within the cluster, counting each bond once. In standard percolation, all configurations with the same number of occupied sites or bonds are treated equivalently. Here, we introduce an additional energetic constraint by assigning an energy cost  $\varepsilon$  to each bond connecting nearest-neighbor occupied sites. The total energy of a cluster is therefore

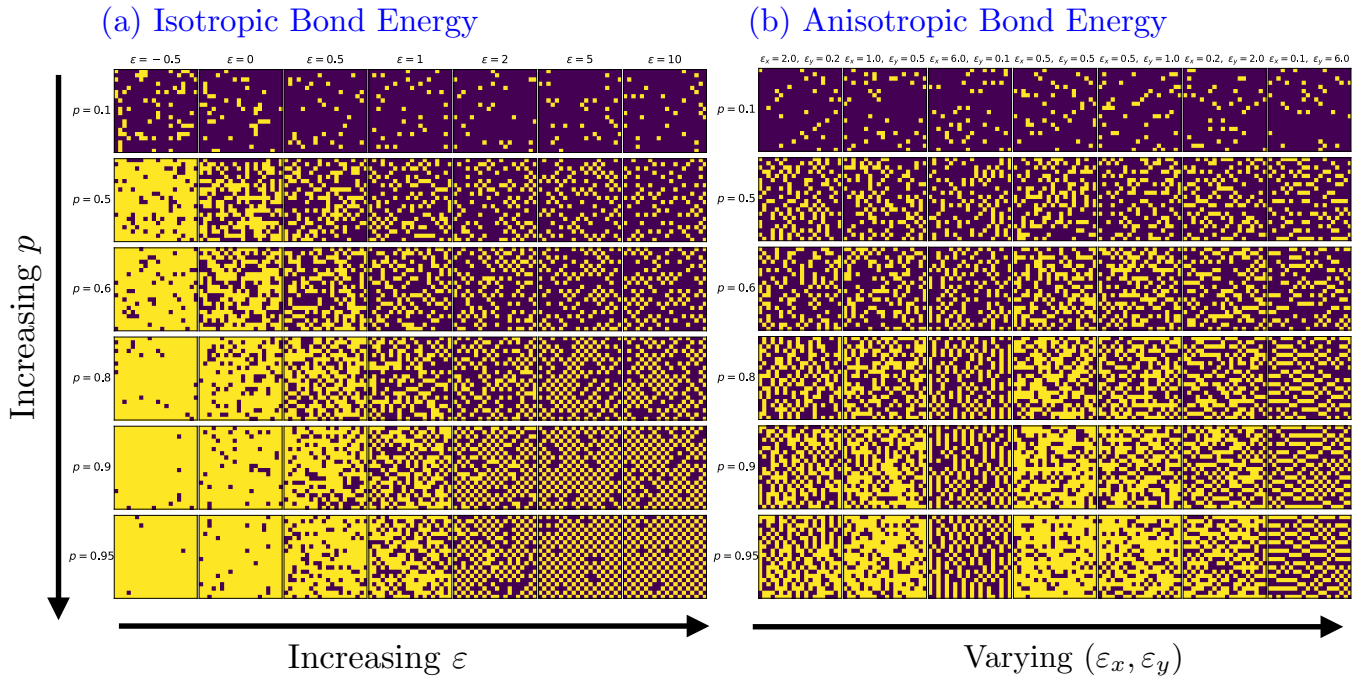


FIG. 2. The left plot (a) illustrates the variation in cluster configurations as a function of site-occupation probability  $p$  and energy cost  $\varepsilon$ , and the right plot (b) illustrates the variation in cluster configurations for different values of site occupation probability  $p$  and energy cost  $\varepsilon_x$  for bonds along the direction of  $x$  axis and  $\varepsilon_y$  for bonds along the direction of  $y$  axis. The computation is performed on a square grid with a total of  $N = 400$  sites. The colors indicate the occupancy variable  $n_{ij}$ , with  $n_{ij} = 1$  indicating that a particular site is occupied (yellow), and  $n_{ij} = 0$  indicating that a particular site is not occupied (blue).

defined as

$$E(C) = \varepsilon T(C). \quad (1)$$

A lattice configuration  $\mathcal{C}$  consists of a collection of disjoint clusters  $\{C_k\}$ . The partition function of the EWSP model can be computed as:

$$Z_{\text{EWSP}} = \sum_{\mathcal{C}} p^{N_{\text{occ}}} (1-p)^{N-N_{\text{occ}}} \prod_k e^{-\varepsilon T(C_k)}. \quad (2)$$

Using the partition function defined in Eq. (2), a probability measure can be assigned to each configuration as:

$$W_{\text{EWSP}}(\mathcal{C}) = \frac{1}{Z_{\text{EWSP}}} p^{N_{\text{occ}}} (1-p)^{N-N_{\text{occ}}} \prod_k e^{-\varepsilon T(C_k)}, \quad (3)$$

where  $N_{\text{occ}} = \sum_i n_i$  is the total number of occupied sites in a particular configuration. This defines a grand-canonical ensemble in which both the number of occupied sites and the internal connectivity of a configuration fluctuate. The site occupation probability  $p$  controls the density of occupied sites, while the parameter  $\varepsilon$  penalizes highly connected clusters for  $\varepsilon > 0$ . From a statistical-mechanical perspective, the model may be viewed as a lattice gas with site fugacity

$$e^\mu = \frac{p}{1-p}, \quad (4)$$

and bond fugacity  $e^{-\varepsilon}$ . The resulting ensemble reflects a competition between entropic combinatorial factors favoring large clusters and energetic suppression or enhancement of configurations with many internal bonds. The lattice gas partition function in terms of this site and bond fugacity can be written as:

$$Z_{\text{LG}} = \sum_{\mathcal{C}} e^{\mu \sum_i n_i - \varepsilon \sum_{\langle ij \rangle} n_i n_j}. \quad (5)$$

In this context,  $n_i$  represents the occupation variable at site  $i$ , taking values of either (0,1). For  $n_i = 0$  the site is unoccupied, while  $n_i = 1$  indicates that the site is occupied,  $\langle ij \rangle$  denotes the nearest neighbor sites  $i$  and  $j$ . One can easily observe the connection between the lattice gas and site percolation in the hard-core on-site exclusion limit. The grand canonical partition function for the lattice gas, as shown in Eq. (5), can be computed in the hard-core on-site exclusion limit as:

$$Z_{\text{HLG}} = (1 + e^\mu)^N, \quad (6)$$

where  $N$  is the local number of sites on the lattice, and **HLG** stands for hard-core lattice gas. Using this partition function, the density of occupied sites can be computed as:

$$\rho(\varepsilon = 0) = \frac{1}{N} \frac{\partial \ln Z_{\text{HLG}}}{\partial \mu} = \frac{e^\mu}{1 + e^\mu} = p. \quad (7)$$

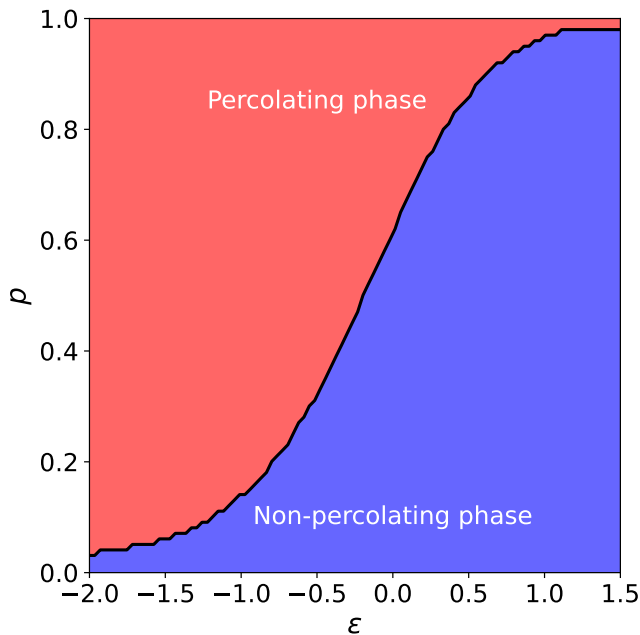


FIG. 3. Phase diagram indicating percolating and non-percolating phase separated by critical curve  $p_c(\varepsilon)$ , in the  $\varepsilon - p$  plane. The phase diagram is obtained using numerical Monte Carlo simulation on a square lattice with dimensions  $N = 40000$ , as detailed in Sec. III. The blue colored region indicates the non-percolating phase, while the red colored region indicates the percolating phase.

Thus, in the case of a hard-core on-site exclusion limit, the density of occupied sites is precisely equal to the site occupation probability. However, in a finite  $\varepsilon$  domain, the density of occupied sites is influenced by both the site occupation probability  $p$ , through site fugacity, and the bond energy  $\varepsilon$ , which reflects the interactions between nearest neighbor occupied sites. This mapping of the lattice gas model in the finite  $\varepsilon$  domain corresponds directly to that of the Ising model in a finite magnetic field [35, 36]. Our study essentially seeks to understand the percolation phase transition in the Lee-Yang lattice gas [36], which is extensively studied in the literature regarding the order-disorder phase transition [37, 38].

Increasing  $p$  promotes connectivity, while increasing  $\varepsilon > 0$  suppresses it. Several limiting cases are of particular interest. When  $\varepsilon = 0$ , all clusters carry weight proportional to the number of occupied sites and the model reduces exactly to standard or classical site percolation, with critical behavior governed by the standard percolation universality class. For  $\varepsilon > 0$ , clusters with many internal bonds are exponentially suppressed, leading to a reduced cutoff in the cluster-size distribution as shown in Fig. 9 and a diminished correlation length as shown in Fig. 10. In the limit  $\varepsilon \rightarrow \infty$ , only minimally connected clusters contribute significantly, and the system is dominated by isolated cluster-like structures reminiscent of dilute lattice walks. The introduction of the energy

cost  $\varepsilon$  therefore provides a continuous tuning parameter that interpolates between classical percolation and an isolated-cluster regime with antiferromagnetic ordering in site occupancy, as shown in Fig. 2. In defining the partition function in Eq. (2), we assume that the energy cost per bond, denoted as  $\varepsilon$ , is a tunable parameter that remains constant across all length scales. This means we allow the bond variable to remain unchanged during the renormalization process in our real-space RG scheme presented in Sec. IV. However, in the lattice gas RG framework presented in Sec. VII, we will relax this assumption. We will consider  $\varepsilon$  not as a constant tunable parameter, but rather as a variable that depends on the length scale, similar to site occupation probability, and that gets renormalized. Unlike classical percolation, which is controlled by a single parameter  $p$ , the present model exhibits a two-parameter phase space  $(p, \varepsilon)$ . As we show below, increasing  $\varepsilon$  shifts the effective percolation threshold  $p_c(\varepsilon)$  and can render the correlation length finite even at the classical percolation point  $p_c(\varepsilon = 0)$ . In what follows, clusters are identified as connected components of occupied sites, and all observables, including cluster-size distributions and correlation lengths, are computed as ensemble averages over the energy-weighted measure defined above. This formulation provides the foundation for the numerical simulations, correlation-length computation, and real-space RG calculations presented in the subsequent sections. Through the lattice-gas simulation discussed in Sec. III, we obtained a phase diagram that illustrates the percolating and non-percolating phases with varying bond energy costs  $\varepsilon$ , as shown in Fig. 3. In this phase diagram, it is evident that as  $\varepsilon$  approaches  $-\infty$ , the critical site occupation probability  $p_c(\varepsilon)$  decreases. This allows for a transition from the non-percolating phase to the percolating phase at very low occupation probabilities, as the bond energy in this scenario enhances connectivity. Conversely, as  $\varepsilon$  approaches  $+\infty$ , the critical site occupation probability  $p_c(\varepsilon)$  increases, leading to a transition from the non-percolating phase to the percolating phase at very high occupation probabilities, since the bond energy in this case suppresses connectivity. Overall, this phase diagram effectively illustrates the competition between entropy and energy costs in determining global connectivity within a disordered network.

Interestingly, this model can be exactly mapped to the high-temperature expansion of the  $O(n)$  spin model [27], especially in the high and low-energy limit, where  $\varepsilon \rightarrow \pm\infty$ . In this regime, we can also apply the results from the Coulomb gas, as demonstrated in Sec. V. In this representation, the partition function can be written as a sum over loop configurations  $\mathcal{G}$ ,

$$Z_{\mathbf{O}(n)} = \sum_{\mathcal{G}} n^{C(\mathcal{G})} K^{|\mathcal{G}|}, \quad (8)$$

where  $C(\mathcal{G})$  is the number of connected components (closed loops),  $|\mathcal{G}|$  is the total number of occupied bonds, and  $K$  is the bond fugacity [39]. This formulation emphasizes that the statistical weight depends explicitly on

the connectivity of the underlying graph, a feature shared with percolation-type models. In the limit  $n \rightarrow 0$ , configurations containing closed loops are suppressed, and only graphs with a single open trajectory contribute. A closely related structure appears in generalized percolation models, particularly in bond percolation [27], most notably in the random-cluster (FK) formulation. The partition function for this is given by:

$$Z_{\text{RC}} = \sum_{\mathcal{G}} p^{|\mathcal{G}|} (1-p)^{|E|-|\mathcal{G}|} q^{C(\mathcal{G})}, \quad (9)$$

where  $p$  represents the bond occupation probability,  $|E|$  is the total number of edges in the lattice, and  $q$  is a fugacity that controls the number of connected clusters. In ordinary percolation, this corresponds to the limit  $q \rightarrow 1$  [23].

The explicit expression for the partition function of our model is derived in Sec. II. Its connection to the FK cluster partition function is established in Sec. V. By comparing Eq. (8) and Eq. (9), it becomes evident that the loop fugacity  $n$  in the  $O(n)$  model plays a role analogous to the cluster fugacity  $q$  in the random-cluster model. The limit of single open trajectories as  $n \rightarrow 0$  corresponds to a generalized percolation regime where closed clusters are suppressed, allowing a single non-self-intersecting path to dominate connectivity.

### III. SIMULATION DETAILS

We simulate our model as an interacting lattice gas with nearest-neighbor interaction strength  $\varepsilon$ . To obtain the cluster configurations for various values of  $\varepsilon$ , we consider a two-dimensional square lattice of linear size  $L$  with binary occupation variables  $n_{ij} \in \{0, 1\}$  defined on each lattice site  $(i, j)$ . For convenience, we map this occupation variable onto an Ising-like spin variable  $S_{ij}$  according to

$$S_{ij} = 2n_{ij} - 1 = \begin{cases} +1, & \text{if } n_{ij} = 1, \\ -1, & \text{if } n_{ij} = 0. \end{cases} \quad (10)$$

Undirected bonds connect nearest-neighbor occupied sites, and connected components of occupied sites define clusters. For each cluster  $\mathcal{C}$ , we define its internal bond number  $T(\mathcal{C})$  as the total number of bonds between nearest neighbor occupied sites within the cluster, counted once per bond, as defined previously in Sec. II. The energy of a configuration is taken to be  $E = \varepsilon \sum_{\mathcal{C}} T(\mathcal{C})$ , where  $T(\mathcal{C})$  denotes the energy of cluster  $\mathcal{C}$ . Configurations with a mean occupation probability  $\rho \neq p$  are generated using a grand-canonical Monte Carlo using a Metropolis scheme [40], where the chemical potential is defined as  $\mu = \ln[p/(1-p)]$ . Because we employ grand-canonical Glauber-like sweeps [41, 42] over occupation variables, the mean occupation probability of the equilibrated configuration will not equal the initial mean occupation probability  $p$ , but depends on both  $p$  and bond energy cost  $\varepsilon$ . This equality holds only when  $\varepsilon = 0$ . The site

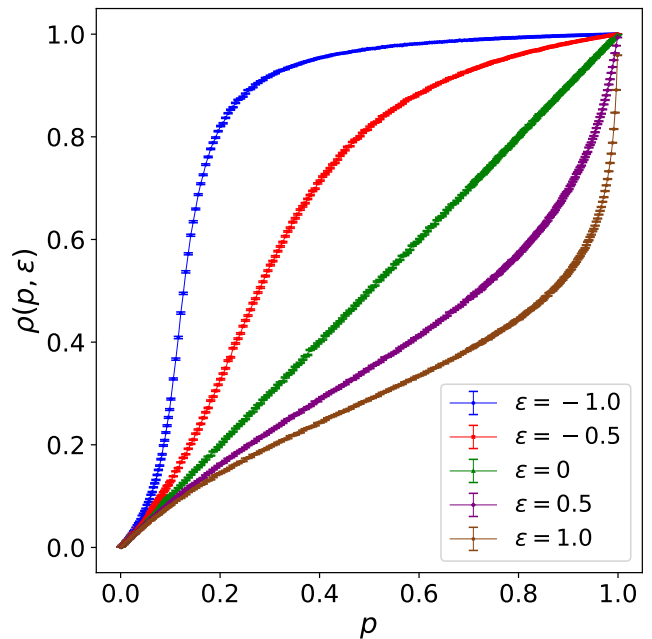


FIG. 4. The plot illustrates the variation in the mean density of occupied sites, defined as  $\rho = \frac{N_{occ}}{L^2}$ , where  $N_{occ}$  represents the number of occupied sites. This value is averaged over 500 realizations of equilibrated samples while varying the site occupation probability  $p$  in the initial configuration, considering different values of bond energy cost  $\varepsilon$ . The system size used in the study is  $N = 400$ . The error bar represents the standard error of the mean density of occupied states. The equilibration is carried out using grand canonical Glauber-like Metropolis sweeps, as outlined in Sec. III.

occupation probability  $p$  of the initial non-equilibrated configuration sets the site fugacity  $e^\mu$ . Therefore, in all subsequent discussions, we will refer to  $p$  instead of  $\rho$ . It is important to note this distinction. The variation in the density of occupied sites  $\rho(p, \varepsilon)$ , averaged over numerous realizations of equilibrated samples with a site occupation probability  $p$ , is presented in Fig. 4. Additionally, Fig. 5 illustrates the variation of the density of occupied sites  $\rho(p_c(\varepsilon))$  evaluated at the critical occupation probability  $p_c(\varepsilon)$  for different values of bond energy cost  $\varepsilon$ . It is evident from computations conducted on finite system sizes that there is a minimum density of occupied sites, at a fixed bond energy cost  $\varepsilon$ , where a percolating phase can exist. This can be explored in greater detail, particularly in the thermodynamic limit, in future studies. In all the cases, the equilibration is performed using Glauber-like Metropolis sweeps [40–42]. On the other hand, one can keep the density of occupied sites constant as the initial configurations by using Kawasaki swaps [42, 43] for equilibration. However, since our objective is to examine the grand canonical partition function, we will not use Kawasaki swaps to compute the observables. Starting from a random initial configuration with occupation probability  $p$ , we perform a sequence of

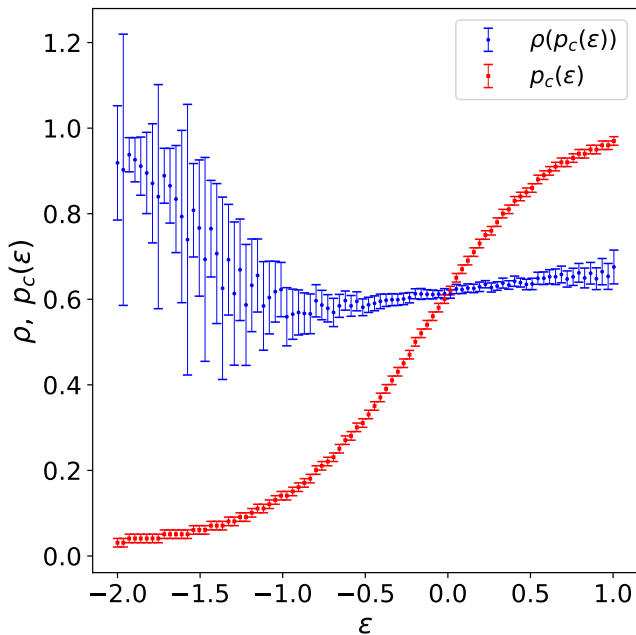


FIG. 5. The plot illustrates the density of occupied sites, denoted as  $\rho(p_c(\varepsilon))$ , which is evaluated at the critical site occupation probability  $p_c(\varepsilon)$  for various values of the bond energy cost  $\varepsilon$ . The data for  $p_c(\varepsilon)$  is also presented in Fig. 3. Additionally, the corresponding mean density of occupied sites,  $\rho(p_c(\varepsilon), \varepsilon)$ , is computed for a system size of  $N = 400$ .

Monte Carlo sweeps [40–42]; during each sweep,  $L^2$  trial updates are attempted by selecting a site at random and proposing to flip its occupation state. The proposed move is accepted with probability  $\min\{1, \exp[-\Delta H]\}$ , where  $\Delta H = \Delta E - \mu \Delta N$ ,  $\Delta E$  is the change in interaction energy induced by the flip, and  $\Delta N = \pm 1$  is the corresponding change in occupation number [44]. After equilibration, clusters are identified by constructing the reduced adjacency matrix of occupied sites and computing its connected components. The cluster configurations are obtained from sampling equilibrated states. They are illustrated in Fig. 2(a) for the isotropic bond energy case and in Fig. 2(b) for the anisotropic bond energy case. As the value of  $\varepsilon$  increases (with  $\varepsilon > 0$ ) in the isotropic case, there is an energetic suppression of clusters, leading to sub-lattice ordered configurations. This ordering phenomenon is discussed in detail in the Supplementary Information [45]. The cluster size distribution  $n_s$  is calculated using this Monte Carlo sampling, as shown in Fig. 12(a) for isotropic bond energy cost and in Fig. 12(b) for anisotropic bond energy cost. From Fig. 12(b), it is evident that compared to the isotropic case, where  $\varepsilon_x = \varepsilon_y$ , the tails of the distribution are reduced and primarily consist of configurations not suppressed by the energetic cost of the bonds. Through this lattice gas simulation, we generate the phase diagram for percolation phases in the  $p - \varepsilon$  plane, as illustrated in Fig. 3. One can also calculate the critical exponent for correla-

tion length, denoted as  $\nu(\varepsilon)$ , along with the critical exponent for susceptibility or average cluster size, denoted as  $\gamma(\varepsilon)$ . These quantities diverge at the critical occupation probability  $p_c(\varepsilon)$ . This is done using finite-size scaling to collapse the wrapping probability  $P(p, L, \varepsilon)$  and the susceptibility  $\chi(p, L, \varepsilon)$  as follows:

$$P(p, L, \varepsilon) \sim G(|p - p_c(\varepsilon)|L^{\frac{1}{\nu(\varepsilon)}}), \quad (11)$$

$$\chi(p, L, \varepsilon) \sim L^{-\frac{\gamma(\varepsilon)}{\nu(\varepsilon)}} H(|p - p_c(\varepsilon)|L^{\frac{1}{\nu(\varepsilon)}}), \quad (12)$$

where  $G$  and  $H$  are universal scaling functions, as shown in Fig. 7 and Fig. 8.

#### IV. REAL SPACE RENORMALIZATION GROUP ANALYSIS

We consider the site-occupation model on the square lattice used in the numerics: each site is occupied independently with probability  $p$ , and a configuration with total  $T$  bonds between nearest-neighbor occupied sites is weighted by

$$W(\mathcal{C}) = \frac{1}{Z_{\text{EWSP}}} p^{N_{\text{occ}}} (1-p)^{N-N_{\text{occ}}} e^{-\varepsilon T}, \quad (13)$$

where the Boltzmann weight for each cluster in the configuration is multiplied, and  $Z_{\text{EWSP}}$  represents the partition function defined in Eq. (2). The set of configurations is labeled by  $\mathcal{C}$ . In this section, we will rescale the original lattice to reduce the degrees of freedom in the system, allowing us to study its large-scale behavior using the real-space renormalization group approach, as presented in [32, 46, 47]. So the partition function for the rescaled block ( $Z_R^b$ ) can be written as:

$$Z_R^b = \sum_k \sum_{\mathcal{C}} p^k (1-p)^{N-k} \prod_{i \in \mathcal{C}} e^{-\varepsilon T_i}, \quad (14)$$

where  $b$  stands for the linear dimension of block, and  $R$  specifies the rule one is considering for renormalization coarse-graining. We rescale the lattice by a factor of  $b$ , the linear dimension of each block. We assign the occupation probability of each block  $\tilde{p}$ , according to a rule. The block occupation probability can be expanded using the usual procedure of real-space renormalization as:

$$\tilde{p} = \frac{1}{Z_R^b} \sum_k \sum_{\mathcal{C}'} p^k (1-p)^{N-k} \prod_{i \in \mathcal{C}'} e^{-\varepsilon T_i}, \quad (15)$$

where  $\mathcal{C}'$  is the set of configurations satisfying the specific rule one is choosing for coarse-graining. For rules like the “majority rule” and other rules termed  $R_0$ ,  $R_1$ , and  $R_2$  as stated below, the numerator in the above expression in Eq. (15) simplifies to:

$$\tilde{p} = \frac{1}{Z_R^b} \sum_k M_k p^k q^{N-k} e^{-\varepsilon T_k}, \quad q = 1 - p, \quad (16)$$

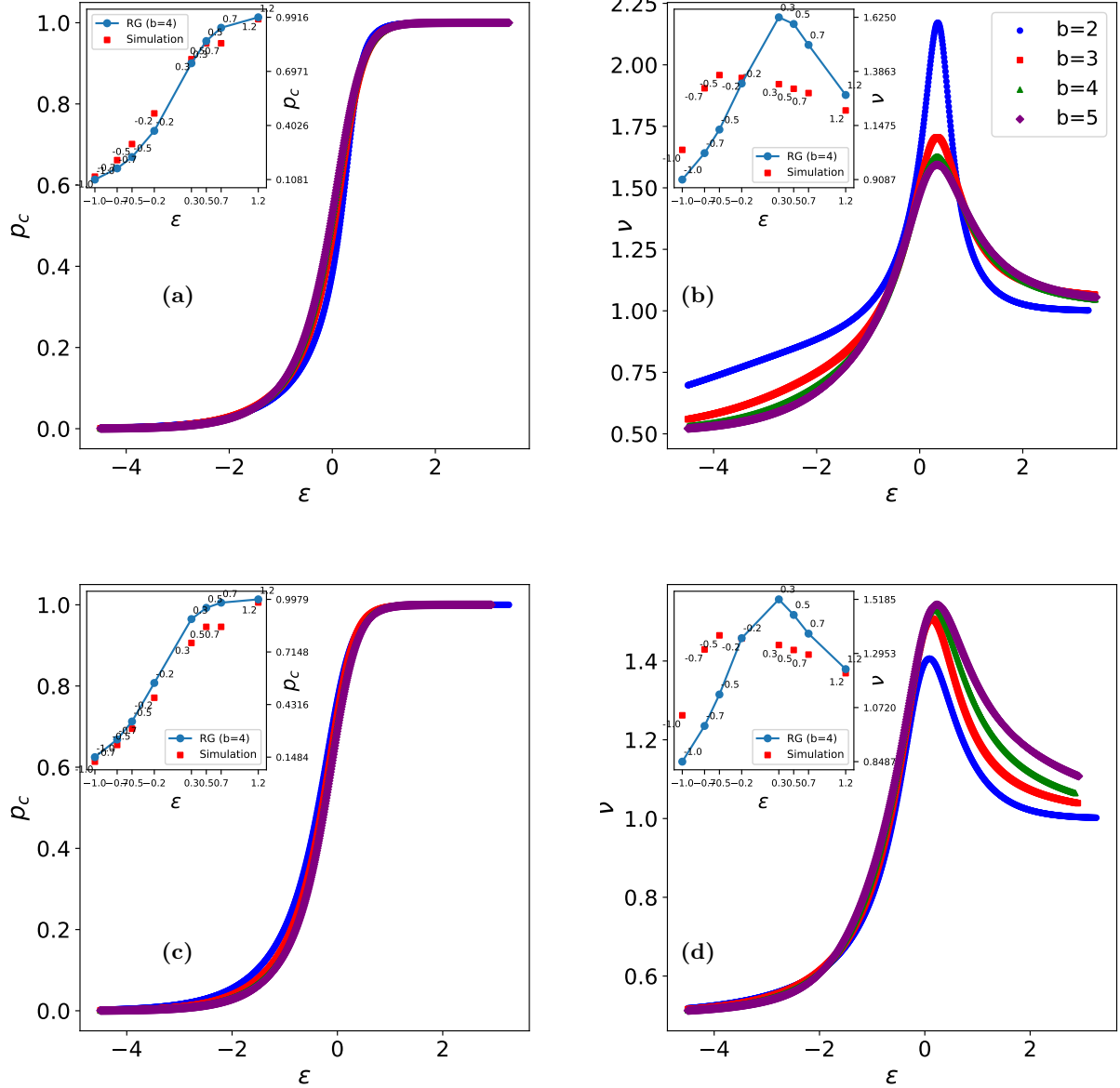


FIG. 6. Panels (a,b):  $p_c$  and  $\nu$  computed using the  $R_0$  rule for block sizes  $b \in [2, 5]$ . Panels (c,d):  $p_c$  and  $\nu$  computed using the  $R_2$  rule for block sizes  $b \in [2, 5]$ . The insets display some values obtained from simulations and the spanning rule. All quantities are shown as functions of the energy cost per bond  $\varepsilon$ .

where  $q$  is the probability that a site is not occupied,  $T_k$  is the total number of bonds across all clusters of  $k$  connected sites in each block, and the coefficient  $M_k$  depends on the rule one chooses. We consider different such rules for analysis below. A rule can be such that  $\tilde{p} = 1$ , where at least  $k = \lceil \frac{N}{2} \rceil$  sites in a block are occupied, which corresponds to “majority rule” in literature. The coefficient  $M_k$  for this type of rule corresponds to the binomial coefficient. Other sets of rules are considered in [48] termed as  $R_0$ ,  $R_1$ , and  $R_2$ . The rules state that a block is considered occupied if it contains a spanning cluster. For rule  $R_0$ , the cluster can span the block either horizon-

tally or vertically. In the case of rule  $R_1$ , the cluster spans the block in a fixed direction (for example, horizontally), while rule  $R_2$  requires the cluster to span the block both horizontally and vertically [48]. These are the conditional statements for all three different rules. For example, below provides the polynomials for the renormalized site occupation probability  $\tilde{p} = R_0^b(p)$  using rule  $R_0$  (for  $\varepsilon = 0$ ) for block sizes  $b \in [2, 5]$  as: For  $b = 2$ :

$$R_0^2(p, q, \varepsilon = 0) = p^4 + 4p^3q + 4p^2q^2, \quad (17)$$

for  $b = 3$ :

$$R_0^3(p, q, \varepsilon = 0) = p^9 + 9p^8q + 36p^7q^2 + 82p^6q^3 + 93p^5q^4 + 44p^4q^5 + 6p^3q^6. \quad (18)$$

for  $b = 4$ :

$$R_0^4(p, q, \varepsilon = 0) = p^{16} + 16p^{15}q + 120p^{14}q^2 + 560p^{13}q^3 + 1818p^{12}q^4 + 4296p^{11}q^5 + 7196p^{10}q^6 + 8136p^9q^7 + 5988p^8q^8 + 2784p^7q^9 + 780p^6q^{10} + 120p^5q^{11} + 8p^4q^{12}, \quad (19)$$

and for  $b = 5$ :

$$R_0^5(p, q, \varepsilon = 0) = p^{25} + 25p^{24}q + 300p^{23}q^2 + 2300p^{22}q^3 + 12650p^{21}q^4 + 53128p^{20}q^5 + 176992p^{19}q^6 + 478316p^{18}q^7 + 1054923p^{17}q^8 + 2666712p^{16}q^{10} + 2963364p^{14}q^{11} + 2556058p^{13}q^{12} + 1699665p^{12}q^{13} + 865132p^{11}q^{14} + 333630p^{10}q^{15} + 95845p^9q^{16} + 19916p^8q^{17} + 2836p^7q^{18} + 248p^6q^{19} + 10p^5q^{20} + 1880864p^{16}q^9, \quad (20)$$

where  $q = 1 - p$ , the probability that a site is not occupied. The expressions are also obtained in [48]. For other rules like  $R_1$  and  $R_2$ , similar expansions can be obtained by changing the conditional statement specified before. A full analytic derivation of this expression for a particular  $b = 3$  (Eq. (18)) using rule  $R_0$  has been shown in Appendix A using the inclusion-exclusion theorem. For  $\varepsilon \neq 0$ , the expansion using spanning rule  $R_0$  can be computed as: For  $b = 2$ :

$$R_0^2(p, q, \varepsilon) = \frac{4p^2q^2e^{-\varepsilon} + 4p^3qe^{-2\varepsilon} + p^4e^{-4\varepsilon}}{Z_0^2}, \quad (21)$$

with scaled block partition function as

$$Z_0^2(p, q, \varepsilon) = q^4 + 4pq^3 + 2p^2q^2 + 4p^2q^2e^{-\varepsilon} + 4p^3qe^{-2\varepsilon} + p^4e^{-4\varepsilon},$$

for  $b = 3$ :

$$R_0^3(p, q, \varepsilon) = \frac{1}{Z_0^3} \left( 6p^3q^6e^{-2\varepsilon} + 12p^4q^5e^{-2\varepsilon} + 32p^4q^5e^{-3\varepsilon} + 4p^5q^4e^{-2\varepsilon} + 40p^5q^4e^{-3\varepsilon} + 33p^5q^4e^{-4\varepsilon} + 16p^5q^4e^{-5\varepsilon} + 4p^6q^3e^{-3\varepsilon} + 22p^6q^3e^{-4\varepsilon} + 28p^6q^3e^{-5\varepsilon} + 24p^6q^3e^{-6\varepsilon} + 4p^6q^3e^{-7\varepsilon} + 14p^7q^2e^{-6\varepsilon} + 8p^7q^2e^{-7\varepsilon} + 14p^7q^2e^{-8\varepsilon} + p^8qe^{-8\varepsilon} + 4p^8qe^{-9\varepsilon} + 4p^8qe^{-10\varepsilon} + p^9e^{-12\varepsilon} \right), \quad (22)$$

with scaled block partition function given as:

$$Z_0^3(p, q, \varepsilon) = q^9 + p^8qe^{-8\varepsilon} + 4p^8qe^{-9\varepsilon} + 4p^8qe^{-10\varepsilon} + p^9e^{-12\varepsilon} + 9pq^8 + 24p^2q^7 + 12p^2q^7e^{-\varepsilon} + 22p^3q^6 + 40p^3q^6e^{-\varepsilon} + 22p^3q^6e^{-2\varepsilon} + 6p^4q^5 + 28p^4q^5e^{-\varepsilon} + 56p^4q^5e^{-2\varepsilon} + 32p^4q^5e^{-3\varepsilon} + 4p^4q^5e^{-4\varepsilon} + p^5q^4 + 24p^5q^4e^{-2\varepsilon} + 48p^5q^4e^{-3\varepsilon} + 37p^5q^4e^{-4\varepsilon} + 16p^5q^4e^{-5\varepsilon} + 4p^6q^3e^{-3\varepsilon} + 24p^6q^3e^{-4\varepsilon} + 28p^6q^3e^{-5\varepsilon} + 24p^6q^3e^{-6\varepsilon} + 4p^6q^3e^{-7\varepsilon} + 14p^7q^2e^{-6\varepsilon} + 8p^7q^2e^{-7\varepsilon} + 14p^7q^2e^{-8\varepsilon}, \quad (23)$$

The expansions of the renormalized occupation probability for a block size of  $b = 4$  are provided in the Supplementary Information [45]. The renormalized spanning probability calculated using the spanning rule  $R_2$  is also detailed in Supplementary Information [45]. We note that by setting  $\varepsilon = 0$ , the polynomial expressions for the spanning probability simplify to the expression presented in [48].

The expression for critical exponent  $\nu$  is given as:

$$\nu(\varepsilon) = \frac{\ln(b)}{\ln\left(\frac{d\bar{p}(p_c(\varepsilon))}{dp}\right)}, \quad (24)$$

with  $p_c(\varepsilon)$  are the fixed points of the RG flow for each value of bond energy cost  $\varepsilon$ . This expression is obtained by linearizing about each fixed point of the RG flow equation (Eq. (16)) [49]. It is important to emphasize that while this real-space RG treatment is accurate for hierarchical lattices [50, 51], it is only approximately correct for regular lattices like the one we are considering.

The variation of the critical site occupation probability  $p_c$  and the critical exponent  $\nu$  using renormalization rules  $R_0$  and  $R_2$  are illustrated in Fig. 6(a,b) and Fig. 6(c,d) respectively. As the block size  $b$  increases, the computed curves for the critical exponent  $\nu$  using rules  $R_0$  and  $R_2$  become closer to each other, as can be inferred from Fig. 6(b,d). In the general case, where  $\varepsilon_x$  represents the energy cost for the bonds in the  $x$  direction and  $\varepsilon_y$  for the bonds in the  $y$  direction, the variation of these quantities using rules  $R_0$  and  $R_2$  are shown in Fig. 18 and Fig. 17, respectively. All the fixed points  $p_c(\varepsilon)$  are unstable fixed points in the  $p$  direction, with the only two basins of attraction or stable fixed points being at  $p_c(\varepsilon_+) = 1$  and  $p_c(\varepsilon_-) = 0$ , respectively. In this context,  $\varepsilon_+$  and  $\varepsilon_-$  represent the values of  $\varepsilon$  when  $p_c = 1$  and  $p_c = 0$ , with  $\varepsilon_- < 0$  and  $\varepsilon_+ > 0$ .

In the Monte Carlo simulations, we obtain equilibrated configurations for each value of the bond energy cost  $\varepsilon$ . This parameter governs the effective interaction between neighboring sites and controls the connectivity properties of the system. As a result, variations in  $\varepsilon$  lead to systematic shifts in the critical site occupation probability  $p_c$ .

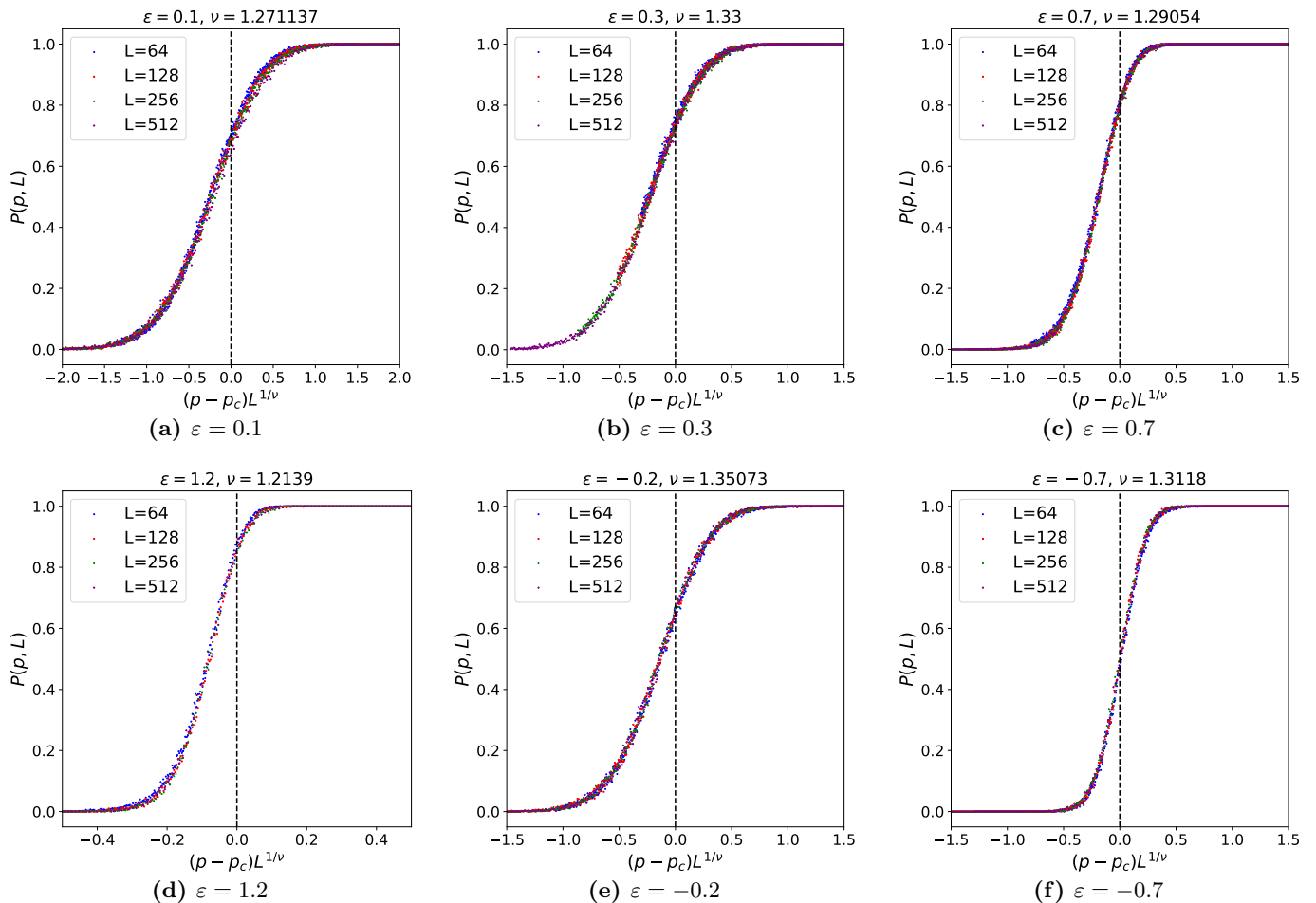


FIG. 7. The plot illustrates the scaling collapse of the wrapping probability function  $P(p, L)$  for different system sizes  $L$ , computed from Metropolis Monte Carlo simulations. The critical exponent  $\nu = \nu(\varepsilon)$  has been extracted for various cases of bond energy cost  $\varepsilon$  as follows: **(a)**  $\varepsilon = 0.1$ :  $p_c \approx 0.65301$ , **(b)**  $\varepsilon = 0.3$ :  $p_c \approx 0.76348$ , **(c)**  $\varepsilon = 0.7$ :  $p_c \approx 0.85$ , **(d)**  $\varepsilon = 1.2$ :  $p_c \approx 0.9812$ , **(e)**  $\varepsilon = -0.2$ :  $p_c \approx 0.4683$ , **(f)**  $\varepsilon = -0.7$ :  $p_c \approx 0.2141$ . These plots illustrate how varying bond energies affect the critical occupation probabilities and exponents.

In this context,  $\varepsilon$  is treated as an externally tunable parameter with no assigned scale dependence. Similarly, in the initial RG schemes considered here, the bond energy is not renormalized. Instead,  $\varepsilon$  is again treated as a fixed, externally adjustable parameter, similar to its role in the Monte Carlo simulations. The RG transformations in this framework act solely on the configurational or connectivity degrees of freedom and do not incorporate feedback on the interaction strength. Consequently, the analysis captures how large-scale connectivity emerges for a given value of  $\varepsilon$ , but it does not address how the interaction parameter itself evolves during coarse-graining. In the lattice gas RG framework discussed in Sec. VII, we consider both scenarios in a unified manner. First, we analyze the case in which the bond energy remains a non-renormalized, tunable parameter, thus maintaining a direct correspondence with the Monte Carlo simulations and the earlier RG treatment. Second, we extend the analysis to include the explicit renormalization of the bond energy, treating  $\varepsilon$  as a scale-

dependent quantity that changes under the RG transformation. This latter approach provides a more complete and self-consistent description, as it accounts for the coupled evolution of both site occupation probability and interaction strength. The implications of these two treatments are examined and compared in detail in that section.

## V. SCALING EXPONENTS

The scaling exponents of our model can be derived by relating it to a known loop model, as discussed in Sec. II. We can expand the partition function in Eq. (14) in the vicinity of the limit as  $\varepsilon \rightarrow \infty$ , focusing on the number of loops. In this limit, no loops are present as  $\varepsilon$  approaches infinity. We define a small expansion parameter as follows:

$$K \equiv e^{-\varepsilon}, \quad K \ll 1 \quad (\varepsilon \rightarrow \infty). \quad (25)$$

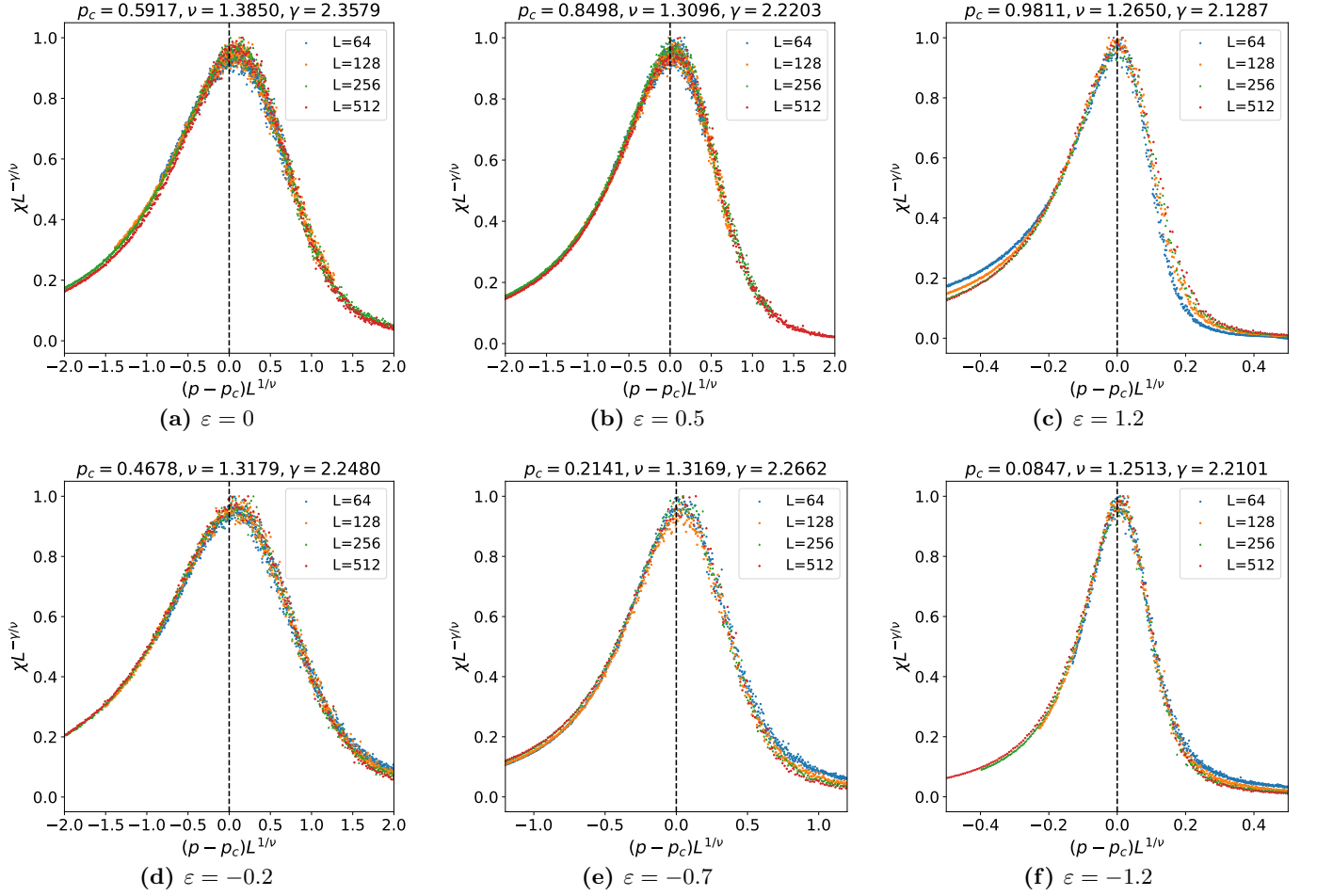


FIG. 8. The plot illustrates the scaling collapse of the susceptibility function  $\chi(p, L)$  for different system sizes  $L$ , computed from Metropolis Monte Carlo simulations. The critical exponents  $\nu = \nu(\varepsilon)$  and  $\gamma = \gamma(\varepsilon)$  has been extracted for various cases of bond energy cost  $\varepsilon$  as follows: (a)  $\varepsilon = 0$ , (b)  $\varepsilon = 0.5$ , (c)  $\varepsilon = 1.2$ , (d)  $\varepsilon = -0.2$ , (e)  $\varepsilon = -0.7$ , (f)  $\varepsilon = -1.2$ .

The partition function in Eq. (14) can be re-written as: and

$$Z = \sum_{s, \ell} \mathcal{N}(s, \ell) z^s K^\ell, \quad (26) \quad Z_2 = \sum_s \mathcal{N}(s, 2) z^s, \quad (30)$$

where  $\mathcal{N}(s, \ell)$  is the number of configurations with  $s$  occupied sites and  $\ell$  loops. Each loop contribute an extra factor of  $e^{-\varepsilon}$  to the partition function in Eq. (26), and  $z = pe^{-\varepsilon}$ . For  $K \rightarrow 0$ , only  $\ell = 0$  contributes:

$$Z_0 = \sum_s \mathcal{N}(s, 0) z^s. \quad (27)$$

Expanding the partition function in terms of the number of loops  $\ell$ , one can obtain, Expand:

$$Z = Z_0 + KZ_1 + K^2Z_2 + O(K^3), \quad (28)$$

where:

$$Z_1 = \sum_s \mathcal{N}(s, 1) z^s, \quad (29)$$

which are the one-loop and two-loop partition functions, respectively. The antiferromagnetic limit (as  $\varepsilon \rightarrow \infty$ ) and the ferromagnetic limit (as  $\varepsilon \rightarrow -\infty$ ) are particularly interesting. In the FK representation of the  $q$ -state Potts model, the partition function, as described in Eq. (14), can be analyzed in the limits of  $\varepsilon \rightarrow 0$  and  $\varepsilon \rightarrow \pm\infty$ . In these cases, it can be mapped to a loop model, whose critical behavior is characterized by a Coulomb gas (CG) or a conformal field theory with a dimensionless coupling constant in two dimensions (2D). This is because in the limit  $\varepsilon \rightarrow \pm\infty$ , the factor  $\prod_{i \in C} e^{-\varepsilon T_i}$  in Eq. (14) becomes  $e^{-\varepsilon C}$ , where  $C$  is the number of clusters in a particular configuration, since

$$\lim_{\varepsilon \rightarrow \infty} e^{-\varepsilon \sum_i T_i} = e^{-\varepsilon C}. \quad (31)$$

The Coulomb gas is a massless free scalar field  $\varphi(z, \bar{z})$

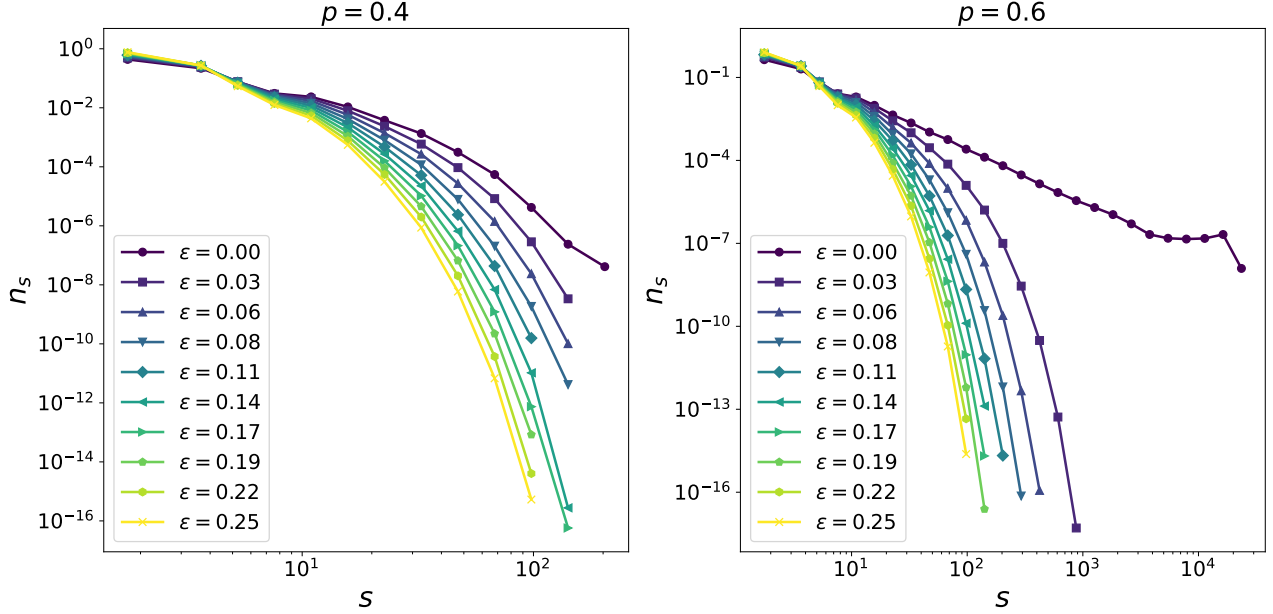


FIG. 9. The plot illustrates the weighted cluster-size distribution for various values of the nearest-neighbor occupied-site bond energy cost, denoted by  $\varepsilon$ , while keeping the site-occupation probability fixed at  $p$ . The x-axis label  $s$  is the cluster size. In both cases, it is evident that as  $\varepsilon$  increases, the cluster size decreases.

with action [29]

$$S = \frac{g}{4\pi} \int d^2r (\nabla\varphi)^2, \quad (32)$$

where  $g$  is the dimensionless coupling parameter, which acts like a stiffness constant controlling the strength of the logarithmic interaction between charges, in the Coulomb gas interpretation. The vertex operators are defined as [30, 31, 52]:

$$V_\alpha(z, \bar{z}) = e^{i\alpha\varphi(z, \bar{z})}. \quad (33)$$

The conformal weight  $h(\alpha)$  of a vertex operator  $V_\alpha$  in the presence of the background charge [53, 54] is

$$h(\alpha) = \frac{\alpha(\alpha - 2\alpha_0)}{4g}, \quad (34)$$

and the full scaling dimension [53–55] is:

$$x(\alpha) = 2h(\alpha) = \frac{\alpha(\alpha - 2\alpha_0)}{2g}. \quad (35)$$

The loop representation of the FK clusters assigns a weight  $n$  to each closed loop. The Coulomb gas analysis gives the exact mapping [29]:

$$n = -2 \cos(2\pi g), \quad (36)$$

where  $g \in [0, 1]$ . The total scaling dimension can be computed in terms of primary field indices  $(r, s)$  as [54]:

$$x_{r,s} = \frac{(r - sg)^2 - (1 - g)^2}{2g}. \quad (37)$$

In two dimensions, the RG eigenvalue  $y$  is related to the scaling dimension  $x$  by

$$y = 2 - x. \quad (38)$$

The thermal operator corresponds to  $(r, s) = (2, 1)$  in the Kac table [56–58]. Substituting into Eq. (37),

$$x_{2,1} = \frac{3}{2g} - 1. \quad (39)$$

The RG eigenvalue [59] is

$$y_t = 2 - x_{2,1} = 3 - \frac{3}{2g}. \quad (40)$$

The critical exponent  $\nu$  is just the reciprocal of this RG eigenvalue, i.e.,  $\nu = \frac{1}{y_t}$ . This form reproduces the exponents of the Potts and percolation universality classes. For percolation ( $\varepsilon = 0$ ), we have  $n = 1$ , which implies  $g = \frac{2}{3}$ , giving  $y_t = \frac{3}{4}$  and therefore  $\nu = \frac{4}{3}$ . In dilute isolated cluster limit ( $\varepsilon \rightarrow \infty$ ), we have  $n = 0$ , yielding  $g = \frac{3}{4}$ , which gives  $y_t = 1$  and, consequently,  $\nu = 1$ . In the dense cluster limit, as  $\varepsilon$  approaches  $-\infty$ , loops become abundant with a loop fugacity of  $n = 2$ . This results in the relation  $g = \frac{3}{2}$ , which indicates a relevant eigenvalue of  $y_t = 2$ . As a consequence, the thermal exponent is computed as  $\nu = \frac{1}{2}$ , which is consistent with the Ornstein-Zernike mean field result as well. Eq. (40) therefore interpolates naturally between dense ferromagnetic like cluster configuration ( $\nu = \frac{1}{2}$ ), ordinary percolation ( $\nu = 4/3$ ) and isolated cluster configuration ( $\nu = 1$ ) as shown in Figs. 6(a), 6(d), and 6(f) corresponding to the isotropic case, where the parameter  $\varepsilon$  is

constant and identical in both the horizontal and vertical directions and Figs. 17(b) and 18(b) corresponding to the anisotropic case, where  $\varepsilon$  is constant but takes different values in the horizontal and vertical directions and can also explain the finite behaviour of correlation length shown in Fig. 10, obtained using the expression for correlation length written in Canonical ensemble, Eq. (61) and using lattice gas Monte Carlo simulations as shown in Fig. 13. A thorough description of the continuous transition between the two extremes as  $\varepsilon$  varies is of interest; however we do not address this in this article.

## VI. CORRELATION LENGTH AND CLUSTER SIZE STATISTICS

### A. Energy Weighted Correlation Length

In ordinary site percolation on a square lattice, the critical point  $p_c \approx 0.5927$  [60] separates a phase with only finite clusters ( $p < p_c$ ) from one in which an infinite connected cluster exists ( $p > p_c$ ). The value of  $p_c$  in the case of bond percolation is analytically derived as  $p_c = 0.5$  [61, 62]. The divergence of the correlation length near the critical point characterizes the critical behavior,

$$\xi(p) \sim |p - p_c|^{-\nu}, \quad \nu = \frac{4}{3} \text{ in 2D.}$$

In a walk-based approach, we define the disorder-averaged generating function as:

$$\overline{F(R; \alpha, p)} = \sum_{T=0}^{T_{\max}} \alpha^{-T} \overline{N_T(R, p)}, \quad (41)$$

where  $\overline{N_T(R, p)}$  counts the number of walks of length  $T$  that end at Euclidean distance  $R$  from the origin, averaged over disorder realizations, and  $\alpha = e^\varepsilon$ . The same partition function for the case  $p = 1$  can be written as:

$$F(R, \alpha) = \sum_T N_T(R) \alpha^{-T}, \quad (42)$$

The parameter  $\alpha$  suppresses long walks: for  $\alpha > \alpha_c$  (the critical threshold for energy cost) the series converges, whereas as  $\alpha \rightarrow \alpha_c^+$  the suppression weakens and long walks proliferate. Thus,  $\alpha$  plays the role of a second control parameter in addition to the percolation probability  $p$ . The partition function presented here is written in a grand canonical formalism, which allows for fluctuations in site occupancy across different realizations. This introduces the fugacity parameter in the average  $\overline{N_T(R, p)}$ . In contrast, the partition function defined in Eq. (42) follows a canonical formalism, where no such fluctuations occur across realizations. From the pair connectedness distribution,  $G(R)$ , obtained from the partition function, one can define the second-moment correlation length [63] as:

$$\xi_2^2 = \frac{1}{2d} \frac{\mu_2}{\mu_0}, \quad \mu_n = \sum_R R^n G(R), \quad (43)$$

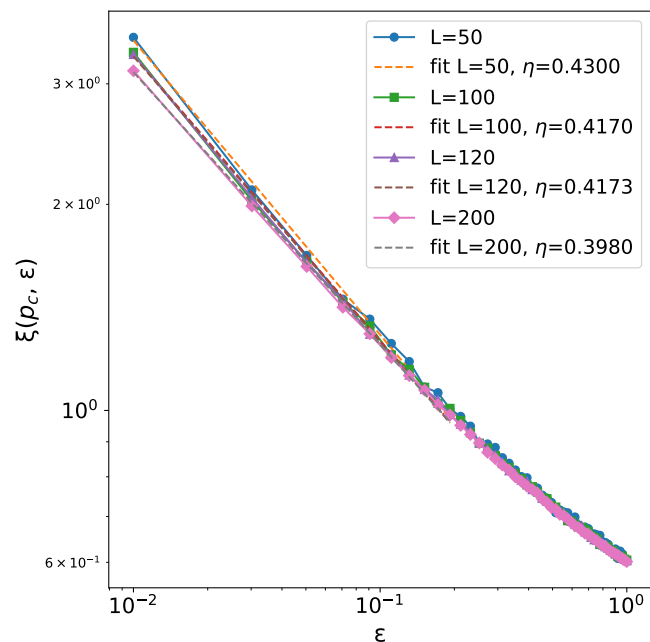


FIG. 10. Scaling of correlation length  $\xi(\varepsilon)$  computed using Eq. (61) for site percolation on a square lattice, with open boundary condition (OBC) at critical percolation threshold (for  $\varepsilon = 0$ ) with varying energy cost  $\varepsilon$ . We fit the numerical values with  $\varepsilon$  as  $\xi(\varepsilon, p_c(\varepsilon = 0)) = A\varepsilon^{-\eta}$ . The legend in the figure indicates the values of the exponent  $\eta$  for different system sizes  $L$ .

which coincides with the standard second-moment definition of correlation length used in statistical mechanics.

In standard site percolation on a lattice, each site is independently occupied with probability  $p$ , and the geometry of clusters is therefore purely random. Let  $k$  index the connected clusters of occupied sites in a given configuration. Each cluster  $k$  contains  $s_k$  occupied sites and  $T_k$  bonds between nearest neighbor occupied sites. We now generalize the percolation problem by assigning an energy cost  $\varepsilon$  to each bond between occupied sites. A cluster with  $T_k$  bonds then carries a total energy:

$$E_k = \varepsilon T_k. \quad (44)$$

Clusters with many bonds (large  $T_k$ ) are thus energetically penalized. In statistical mechanical language, for each cluster labelled as  $k$ , there is a total of  $s_k$  number of microstate degeneracies contributing to the same energy  $E_k = \varepsilon T_k$ , so the probability for a randomly chosen site in a system to belong to cluster  $k$  is:

$$P_\varepsilon^k = \frac{s_k e^{-E_k}}{\sum_k s_k e^{-E_k}}, \quad (45)$$

where  $s_k$  is the count of microstates or degeneracy associated with energy state  $k$ , and  $e^{-E_k}$  is the Boltzmann weight associated with the canonical ensemble of clusters or energy states. Here  $E_k$  is the total energy of the energy

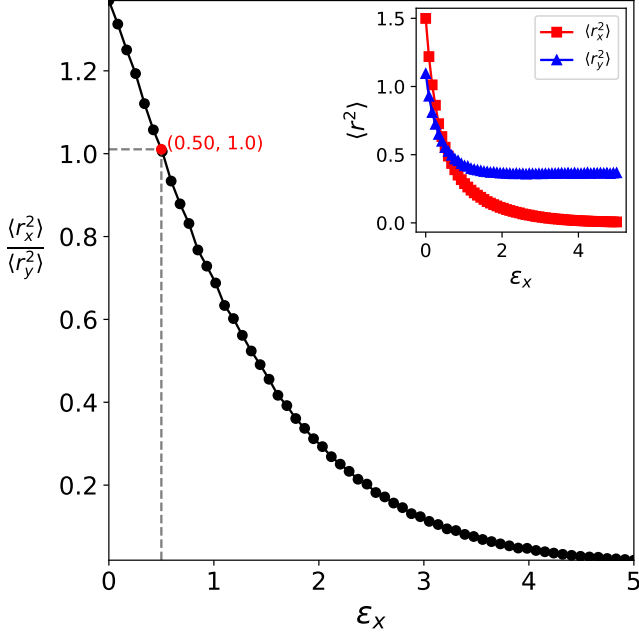


FIG. 11. The plot shows the ratio of the mean-squared radii in the  $x$  and  $y$  directions for the energy-weighted cluster ensemble. This ratio is shown relative to the anisotropic energy costs, with  $\varepsilon_x$  denoting the cost in the  $x$  direction and  $\varepsilon_y = 0.5$  in the  $y$  direction. The inset illustrates the mean square radius in the  $x$  and  $y$  directions. The site-occupation probability is fixed at  $p = 0.6$ , and the system size is  $N = 1600$ . The mean square radius in both directions is computed using Eq. (46) and the equilibration is reached by performing a Glauber-like Metropolis sweep. At  $\varepsilon_x = 0.5$ , the two inset curves of mean-squared radii in the  $x$  and  $y$  directions cross each other. This indicates that at  $\varepsilon_x = \varepsilon_y = \varepsilon = 0.5$ , the anisotropic ratio is equal to 1, as can be observed in the main plot.

state  $k$  or  $E_{\text{cluster},k}$  for cluster  $k$ . Here  $\beta$ , the inverse temperature parameter is set as 1. This construction introduces an additional control parameter  $\varepsilon$  that suppresses large connected clusters. In the absence of an energy penalty ( $\varepsilon = 0$ ), the system reduces to standard percolation, whereas for finite  $\varepsilon > 0$  the formation of large clusters is exponentially suppressed; in the strongly penalized regime  $\varepsilon \gg \varepsilon_c$ , the system is therefore dominated by small, isolated clusters. Above a critical value  $\varepsilon_c$ , the correlation length becomes finite even if  $p \approx p_c$ . Here  $p_c$  denotes the critical percolation threshold for  $\varepsilon = 0$ . In classical percolation, the correlation length is obtained from cluster statistics as:

$$\xi^2 = \frac{\sum_k s_k \langle r^2 \rangle_k}{\sum_k s_k}, \quad (46)$$

where,  $\langle r^2 \rangle_k$ , the mean-square distance of sites within cluster  $k$  from its center of mass is defined as:

$$\langle r^2 \rangle_k = \frac{1}{s_k} \sum_{i=1}^{s_k} |\mathbf{r}_i - \mathbf{r}_{\text{cm},k}|^2, \quad \mathbf{r}_{\text{cm},k} = \frac{1}{s_k} \sum_{i=1}^{s_k} \mathbf{r}_i, \quad (47)$$

and the weighting is done in terms of cluster size or number of sites in the cluster. This results from the probability that the chosen site belongs to cluster  $k$  is

$$\mathbb{P}(\text{site} \in k) = \frac{s_k}{\sum_j s_j}, \quad (48)$$

and the correlation length is just the mean of the mean-square distance of sites within a cluster.

The correlation length result presented in Eq. (46) can be readily reduced to the standard definition found in the literature [64], which involves the number of clusters of size  $s$  per site, denoted as  $n_s(p)$  or the cluster size strength. We can reorganize the summation over clusters into a summation over cluster sizes as:

$$\sum_k s_k = \sum_s \sum_{k=1}^{N_s} s = \sum_s s N_s, \quad (49)$$

where,  $N_s$  is the total of clusters of size  $s$ . Using  $N_s = N n_s$ , where  $N$  is the total number of clusters of all sizes, this can be written as:

$$\sum_k s_k = N \sum_s s n_s. \quad (50)$$

Hence, the radius of gyration squared summed over each cluster can be written as:

$$\sum_k s_k \langle r^2 \rangle_k = \sum_s \sum_{k=1}^{N_s} s \langle r^2 \rangle_k. \quad (51)$$

We next define the average radius of gyration squared for clusters of size  $s$ :

$$\langle R^2 \rangle_s = \frac{1}{N_s} \sum_{k=1}^{N_s} \langle r^2 \rangle_k. \quad (52)$$

Then,

$$\sum_{k=1}^{N_s} s \langle r^2 \rangle_k = s N_s \langle R^2 \rangle_s. \quad (53)$$

Thus, the radius of gyration squared summed over each cluster can be written as:

$$\sum_k s_k \langle r^2 \rangle_k = \sum_s s N_s \langle R^2 \rangle_s = N \sum_s s n_s \langle R^2 \rangle_s. \quad (54)$$

Substituting the expressions in Eq. (50) and Eq. (54) into Eq. (46), we can obtain the standard expression for correlation length as:

$$\xi^2 = \frac{N \sum_s s n_s \langle R^2 \rangle_s}{N \sum_s s n_s} = \frac{\sum_s s n_s \langle R^2 \rangle_s}{\sum_s s n_s}, \quad (55)$$

where the sum runs over finite clusters only.

The correlation length can also be expressed in terms of the pair-connectedness function. The connectedness

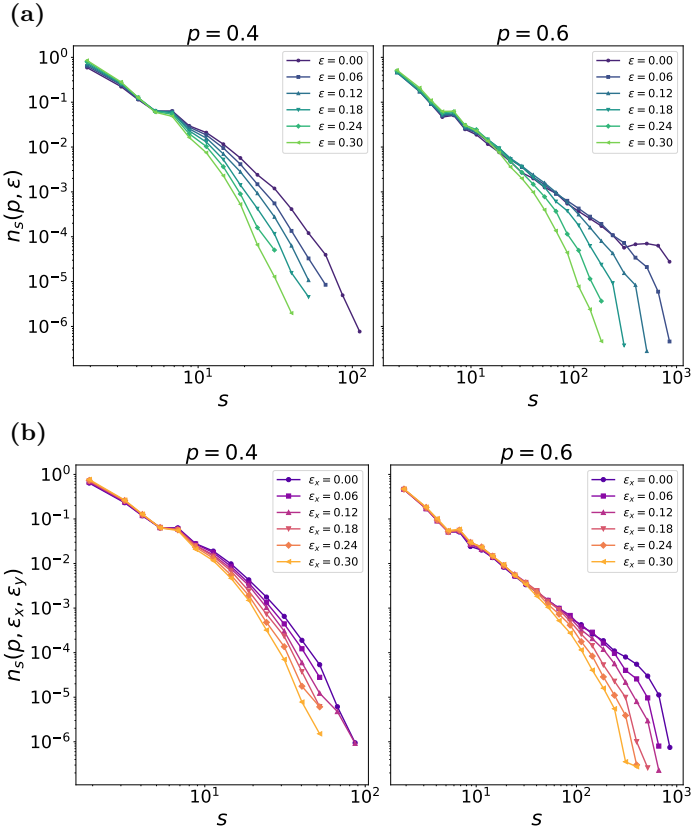


FIG. 12. Cluster-size distributions obtained from Glauber-like Metropolis sweep as described in Sec. III for a square lattice of size  $N = 1600$ . The top figure (a) shows the distributions for isotropic bond energy costs  $\epsilon$  for different site occupation probabilities  $p = 0.4$  and  $p = 0.6$ . The bottom figure (b) shows the corresponding distributions for anisotropic bond energy costs, with  $\epsilon_x$  and  $\epsilon_y$  associated with bonds along the  $x$  and  $y$  directions, respectively. The bond energy cost along the  $y$ -direction is taken as  $\epsilon_y = 0.1$ . The bond energy values along the  $x$  direction are indicated in the legends.

function  $G(\mathbf{r})$  is defined as the probability that two sites separated by vector  $\mathbf{r}$  lie in the same cluster.

The second-moment correlation length is often defined by as Eq. (43):

$$\xi^2 = \frac{\sum_{\mathbf{r}} |\mathbf{r}|^2 G(\mathbf{r})}{\sum_{\mathbf{r}} G(\mathbf{r})}, \quad (56)$$

where  $d$  is the spatial dimension. We now relate this to cluster sums for a single configuration. Summing over all sites pairs in a given configuration, one can have:

$$\sum_{i,j} \mathbf{1}_{ij} = \sum_k \sum_{i \in k} \sum_{j \in k} 1 = \sum_k s_k^2, \quad (57)$$

and

$$\sum_{i,j} |\mathbf{r}_j - \mathbf{r}_i|^2 \mathbf{1}_{ij} = \sum_k \sum_{i \in k} \sum_{j \in k} |\mathbf{r}_j - \mathbf{r}_i|^2, \quad (58)$$

and performing the summation in Eq. (58) over  $i \in k$  and  $j \in k$ , [64]:

$$\sum_{i \in k} \sum_{j \in k} |\mathbf{r}_j - \mathbf{r}_i|^2 = 2 s_k \sum_{i \in k} |\mathbf{r}_i - \mathbf{r}_{\text{cm},k}|^2 = 2 s_k \cdot s_k \cdot \langle r^2 \rangle_k. \quad (59)$$

Therefore, inserting these into Eq. (56), yields a cluster representation with  $s_k^2$  weights:

$$\xi^2(\epsilon = 0) = \frac{\sum_k s_k^2 \langle r^2 \rangle_k}{\sum_k s_k^2}. \quad (60)$$

When the energy cost per link is introduced, each cluster is weighted by its Boltzmann factor  $e^{-\epsilon T_k}$  following Eq. (45), yielding the generalized, energy-weighted correlation length per realization of occupation probability ( $p$ ):

$$\xi^2(\epsilon) = \frac{\sum_k s_k e^{-\epsilon T_k} \langle r^2 \rangle_k}{\sum_k s_k e^{-\epsilon T_k}}. \quad (61)$$

This definition reduces to the usual percolation result (46) when  $\epsilon = 0$ . The above result Eq. (61) is the energy-weighted correlation length for a particular realization, which can be treated in the canonical formalism. For an infinite system size, near the critical occupation probability  $p_c(\epsilon)$ , the correlation length  $\xi(\epsilon)$  diverges as [64]:

$$\xi(\epsilon) \sim |p - p_c(\epsilon)|^{-\nu(\epsilon)}. \quad (62)$$

For finite system size  $L$ ,

$$\xi(p, \epsilon) \sim L \sim |p - p_c(\epsilon)|^{-\nu(\epsilon)}, \quad (63)$$

which implies

$$|p - p_c(\epsilon)| L^{\frac{1}{\nu(\epsilon)}} \sim 1. \quad (64)$$

The plot of the ratio  $\frac{\langle r^2 \rangle}{\langle r^2 \rangle}$ , computed using Eq. (61), as shown in Fig. 11, indicates that as the anisotropy in bond energy increases, the clusters effectively become one-dimensional. Summing Eq. (61) over the fluctuations of site occupation across each realization will yield the complete energy-weighted correlation length in a grand canonical ensemble, similar in fashion to that which can be obtainable from the partition function expressed in Eq. (41).

## B. Energy Weighted Cluster Size Statistics

For  $\epsilon = 0$ , large clusters contribute significantly, leading to a diverging correlation length as  $p \rightarrow p_c(\epsilon = 0)$ . As  $\epsilon$  increases, the contribution of large clusters is exponentially suppressed, and the effective correlation length

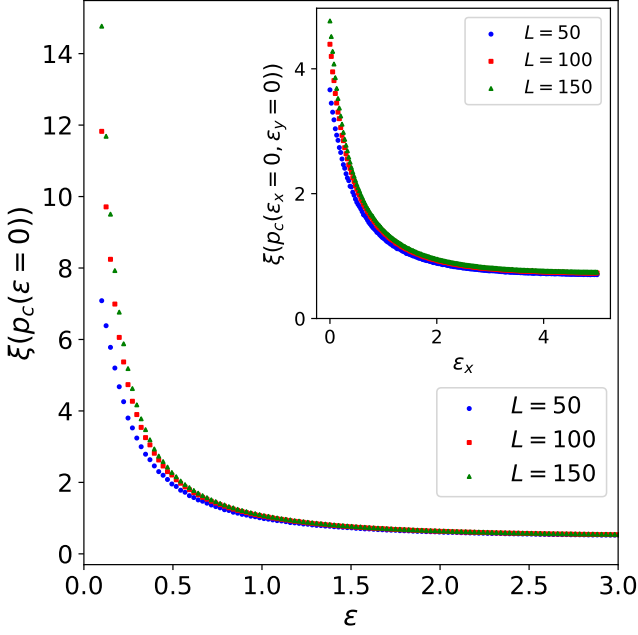


FIG. 13. Correlation length for site percolation on a square lattice with open boundary conditions at the critical percolation threshold  $p_c(\varepsilon = 0)$ , computed using Eq. (46). The main plot shows the variation of  $\xi(p_c(\varepsilon = 0))$  with the energy cost  $\varepsilon$  for different system sizes  $L$ . The inset plot shows the variation of  $\xi(p_c(\varepsilon_x = 0, \varepsilon_y = 0))$  with anisotropic energy cost  $\varepsilon_x$  for fixed  $\varepsilon_y = 0.5$  and different  $L$ . The data are obtained from Monte Carlo simulations in the grand canonical ensemble as described in Sec. III. The system size used is  $N = 1600$ .

$\xi(\varepsilon)$  decreases sharply. For  $\varepsilon > \varepsilon_c(p)$ , the system is dominated by small clusters and  $\xi(\varepsilon)$  becomes finite even at  $p = p_c(\varepsilon = 0)$ , as shown in Fig. 13. Near the percolation threshold  $p_c$ , the cluster size distribution follows [64]:

$$n_s(p) \sim s^{-\tau} e^{-s/s^*(p)}, \quad (65)$$

where the size  $s^*(p)$  scales as

$$s^*(p) \sim |p - p_c|^{-1/\sigma}, \quad (66)$$

where the critical exponent is calculated as  $\sigma = \frac{1}{\nu D}$ , with  $D$  as the fractal dimension of the lattice, and  $\nu$  as defined previously, the critical exponent for correlation length. Now, suppose each bond between nearest-neighbor occupied sites carries an additional energy cost  $\varepsilon$ . Then a cluster of size  $s$  acquires a Boltzmann weight  $e^{-\varepsilon s}$ . The modified cluster size distribution takes the form

$$n_s(p, \varepsilon) \sim s^{-\tau} e^{-s/s^*(p, \varepsilon)}, \quad (67)$$

where the effect of bond energy cost impacts only the cut-off cluster size; thus, the overall cluster size distribution retains the same form. From Fig. 9 one can verify that the energy-dependent cut-off cluster size  $s^*(p, \varepsilon)$  decreases with increasing bond energy cost  $\varepsilon$ . In the full

grand canonical ensemble, as shown in Fig. 12, the cluster size cut-off diverges near the critical occupation threshold  $p_c(\varepsilon)$  as:

$$s^*(p, \varepsilon) \sim |p - p_c(\varepsilon)|^{-\frac{1}{\sigma(\varepsilon)}}, \quad (68)$$

where  $\sigma(\varepsilon)$  is the cut-off cluster size critical exponent as mentioned before. In this article, we have not included the observed values for  $\sigma(\varepsilon)$  at varying bond energy costs  $\varepsilon$ . This is because we have already presented the relationship between the thermal exponent  $\nu(\varepsilon)$  and the bond energy cost  $\varepsilon$ . Additionally, the cluster cut-off size critical exponent is not an independent critical exponent; it inversely depends on the thermal critical exponent by a constant proportionality factor, which is a function of the fractal dimension  $D$ , as specified earlier.

## VII. RENORMALIZATION GROUP ANALYSIS FOR THE LATTICE GAS

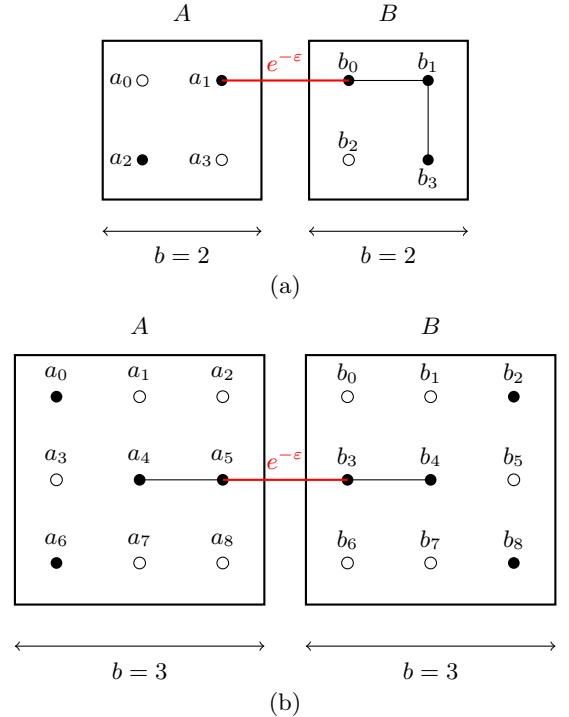


FIG. 14. Schematic diagram showing two neighboring Kadanoff blocks, labeled  $A$  and  $B$ . Filled circles denote occupied sites (weight  $e^\mu$ ), open circles denote empty sites, and bonds carry weight  $e^J = e^{-\varepsilon}$ . The red bond illustrates an inter-block connection contributing to coarse-grained bond energy. (a) Block size  $b = 2$ . (b) Block size  $b = 3$ .

We have a site percolation on a square lattice where each bond between two nearest-neighbor occupied sites has an energy cost  $\varepsilon$ . A cluster of  $T$  bonds contributes weight

$$W_T = e^{-\varepsilon T}. \quad (69)$$

Now we introduce fugacity  $e^\mu = p/(1-p)$  for sites and  $e^J = e^{-\varepsilon}$  for bonds. Consider two nearest neighbor Kadanoff blocks  $A$  and  $B$  (after rescaling the original lattice into blocks of linear dimensions  $b$ ) [32, 46, 47], the weight of a block  $A$  is given as:

$$W_A = e^{\mu k_A + J b_A}, \quad k_A = \sum_{i \in A} a_i, \quad b_A = \sum_{(ij) \in A} a_i a_j. \quad (70)$$

The sites of the  $A$  block (with  $b = 2$ ) are numbered as 0, 1, 2, 3 as shown in Fig. 14, and the internal bonds are given by (0, 1), (0, 2), (1, 3), (2, 3) for a fully occupied block. Fixing the right-column boundary sites ( $a_1, a_3$ ) and sum over the rest sites ( $a_0, a_2$ ) gives the weights corresponding to boundary sites:

$$\begin{aligned} Q(a_1, a_3) &= \sum_{a_0, a_2 \in \{0, 1\}} e^{\mu k_A + J b_A}, \\ k_A &= a_1 + a_3 + a_0 + a_2, \\ b_A &= a_1 a_3 + a_0 a_1 + a_0 a_2 + a_2 a_3, \end{aligned} \quad (71)$$

Here  $a_0, a_1, a_2$  and  $a_3$  label the occupancy variables for sites 0, 1, 2 and 3 for  $A$  block respectively as shown in Fig. 14. Explicitly evaluating the weights gives:

$$\begin{aligned} Q(0, 0) &= 2e^\mu + e^{2\mu} e^J, \\ Q(1, 0) &= Q(0, 1) = e^\mu + e^{2\mu}(1 + e^J) + e^{3\mu} e^{2J}, \\ Q(1, 1) &= e^{2\mu} e^J + 2e^{3\mu} e^{2J} + e^{4\mu} e^{4J}. \end{aligned} \quad (72)$$

We denote  $Q(0, 0) = Q_{00}$ ,  $Q(1, 0) = Q(0, 1) = Q_1$ , and  $Q(1, 1) = Q_{11}$  for the sake of convenience. The combined weights of two nearest neighbor blocks  $A$  and  $B$ , denoted as  $W_{N_A, N_B}$ , where  $N_A$  and  $N_B$  denotes the occupancy variables of blocks  $A$  and  $B$  respectively. The weights can be computed as:

$$W_{00} = 1, \quad (73)$$

$$\begin{aligned} W_{10} &= Q_{00} + 2Q_1 + Q_{11} = 4e^\mu + 2e^{2\mu} + 4e^{2\mu} e^J \\ &\quad + 4e^{3\mu} e^{2J} + e^{4\mu} e^{4J}. \end{aligned} \quad (74)$$

$$W_{11} = \sum_{a_1, a_3, b_0, b_2} Q(a_1, a_3) Q(b_0, b_2) e^{J(a_1 b_0 + a_3 b_2)}, \quad (75)$$

Grouping the terms in the above expression by boundary configurations, the expression simplifies to:

$$\begin{aligned} W_{11} &= Q_{00}(Q_{00} + 2Q_1 + Q_{11}) \\ &\quad + 2Q_1(Q_{00} + Q_1(1 + e^J) + Q_{11}e^J) \\ &\quad + Q_{11}(Q_{00} + 2Q_1e^J + Q_{11}e^{2J}) \end{aligned} \quad (76)$$

Substituting the  $Q$  values and collecting terms for each power of  $e^\mu$  and  $e^J$  yields a finite polynomial expansion

as:

$$\begin{aligned} W_{11} &= 14e^{2\mu} + 2e^{2\mu} e^J + 12e^{3\mu} + 28e^{3\mu} e^J + 8e^{3\mu} e^{2J} \\ &\quad + 2e^{4\mu} + 10e^{4\mu} e^J + 37e^{4\mu} e^{2J} + 18e^{4\mu} e^{3J} + e^{4\mu} e^{4J} \\ &\quad + 4e^{5\mu} e^{2J} + 24e^{5\mu} e^{3J} + 20e^{5\mu} e^{4J} + 8e^{5\mu} e^{5J} \\ &\quad + 2e^{6\mu} e^{4J} + 16e^{6\mu} e^{5J} + 8e^{6\mu} e^{6J} + 2e^{6\mu} e^{7J} \\ &\quad + 4e^{7\mu} e^{7J} + 4e^{7\mu} e^{8J} + e^{8\mu} e^{10J}. \end{aligned} \quad (77)$$

After rescaling, the weight for the two nearest neighbor blocks  $A$  and  $B$  can be written as:

$$W_{N_A, N_B} = e^{\tilde{\mu}(N_A + N_B) + \tilde{J} N_A N_B}, \quad (78)$$

where  $\tilde{\mu}$  and  $\tilde{J}$  are the renormalized site fugacity and bond energy respectively. The RG flow equation for the renormalized site occupation probability and the renormalized energy cost can be derived by comparing the derived weights with those given in Eq. (78):

$$p' = \frac{W_{10}(p, y)}{1 + W_{10}(p, y)} \quad (79)$$

$$J' = \ln \left( \frac{W_{11}(p, y)}{W_{10}^2(p, y)} \right), \quad (80)$$

where  $W_{10}, W_{11}$  are explicit polynomials in  $e^\mu = p/(1-p)$  and  $y = e^J$  as derived above. In a real-space RG, we coarse-grain the lattice by a factor  $b$  (here  $b = 2$  for  $2 \times 2$  blocks). At the critical fixed point  $(\mu^*, J^*)$ , the system is scale-invariant [32]. Small deviations from the fixed point obey the linearized RG map as:

$$\begin{pmatrix} \delta\mu' \\ \delta J' \end{pmatrix} = \mathcal{J} \begin{pmatrix} \delta\mu \\ \delta J \end{pmatrix}, \quad \mathcal{J} = \left. \frac{\partial(\mu', J')}{\partial(\mu, J)} \right|_{(\mu^*, J^*)}, \quad (81)$$

where  $\delta\mu = \mu - \mu^*$ ,  $\delta J = J - J^*$ . Let  $\lambda_{\text{rel}}$  be the eigenvalue corresponding to the relevant direction (growing under RG), so that

$$\delta' = \lambda_r \delta. \quad (82)$$

The correlation length  $\xi$  transforms under RG as:

$$\xi(\delta') = \frac{\xi(\delta)}{b}. \quad (83)$$

Near criticality,  $\xi \sim \delta^{-\nu}$ . Using  $\delta' = \lambda_r \delta$ , we have

$$\xi(\delta') = (\delta')^{-\nu} = (\lambda_r \delta)^{-\nu} = \lambda_r^{-\nu} \delta^{-\nu}. \quad (84)$$

Comparing with  $\xi(\delta') = \xi(\delta)/b = \delta^{-\nu}/b$  gives:

$$\lambda_r^{-\nu} = \frac{1}{b} \quad \Rightarrow \quad \nu = \frac{\ln b}{\ln \lambda_r}. \quad (85)$$

Let  $X \equiv W_{10}$ ,  $Y \equiv W_{11}$ . Then

$$\mu' = \ln X, \quad J' = \ln Y - 2 \ln X.$$

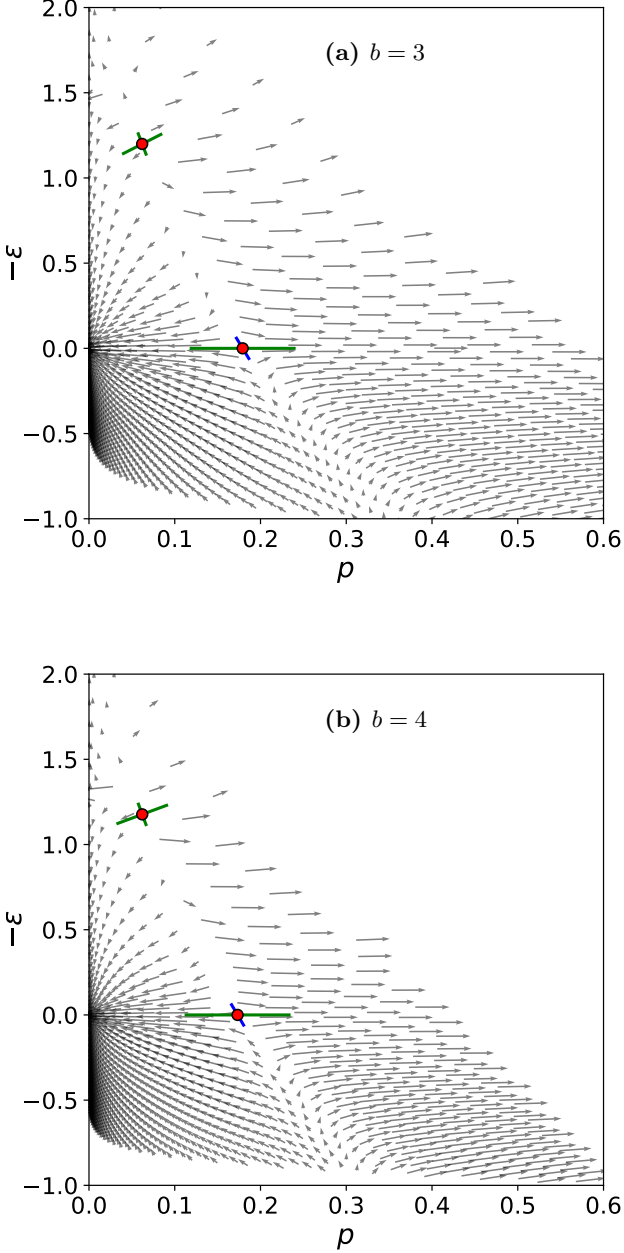


FIG. 15. RG flows around the fixed points, where both the site occupation probability  $p$  and the energy cost  $\varepsilon$  are renormalized for block sizes  $b = 3, 4$ , respectively. The y-axis represents the energy cost  $-\varepsilon$ . The linearized flow around each fixed point is depicted by blue and green lines. The green lines indicate the relevant direction, while the blue lines represent the irrelevant direction. The overall flow is illustrated with black arrows.

The Jacobian entries at any point  $(\mu, J)$  can be computed

as:

$$\begin{aligned}\frac{\partial \mu'}{\partial \mu} &= \frac{\tilde{X}_\mu}{X}, \\ \frac{\partial \mu'}{\partial J} &= \frac{\tilde{X}_J}{X}, \\ \frac{\partial J'}{\partial \mu} &= \frac{\tilde{Y}_\mu}{Y} - 2\frac{\tilde{X}_\mu}{X}, \\ \frac{\partial J'}{\partial J} &= \frac{\tilde{Y}_J}{Y} - 2\frac{\tilde{X}_J}{X},\end{aligned}\quad (86)$$

where

$$\begin{aligned}\tilde{X}_\mu &= e^\mu \frac{\partial X}{\partial e^\mu}, \\ \tilde{X}_J &= e^J \frac{\partial X}{\partial e^J}, \\ \tilde{Y}_\mu &= e^\mu \frac{\partial Y}{\partial e^\mu}, \\ \tilde{Y}_J &= e^J \frac{\partial Y}{\partial e^J}.\end{aligned}\quad (87)$$

The RG eigenvalue  $\{\lambda_i\}$  can be evaluated from Eq. (86) at fixed point  $(\mu^*, J^*)$ . The relevant eigenvalue  $\lambda_{\text{rel}}$  gives the correlation-length exponent similar to that done in Sec. IV, with

$$\nu = \frac{\ln b}{\ln \lambda_r}. \quad (88)$$

In our original model,  $\varepsilon$  represents a constant energy cost at all length scales, which we will later relax to consider the renormalization of the energy cost as well. Therefore, we focus on the flow equation for the site-occupation probability, as derived in Eq. (79). For  $b = 2$ , the way we derived Eq. (74) indicates that a block is considered occupied if at least one site within it is occupied. When we solve the flow equation for the fixed point, we find only the trivial fixed point  $p_c = 1$  and do not obtain any non-trivial fixed points for values of  $\varepsilon$ . For that reason, while deriving Eq. (74), we now only consider configurations that span the block along at least one direction. To compute the modified weight for a block size of  $b = 2$ , we can use Eq. (75) along with the modified boundary weights computed as follows:

$$\begin{aligned}\tilde{Q}_{00} &= e^{2\mu} e^J, \\ \tilde{Q}_{10} = \tilde{Q}_{01} &= e^{2\mu} e^J + e^{3\mu} e^{2J}, \\ \tilde{Q}_{11} &= e^{2\mu} e^J + 2e^{3\mu} e^{2J} + e^{4\mu} e^{4J}.\end{aligned}\quad (89)$$

Taking this into account, the modified weights become:

$$\tilde{W}_{10}^2 = 4e^{2\mu} e^J + 4e^{3\mu} e^{2J} + e^{4\mu} e^{4J}, \quad (90)$$

where  $\tilde{W}^2$  indicates the modified weights for block size  $b = 2$ . For block size  $b = 3$ , the modified boundary

weights can be computed similarly as:

$$\begin{aligned}
\tilde{Q}(0,0,0) &= 2y^2z^3 + 8y^3z^4 + 2y^4z^5 + 4y^5z^5 + y^7z^6, \\
\tilde{Q}(1,0,0) &= y^2z^3 + 3y^2z^4 + 4y^3z^4 + 6y^3z^5 + 5y^4z^5 + y^5z^5 \\
&\quad + 3y^5z^6 + 3y^6z^6 + y^8z^7, \\
\tilde{Q}(0,1,0) &= y^2z^3 + y^2z^4 + 7y^3z^4 + 4y^3z^5 + 6y^4z^5 + y^4z^6 \\
&\quad + 2y^5z^5 + y^5z^6 + 4y^6z^6 + y^8z^7, \\
\tilde{Q}(0,0,1) &= y^2z^3 + 3y^2z^4 + 4y^3z^4 + 6y^3z^5 + 5y^4z^5 + y^5z^5 \\
&\quad + 3y^5z^6 + 3y^6z^6 + y^8z^7, \\
\tilde{Q}(0,1,1) &= 3y^3z^4 + 4y^3z^5 + 5y^4z^5 + 4y^4z^6 + 3y^5z^5 \\
&\quad + 4y^5z^6 + 5y^6z^6 + y^6z^7 + y^7z^6 + 2y^7z^7 \\
&\quad + 3y^8z^7 + y^{10}z^8, \\
\tilde{Q}(1,1,0) &= 3y^3z^4 + 4y^3z^5 + 5y^4z^5 + 4y^4z^6 + 3y^5z^5 \\
&\quad + 4y^5z^6 + 5y^6z^6 + y^6z^7 + y^7z^6 + 2y^7z^7 \\
&\quad + 3y^8z^7 + y^{10}z^8, \\
\tilde{Q}(1,0,1) &= 2y^2z^4 + 3y^2z^5 + 8y^3z^5 + 3y^3z^6 + y^4z^5 + 7y^4z^6 \\
&\quad + 5y^5z^6 + 4y^6z^7 + 2y^7z^7 + y^9z^8, \\
\tilde{Q}(1,1,1) &= y^2z^3 + 3y^2z^4 + y^2z^5 + 3y^3z^4 + 8y^3z^5 + y^3z^6 \\
&\quad + 4y^4z^5 + 6y^4z^6 + 2y^5z^5 + 8y^5z^6 + 4y^6z^6 \\
&\quad + 8y^6z^7 + y^7z^6 + 2y^7z^7 + 5y^8z^7 + y^8z^8 \\
&\quad + 3y^9z^8 + 2y^{10}z^8 + y^{12}z^9,
\end{aligned} \tag{91}$$

with permutation symmetry  $\tilde{Q}(1,0,0) = \tilde{Q}(0,0,1)$ , and  $\tilde{Q}(1,1,0) = \tilde{Q}(0,1,1)$  respectively. A similar calculation for block size  $b = 3$  yields the following expansion for modified weight:

$$\begin{aligned}
\tilde{W}_{10}^3(z, y) &= 6z^3y^2 + 12z^4y^2 + 32z^4y^3 \\
&\quad + 4z^5y^2 + 40z^5y^3 + 33z^5y^4 + 16z^5y^5 \\
&\quad + 4z^6y^3 + 22z^6y^4 + 28z^6y^5 + 24z^6y^6 + 4z^6y^7 \\
&\quad + 14z^7y^6 + 8z^7y^7 + 14z^7y^8 \\
&\quad + z^8y^8 + 4z^8y^9 + 4z^8y^{10} + z^9y^{12},
\end{aligned} \tag{92}$$

where  $z = e^\mu$  and  $y = e^{-\varepsilon}$  respectively. Setting  $y = 1$ , i.e.  $\varepsilon = 0$ , the coefficients in the above expansion, Eq. (92) reduces to that in Eq. (18), and that given in [48]. Thus, one can reproduce the same RG flow equations directly from a hard-core lattice gas analogy relevant to our model and Monte Carlo simulation. The critical threshold for site occupation probability is calculated using the recursion relation outlined in Eq. (79). This calculation is performed for various bond energy costs, denoted as  $\varepsilon$ , and for block sizes of  $b = 2$  and  $b = 3$ . The corresponding weights used in this analysis are given in Eq. (90) and Eq. (92). The results are illustrated in Fig. 16.

One significant advantage of this technique arises when

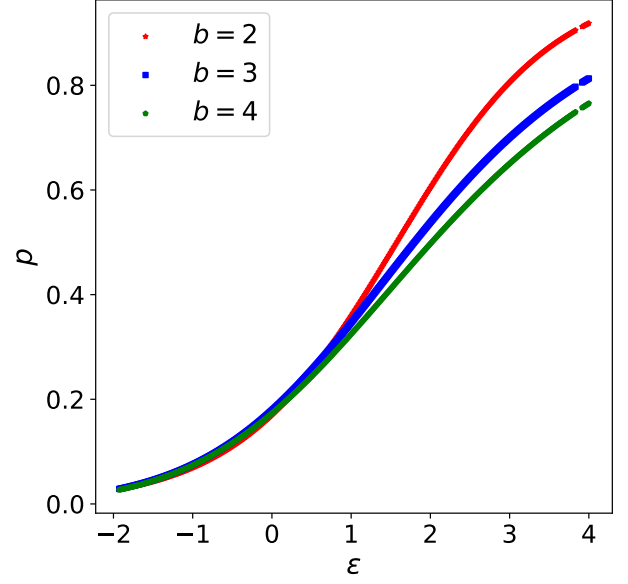


FIG. 16. Variation of the critical site percolation threshold, calculated from renormalization group analysis in real space using Kadanoff blocks with a lattice gas analogy, as described in Eq. (79), considering an energy cost per link  $\varepsilon$  for different block sizes,  $b$  as  $b = 2$ ,  $b = 3$  and  $b = 4$  as derived in Eq. (90), Eq. (92) and Eq. S25 given in the Supplementary Information [45]. We can infer from the plot that as  $\varepsilon$  increases, the critical occupation probability  $p_c$  tends to unity.

the energy cost  $\varepsilon$  is not constant but renormalizable, similar to the probability of site occupation. In this case, it is possible to express the weights  $W_{11}$  as derived in Eq. (77) and the modified weight  $\tilde{W}_{11}$ . By solving the coupled polynomial equations in Eq. (79) and Eq. (80), we can determine the fixed point for  $p$ , and  $\varepsilon$  flows.

Using this modified boundary weights derived in Eq. (89), we can determine the modified weight  $\tilde{W}_{11}^2$  for block size  $b = 2$  similar to that in Eq. (92) following Eq. (75) as:

$$\begin{aligned}
\tilde{W}_{11}^2(z, y) &= 9y^2z^4 + 6y^3z^4 + 12y^3z^5 + y^4z^4 + 16y^4z^5 + 2y^4z^6 \\
&\quad + 4y^5z^5 + 12y^5z^6 + 8y^6z^6 + 2y^7z^6 \\
&\quad + 4y^7z^7 + 4y^8z^7 + y^{10}z^8.
\end{aligned} \tag{93}$$

where  $z = e^\mu$ , and  $y = e^{-\varepsilon}$  as defined previously. Similarly, the modified block weight  $\tilde{W}_{11}^3$  for block size  $b = 3$  can be computed as:

$$\tilde{W}_{11}^3 = \sum_{\substack{a_2, a_5, a_8 \in [0,1] \\ b_0, b_3, b_6 \in [0,1]}} Q(a_2, a_5, a_8) Q(b_0, b_3, b_6) e^{J(a_2b_0 + a_5b_3 + a_8b_6)}. \tag{94}$$

By explicitly computing the summation using the modified boundary weights derived in Eq. (91), one can arrive

at,

$$\begin{aligned}
\tilde{W}_{11}^3(z, y) = & z^{18}y^{27} + 4z^{17}y^{25} + 10z^{17}y^{24} + 4z^{17}y^{23} \\
& + 14z^{16}y^{23} + 38z^{16}y^{22} + 65z^{16}y^{21} + 33z^{16}y^{20} \\
& + 3z^{16}y^{19} + 2z^{15}y^{22} + 32z^{15}y^{21} + 130z^{15}y^{20} \\
& + 232z^{15}y^{19} + 268z^{15}y^{18} + 132z^{15}y^{17} \\
& + 8z^{14}y^{20} + 83z^{14}y^{19} + 296z^{14}y^{18} + 693z^{14}y^{17} \\
& + 852z^{14}y^{16} + 716z^{14}y^{15} + 282z^{14}y^{14} \\
& + 24z^{13}y^{18} + 178z^{13}y^{17} + 628z^{13}y^{16} \\
& + 2088z^{13}y^{14} + 1928z^{13}y^{13} + 1054z^{13}y^{12} \\
& + 20z^{13}y^{10} + 3z^{12}y^{17} + 46z^{12}y^{16} + 342z^{12}y^{15} \\
& + 1135z^{12}y^{14} + 2638z^{12}y^{13} + 3745z^{12}y^{12} \\
& + 2033z^{12}y^{10} + 556z^{12}y^9 + 85z^{12}y^8 + 6z^{12}y^7 \\
& + 8z^{11}y^{15} + 96z^{11}y^{14} + 516z^{11}y^{13} + 1794z^{11}y^{12} \\
& + 3796z^{11}y^{11} + 5186z^{11}y^{10} + 4562z^{11}y^9 \\
& + 482z^{11}y^7 + 30z^{11}y^6 + 20z^{10}y^{13} + 140z^{10}y^{12} \\
& + 712z^{10}y^{11} + 2281z^{10}y^{10} + 4302z^{10}y^9 \\
& + 3061z^{10}y^7 + 781z^{10}y^6 + 30z^{10}y^5 + 2z^9y^{12} \\
& + 24z^9y^{11} + 182z^9y^{10} + 792z^9y^9 + 2248z^9y^8 \\
& + 3194z^9y^7 + 2318z^9y^6 + 402z^9y^5 + 6z^9y^4 \\
& + 4z^8y^{10} + 29z^8y^9 + 162z^8y^8 + 668z^8y^7 \\
& + 1370z^8y^6 + 763z^8y^5 + 56z^8y^4 + 6z^7y^8 \\
& + 18z^7y^7 + 106z^7y^6 + 318z^7y^5 + 80z^7y^4 \\
& + z^6y^7 + 9z^6y^5 + 26z^6y^4 + 266z^{13}y^{11} \\
& + 16z^{15}y^{16} + 28z^{14}y^{13} + 1480z^{13}y^{15} \\
& + 3635z^{12}y^{11} + 2058z^{11}y^8 + 4970z^{10}y^8.
\end{aligned} \tag{95}$$

Solving the coupled equations in Eq. (79) and Eq. (80) using the modified block weights derived in Eq. (92), Eq. (95) and in Supplementary Information [45], we obtain the fixed points  $(p^*, \varepsilon^*)$  for block sizes  $b = 3, 4$  as listed below, with the RG flows around the fixed points shown in Fig. 15. For a block size  $b = 3$ , there are three fixed points: (a)  $p = 0, \varepsilon = 0$ : This is a stable fixed point, remaining stable in both directions. (b)  $p \approx 0.179202, \varepsilon = 0$ : This is a critical fixed point, stable in one direction and unstable in the other. The scaling dimension associated with the relevant direction is  $y_t \approx 1.2662$ , which implies a thermal exponent of  $\nu \approx 0.78975$ . (c)  $p \approx 0.062002, \varepsilon \approx -1.199417$ : This is an unstable fixed point, unstable in both directions. The scaling dimension associated with the maximum relevant direction is  $y_t \approx 1.407$ , leading to a thermal exponent of  $\nu \approx 0.710751$ . Additionally, the crossover exponent corresponding to the second relevant direction is  $\phi \approx 0.295249$ . For a block size of  $b = 4$ , the unstable fixed point is at  $p \approx 0.0620712$  with  $\varepsilon \approx -1.177$ . The thermal exponent associated with the maximum relevant direction is  $\nu \approx 0.69934$ , and the crossover exponent due to

the second relevant direction is  $\phi \approx 0.290565$ . The critical fixed point for this block size is at  $p \approx 0.17328$  with  $\varepsilon = 0$ , and the thermal exponent associated with the relevant direction is  $\nu \approx 0.760657$ . We do not want to comment on the universality class, since to do so would require performing similar calculations for very large block sizes. The full expressions for the modified boundary weights and modified block weights for block size  $b = 4$  are given in the Supplementary Information [45]. The critical fixed point is the classical percolation phase transition point  $(p_c, \varepsilon = 0)$ , indicating a phase transition from non-percolating to percolating phase as shown in Fig. 16 with tunable bond energy cost. At the unstable fixed point with two relevant directions, the crossover exponent  $\phi \approx 0.3$  indicates that the second relevant direction affects the system much more slowly than the primary one. As a result, the system behaves according to the first scaling regime over a wide range before eventually transitioning to a different behavior. The second relevant direction becomes useful at large length scales and guides the flow towards  $\varepsilon = 0$  and  $\varepsilon = -\infty$ , respectively.

## VIII. CONCLUSIONS AND DISCUSSION

In this work, we generalized site percolation on the two-dimensional square lattice by assigning an energy cost  $\varepsilon$  to each occupied nearest-neighbor bond. This model interpolates between a dense-cluster phase ( $\varepsilon \rightarrow -\infty$ ) and a dilute phase ( $\varepsilon \rightarrow \infty$ ) dominated by minimally connected clusters. Numerically, energy weighting preserved the universal scaling form of the cluster-size distribution  $n_s(p, \varepsilon)$  but lowered its cutoff  $s^*(p, \varepsilon)$  as  $\varepsilon$  increased, suppressing large, highly connected clusters. The correlation length  $\xi(\varepsilon)$  similarly decreased with  $\varepsilon$  and became finite at the classical percolation threshold, showing that energetic constraints could destroy geometric criticality even at fixed  $p = p_c(\varepsilon = 0)$ . Monte Carlo simulations with varying  $p$  and bond energy  $\varepsilon$  revealed antiferromagnetic site ordering at large  $\varepsilon$  as  $p$  increases. Using Kadanoff block scaling, we constructed a real-space RG that tracks the renormalization of effective occupation and bond-energy parameters under coarse-graining. For both spanning rules  $(R_0, R_2)$ , the RG recursions showed that increasing  $\varepsilon$  continuously drove the correlation-length exponent from the classical  $\nu = 4/3$  to the isolated-cluster limit  $\nu = 1$ , in quantitative agreement with Coulomb-gas predictions that identified the bond energy cost  $\varepsilon$  as tuning the effective loop fugacity. A complementary lattice-gas RG with site and bond fugacities gave exact recursion relations for block-renormalized parameters and reproduced the full flow, including the shift of  $p_c(\varepsilon)$  toward unity as  $\varepsilon$  grows. This framework can therefore provide a unified description of the crossover behavior between percolation, self-avoiding walks, and broader random-cluster universality classes.

Energy-Weighted Site Percolation could capture the behavior relevant to polymer networks [22, 65], random

media with activation energies [66], force-balanced soft materials [66–69], and networks with resource-dependent connectivity costs [70–72]. This framework can also be extended to analyze lattice walks, where the asymptotic behavior of the generating function in Eq. (41) becomes directly connected to the emergence of the critical percolation threshold and the growth of connected trajectories across the lattice. Generalizing the framework to non-equilibrium directed percolation (DP) [73] further allows the incorporation of non-Hermitian transition rates and irreversible stochastic dynamics that violate detailed balance. Applying large anisotropic bond energy penalties ( $\varepsilon_x \gg \varepsilon_y$  or  $\varepsilon_y \gg \varepsilon_x$ ) biases cluster formation preferentially along selected directions, producing anisotropic connected structures that map onto directed rigidity percolation [74]. This is particularly relevant in gels, fiber networks, and polymers [22, 65, 66], where microscopic deformation energies associated with bending, stretching, and cross-linking [1, 4, 20–22] constrain connectivity and alter gelation thresholds [67, 68, 75, 76]. It can also connect to modern elasticity-based descriptions of polymer networks [77–80] by integrating additional structural features—such as loop formation and spatial correlations. In the context of transport in disordered media,  $\varepsilon$  acts as a thermal activation factor representing microscopic

energy barriers for hopping or tunneling [5–7], exponentially suppressing energetically costly connections and producing heterogeneous transport pathways. In neural networks, the same framework captures how resource constraints shift robustness thresholds and modify cluster statistics [70–72]. In disordered solids and jamming transitions [4, 81], penalizing internal connectivity can suppress highly coordinated bond configurations, providing a connection between rigidity percolation and mechanically marginal elastic states. Finally, extensions to higher dimensions and higher-order neighbor interactions could provide a statistical framework for slow relaxation dynamics in supercooled liquids [82] and lattice glass models [82–84] near the solid-to-liquid transition.

## IX. ACKNOWLEDGMENTS

We thank Mustansir Barma and Sreenath K. Manikandan for useful discussions. This project was funded by intramural funds at TIFR Hyderabad from the Department of Atomic Energy (DAE), Government of India, under Project Identification No. RTI4007.

- 
- [1] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, 1994).
  - [2] B. Bollobás and O. Riordan, *Percolation* (Cambridge University Press, 2006).
  - [3] H. Kesten, *Percolation Theory for Mathematicians* (Birkhäuser, 1982).
  - [4] M. Sahimi, *Applications of Percolation Theory* (Taylor & Francis, 1994).
  - [5] K.-H. Müller, J. Herrmann, B. Raguse, G. Baxter, and T. Reda, *Phys. Rev. B* **66**, 075417 (2002).
  - [6] S. Kirkpatrick, *Reviews of Modern Physics* **45**, 574 (1973).
  - [7] I. S. Beloborodov, A. V. Lopatin, and V. M. Vinokur, *Phys. Rev. B* **72**, 125121 (2005).
  - [8] D. J. Jacobs and M. F. Thorpe, *Physical Review Letters* **75**, 4051 (1995).
  - [9] J. Zhang, E. M. Duxbury, and M. F. Thorpe, *Physical Review E* **91**, 032124 (2015).
  - [10] P. M. Duxbury, D. J. Jacobs, M. F. Thorpe, and C. F. Moukarzel, *Physical Review E* **59**, 2084 (1999).
  - [11] V. K. de Souza and P. Harrowell, *Proceedings of the National Academy of Sciences* **106**, 15136 (2009).
  - [12] J. Chalupa, P. L. Leath, and G. R. Reich, *Journal of Physics C: Solid State Physics* **12**, L31 (1979).
  - [13] A. E. Holroyd, *Probability Theory and Related Fields* **125**, 195 (2003).
  - [14] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris, *Transactions of the American Mathematical Society* **364**, 2667 (2012).
  - [15] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Physical Review E* **82**, 011103 (2010).
  - [16] C. M. Soukoulis and G. S. Grest, *Physical Review B* **44**, 4685 (1991).
  - [17] C. M. Soukoulis, Q. Li, and G. S. Grest, *Phys. Rev. B* **45**, 7724 (1992).
  - [18] L. Ujfalusi and I. Varga, *Phys. Rev. B* **90**, 174203 (2014).
  - [19] R. Bholra, S. Biswas, M. M. Islam, and K. Damle, *Phys. Rev. X* **12**, 021058 (2022).
  - [20] P. R. Onck, T. Koeman, T. van Dillen, and E. van der Giessen, *Phys. Rev. Lett.* **95**, 178102 (2005).
  - [21] C. Heussinger, B. Schaefer, and E. Frey, *Phys. Rev. E* **76**, 031906 (2007).
  - [22] D. A. Head, A. J. Levine, and F. C. MacKintosh, *Phys. Rev. Lett.* **91**, 108102 (2003).
  - [23] C. M. Fortuin and P. W. Kasteleyn, *Physica* **57**, 536 (1972).
  - [24] R. H. Swendsen and J.-S. Wang, *Physical Review Letters* **58**, 86 (1987).
  - [25] C. Ding, Y. Deng, W. Guo, and H. W. J. Blöte, *Phys. Rev. E* **79**, 061118 (2009).
  - [26] C. Fortuin and P. Kasteleyn, *Physica* **57**, 536 (1972).
  - [27] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987).
  - [28] F. Y. Wu, *Rev. Mod. Phys.* **54**, 235 (1982).
  - [29] B. Nienhuis, *Journal of Statistical Physics* **34**, 731 (1984).
  - [30] V. S. Dotsenko, J. L. Jacobsen, and M. Picco, *Nuclear Physics B* **618**, 523 (2001).
  - [31] D. Kapec and R. Mahajan, *Journal of High Energy Physics* **2021**, 136 (2021).
  - [32] K. G. Wilson, *Phys. Rev. B* **4**, 3174 (1971).
  - [33] S. R. Anderson and F. Family, *Phys. Rev. A* **38**, 4198 (1988).

- [34] H. Sun, F. Radicchi, J. Kurths, and G. Bianconi, *Nature Communications* **14**, 1308 (2023).
- [35] C. N. Yang, *Phys. Rev.* **85**, 808 (1952).
- [36] T. D. Lee and C. N. Yang, *Phys. Rev.* **87**, 410 (1952).
- [37] T. Nath and R. Rajesh, *Phys. Rev. E* **90**, 012120 (2014).
- [38] A. P. Furlan, T. J. Oliveira, J. F. Stilck, and R. Dickman, *Phys. Rev. E* **100**, 022109 (2019).
- [39] H. E. Stanley, *Phys. Rev.* **158**, 546 (1967).
- [40] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *The Journal of Chemical Physics* **21**, 1087 (1953).
- [41] R. J. Glauber, *Journal of Mathematical Physics* **4**, 294 (1963).
- [42] K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction*, 6th ed., Graduate Texts in Physics (Springer, 2019).
- [43] K. Kawasaki, *Physical Review* **145**, 224 (1966).
- [44] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, San Diego, 2002).
- [45] See Supplemental Material for details.
- [46] L. P. Kadanoff, *Physics Physique Fizika* **2**, 263 (1966).
- [47] K. G. Wilson, *Phys. Rev. B* **4**, 3184 (1971).
- [48] P. J. Reynolds, H. E. Stanley, and W. Klein, *Phys. Rev. B* **21**, 1223 (1980).
- [49] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, Frontiers in Physics, Vol. 85 (Westview Press, 1992).
- [50] D. C. Hong, *Journal of Physics A: Mathematical and General* **17**, L983 (1984).
- [51] M. Barma and P. Ray, *Phys. Rev. B* **34**, 3403 (1986).
- [52] M. Picco, R. Santachiara, J. Viti, and G. Delfino, *Nuclear Physics B* **875**, 719 (2013).
- [53] P. D. Francesco, P. Mathieu, and D. Sénéchal, *Conformal Field Theory*, Graduate Texts in Contemporary Physics (Springer, New York, 1997).
- [54] V. Dotsenko and V. Fateev, *Nuclear Physics B* **240**, 312 (1984).
- [55] J. L. Jacobsen, *Algèbres, intégrabilité et modèles exactement solubles (aimes)* (2024), complete lecture notes.
- [56] V. G. Kac, in *Group Theoretical Methods in Physics*, Lecture Notes in Physics, Vol. 94 (Springer, 1979) pp. 441–445.
- [57] D. Friedan, Z. Qiu, and S. Shenker, *Phys. Rev. Lett.* **52**, 1575 (1984).
- [58] A. A. Belavin, A. M. Polyakov, and A. B. Zamolodchikov, *Journal of Statistical Physics* **34**, 763 (1984).
- [59] Y. Xu, T. Chen, Z. Zhou, J. Salas, and Y. Deng, *Phys. Rev. E* **111**, 034108 (2025).
- [60] S. Mertens, *Journal of Physics A: Mathematical and Theoretical* **55**, 334002 (2022).
- [61] B. Bollobás and O. Riordan, *Random Structures & Algorithms* **29**, 524 (2006).
- [62] P. J. Reynolds, H. E. Stanley, and W. Klein, *Journal of Physics C: Solid State Physics* **10**, L167 (1977).
- [63] A. Pelissetto and E. Vicari, *Physics Reports* **368**, 549–727 (2002).
- [64] D. Stauffer, *Physics Reports* **54**, 1 (1979).
- [65] C. P. Broedersz and F. C. MacKintosh, *Rev. Mod. Phys.* **86**, 995 (2014).
- [66] C. Storm, J. J. Pastore, F. C. MacKintosh, T. C. Lubensky, and P. A. Janmey, *Nature* **435**, 191 (2005).
- [67] P. J. Flory, *Journal of the American Chemical Society* **63**, 3091 (1941).
- [68] W. H. Stockmayer, *The Journal of Chemical Physics* **12**, 125 (1944).
- [69] K. Ramola and B. Chakraborty, *Journal of Statistical Mechanics: Theory and Experiment* **2016**, 114002 (2016).
- [70] G. Li, L. A. Braunstein, S. V. Buldyrev, S. Havlin, and H. E. Stanley, *Phys. Rev. E* **75**, 045103 (2007).
- [71] M. Schröder, N. A. M. Araújo, D. Sornette, and J. Nagler, *Phys. Rev. E* **96**, 062302 (2017).
- [72] M. Kim, C. T. Diggans, and F. Radicchi, *Nature Communications* **16**, 10.1038/s41467-025-63489-w (2025).
- [73] R. Rajesh and D. Dhar, *Phys. Rev. Lett.* **81**, 1646 (1998).
- [74] M. A. de Menezes and C. F. Moukarzel, *Phys. Rev. E* **60**, 5699 (1999).
- [75] C. Allain, L. Limat, and L. Salomé, *Phys. Rev. A* **43**, 5412 (1991).
- [76] P. J. Flory, *Journal of the American Chemical Society* **63**, 3096 (1941).
- [77] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, International Series of Monographs on Physics, Vol. 73 (Clarendon Press, Oxford University Press, Oxford and New York, 1986) publisher page: <https://global.oup.com/academic/product/the-theory-of-polymer-dynamics-9780198520337>.
- [78] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979) cornell Press page: <https://www.cornellpress.cornell.edu/book/9780801412035/scaling-concepts-in-polymer-physics/>.
- [79] L. R. G. Treloar, *The Physics of Rubber Elasticity*, 3rd ed. (Clarendon Press, Oxford, 1975) online version at Oxford Acad. <https://doi.org/10.1093/oso/9780198570271.001.0001>.
- [80] S. F. Edwards, *Polymer Journal* **17**, 271 (1985).
- [81] G. Lois, J. Blawdziewicz, and C. S. O'Hern, *Phys. Rev. Lett.* **100**, 028001 (2008).
- [82] Y. Nishikawa and L. Berthier, *Phys. Rev. Lett.* **132**, 067101 (2024).
- [83] C. Merrigan, K. Ramola, R. Chatterjee, N. Segall, Y. Shokef, and B. Chakraborty, *Physical Review Research* **2**, 013260 (2020).
- [84] G. Lois, J. Blawdziewicz, and C. S. O'Hern, *Phys. Rev. Lett.* **102**, 015702 (2009).

#### Appendix A: COMBINATORIAL DERIVATION OF THE COEFFICIENTS $M_k$ FOR RULE $R_0$ , $b = 3$

We consider a  $3 \times 3$  block of sites, which contains a total of  $N = 9$  sites. Each site has an occupation probability of  $p$ , while the probability of being unoccupied is  $q = 1 - p$ . A configuration is said to span horizontally if there exists a connected path of occupied sites, using nearest-neighbor adjacency, that connects the left edge (at  $x = 1$ ) to the right edge (at  $x = 3$ ). Similarly, a configuration spans vertically if a path connects the top edge (at  $y = 1$ ) to the bottom edge (at  $y = 3$ ). The normalized occupation probability according to Rule  $R_0$  is then

$$p' = \sum_{k=0}^9 M_k p^k q^{9-k}, \quad (\text{A1})$$

where  $M_k$  counts the number of distinct  $k$ -site configurations that span either horizontally or vertically. We index the  $3 \times 3$  sites in row-major order:

$$\begin{array}{ccc} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{array}$$

A subset  $A \subset \{0, \dots, 8\}$  is considered a horizontal minimal spanning set if it contains a connected path that connects the left edge to the right edge, and no proper subset of  $A$  possesses this property. The minimal spanning sets that span the block from left to right for  $b = 3$  can be listed as follows:

$$\begin{aligned} A_1 &= \{0, 1, 2\}, & A_2 &= \{3, 4, 5\}, & A_3 &= \{6, 7, 8\}, \\ A_4 &= \{0, 1, 4, 5\}, & A_5 &= \{1, 2, 3, 4\}, & A_6 &= \{3, 4, 7, 8\}, \\ A_7 &= \{4, 5, 6, 7\}, & A_8 &= \{0, 1, 4, 7, 8\}, & A_9 &= \{1, 2, 4, 6, 7\}. \end{aligned} \quad (\text{A2})$$

Thus, there are 9 minimal sets in total: three of size 3, four of size 4, and two of size 5.

### Inclusion–Exclusion Expansion

According to the principle of inclusion-exclusion, for a finite collection of finite sets  $X_1, \dots, X_n$ ,

$$\left| \bigcup_{i=1}^n X_i \right| = \sum_{k=1}^n (-1)^{k+1} \left( \sum_{1 \leq i_1 < \dots < i_k \leq n} |X_{i_1} \cap \dots \cap X_{i_k}| \right), \quad (\text{A3})$$

which can be written compactly as:

$$\left| \bigcup_{i=1}^n X_i \right| = \sum_{\emptyset \neq I \subseteq \{1, \dots, n\}} (-1)^{|I|+1} \left| \bigcap_{i \in I} X_i \right|. \quad (\text{A4})$$

Let  $\mathcal{A} = \{A_1, \dots, A_9\}$  be the set of all minimal sets. For any subset  $I \subseteq \{1, \dots, 9\}$ , let  $U_I = \bigcup_{i \in I} A_i$  represent its union. By inclusion-exclusion, the number of configurations with exactly  $k$  occupied sites that span from left to right is:

$$H_k = \sum_{\emptyset \neq I \subseteq \mathcal{A}} (-1)^{|I|+1} \binom{9 - |U_I|}{k - |U_I|}, \quad (\text{A5})$$

with the convention that  $\binom{n}{r} = 0$  if  $r < 0$  or  $r > n$ .

Let  $\mathcal{A}$  represent the set of all minimal spanning subsets of sites in a  $b \times b$  block that extend from the left side to the right side of the block. From the principle of inclusion-exclusion, the number of configurations in which exactly  $k$  sites are occupied, and span from left to right, is given by:

$$H_k = \sum_{\emptyset \neq I \subseteq \mathcal{A}} (-1)^{|I|+1} \binom{b^2 - |U_I|}{k - |U_I|}, \quad U_I = \bigcup_{A \in I} A, \quad (\text{A6})$$

where we adopt the convention that  $\binom{n}{r} = 0$  if  $r < 0$  or  $r > n$ .

*Proof.* A configuration with exactly  $k$  occupied sites corresponds to a subset  $S \subseteq [b^2]$  where  $|S| = k$ . This configuration is said to span left to right if it includes at least one of the minimal spanning subsets  $A \in \mathcal{A}$ . Therefore, the collection of configurations of size  $k$  that span left to right can be expressed as:

$$\mathcal{H}_k = \bigcup_{A \in \mathcal{A}} \mathcal{X}_A, \quad \text{where } \mathcal{X}_A = \{S \subseteq [b^2] : |S| = k, S \supseteq A\}. \quad (\text{A7})$$

Therefore,

$$H_k = |\mathcal{H}_k| = \left| \bigcup_{A \in \mathcal{A}} \mathcal{X}_A \right|. \quad (\text{A8})$$

Applying the inclusion–exclusion principle gives:

$$H_k = \sum_{\emptyset \neq I \subseteq \mathcal{A}} (-1)^{|I|+1} \left| \bigcap_{A \in I} \mathcal{X}_A \right|. \quad (\text{A9})$$

For any nonempty subset  $I \subseteq \mathcal{A}$ , the intersection  $\bigcap_{A \in I} \mathcal{X}_A$  consists of all  $k$ -site configurations that contain every  $A \in I$ , i.e.

$$\bigcap_{A \in I} \mathcal{X}_A = \{S \subseteq [b^2] : |S| = k, S \supseteq U_I\}, \quad U_I = \bigcup_{A \in I} A. \quad (\text{A10})$$

Given  $U_I$ , the number of  $k$ -site super sets  $S \supseteq U_I$  is the number of ways to choose the remaining  $k - |U_I|$  sites from the  $b^2 - |U_I|$  sites that are not in  $U_I$ :

$$\left| \bigcap_{A \in I} \mathcal{X}_A \right| = \binom{b^2 - |U_I|}{k - |U_I|}, \quad (\text{A11})$$

where  $\binom{n}{r} = 0$  if  $r < 0$  or  $r > n$ . Substituting this expression into the inclusion–exclusion formula yields

$$H_k = \sum_{\emptyset \neq I \subseteq \mathcal{A}} (-1)^{|I|+1} \binom{b^2 - |U_I|}{k - |U_I|}, \quad (\text{A12})$$

which proves the stated result.  $\square$

We group all such non-empty subsets  $I$  by their union size  $u = |U_I|$  and define

$$S_u = \sum_{\substack{\emptyset \neq I \subseteq \mathcal{A} \\ |U_I| = u}} (-1)^{|I|+1}. \quad (\text{A13})$$

Then the inclusion–exclusion formula for  $b = 3$  becomes the compact expression

$$H_k = \sum_{u=1}^9 S_u \binom{9 - u}{k - u}. \quad (\text{A14})$$

Enumerating all such subsets of  $\mathcal{A}$  and summing signs grouped by  $u = |U_I|$  gives the following signed multiplicities:

$$\begin{aligned} S_1 &= 0, & S_2 &= 0, & S_3 &= 3, & S_4 &= 4, \\ S_5 &= -6, & S_6 &= -9, & S_7 &= 14, & S_8 &= -6, & S_9 &= 1. \end{aligned} \quad (\text{A15})$$

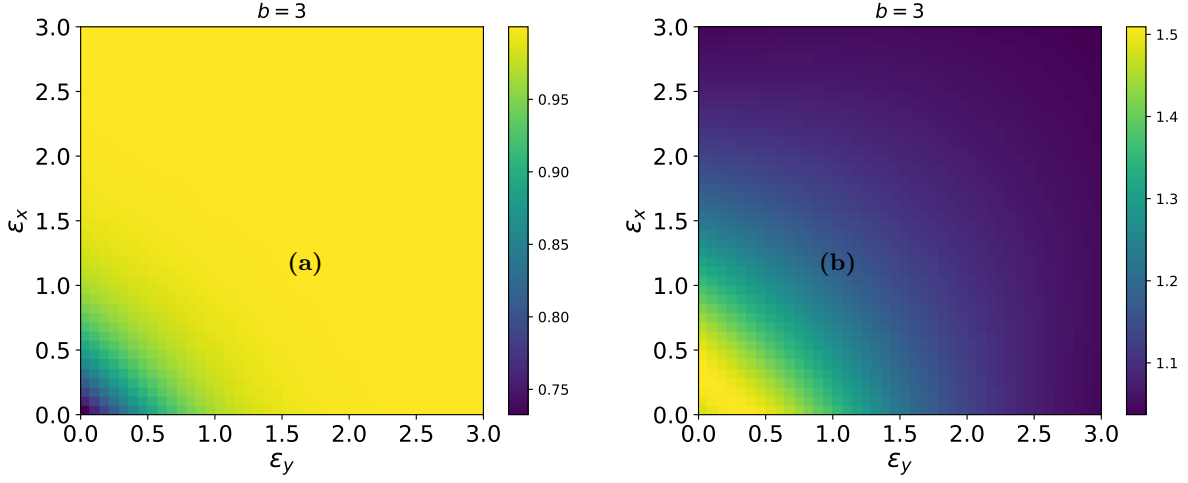


FIG. 17. (a) The variation of the critical site percolation threshold was computed from RG analysis in real space using Kadanoff blocks. This analysis utilized the  $R_2$  rule, incorporating an anisotropic energy cost per link in the  $x$  direction, denoted as  $\epsilon_x$ , and an energy cost per link in the  $y$  direction, denoted as  $\epsilon_y$ , for a block size of  $b = 3$ , (b) variation of the correlation length exponent  $\nu$ , also computed from RG analysis in real space using Kadanoff blocks, following the same  $R_2$  rule and considering the anisotropic energy costs  $\epsilon_x$  and  $\epsilon_y$  for the block size  $b = 3$ .

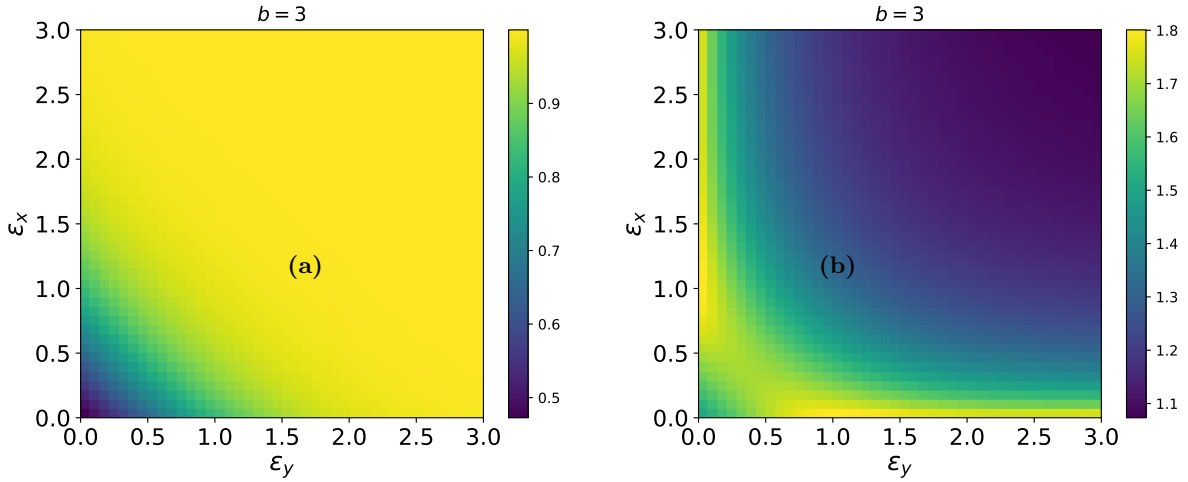


FIG. 18. (a) The variation of the critical site percolation threshold was computed from RG analysis in real space using Kadanoff blocks. This analysis utilized the  $R_0$  rule, incorporating an anisotropic energy cost per link in the  $x$  direction, denoted as  $\epsilon_x$ , and an energy cost per link in the  $y$  direction, denoted as  $\epsilon_y$ , for a block size of  $b = 3$ , (b) variation of the correlation length exponent  $\nu$ , also computed from RG analysis in real space using Kadanoff blocks, following the same  $R_0$  rule and considering the anisotropic energy costs  $\epsilon_x$  and  $\epsilon_y$  for the block size  $b = 3$ .

Evaluating the values of  $H_k$  for all possible values of  $k$  gives, for  $k = 3$ :

$$H_3 = S_3 \binom{6}{0} = 3 \cdot 1 = 3. \quad (\text{A16})$$

Since vertical symmetry gives  $V_3 = 3$  and no 3-site configuration spans both directions,

$$M_3 = H_3 + V_3 = 6. \quad (\text{A17})$$

For  $k = 4$ :

$$H_4 = S_3 \binom{6}{1} + S_4 \binom{5}{0} = 3 \cdot 6 + 4 \cdot 1 = 22, \quad (\text{A18})$$

$$M_4 = 2H_4 = 44. \quad (\text{A19})$$

Similarly for  $k = 5$ :

$$\begin{aligned} H_5 &= S_3 \binom{6}{2} + S_4 \binom{5}{1} + S_5 \binom{4}{0} \\ &= 3 \cdot 15 + 4 \cdot 5 - 6 \cdot 1 = 59, \end{aligned} \quad (\text{A20})$$

$$(H \cap V)_5 = 25,$$

$$M_5 = 2H_5 - (H \cap V)_5 = 2 \cdot 59 - 25 = 93. \quad (\text{A21})$$

Among the  $\binom{9}{5} = 126$  configurations with 5 occupied sites, exactly 25 span both horizontally and vertically. Enumerating them in row-major site indices (0–8):

$$\begin{aligned} [H \cap V]_5 = \{ & (0, 1, 2, 3, 6), (0, 1, 2, 4, 7), (0, 1, 2, 5, 8), \\ & (0, 1, 4, 5, 7), (0, 1, 4, 5, 8), (0, 1, 4, 7, 8), \\ & (0, 3, 4, 5, 6), (0, 3, 4, 5, 7), (0, 3, 4, 5, 8), \\ & (0, 3, 4, 7, 8), (0, 4, 5, 6, 7), (0, 4, 5, 7, 8), \\ & (1, 2, 3, 4, 6), (1, 2, 3, 4, 7), (1, 2, 3, 4, 8), \\ & (1, 2, 4, 5, 7), (1, 2, 4, 6, 7), (1, 4, 5, 6, 7), \\ & (1, 4, 5, 7, 8), (2, 3, 4, 5, 6), (2, 3, 4, 6, 7), \\ & (2, 3, 4, 6, 8), (2, 4, 5, 6, 7), (3, 4, 5, 6, 7), \\ & (3, 4, 5, 7, 8) \}. \end{aligned} \quad (\text{A22})$$

From the same procedure, one obtains

$$H_k = [0, 0, 0, 3, 22, 59, 67, 36, 9, 1], \quad (\text{A23})$$

$$(H \cap V)_k = [0, 0, 0, 0, 0, 25, 52, 36, 9, 1], \quad (\text{A24})$$

$$M_k = [0, 0, 0, 6, 44, 93, 82, 36, 9, 1]. \quad (\text{A25})$$

The polynomial for rule  $R_0$ ,  $b = 3$  is therefore:

$$\begin{aligned} p' &= \sum_{k=0}^9 M_k p^k q^{9-k} \\ &= p^9 + 9p^8q + 36p^7q^2 + 82p^6q^3 + 93p^5q^4 \\ &\quad + 44p^4q^5 + 6p^3q^6. \end{aligned} \quad (\text{A26})$$

where  $q = 1 - p$ . The polynomial expansion for other  $b$  cases and other rules like  $R_1$  and  $R_2$  as described in [48] can be similarly derived using the Inclusion-Exclusion principle. Nevertheless, they can be easily computed by writing a computer program, up to  $b = 5$ . For  $b \geq 6$ , Monte Carlo techniques need to be implemented.