How complexity emerges in urban systems: Theory of urban morphology

Segun Goh,[1] M. Y. Choi,[1,*] Keumsook Lee,[2,†] and Kyung-min Kim[3]
[1]*Department of Physics and Astronomy and Center for Theoretical Physics, Seoul National University, Seoul 151-747, Korea*
[2]*Department of Geography, Sungshin Women's University, Seoul 136-742, Korea*
[3]*Graduate School of Environmental Studies, Seoul National University, Seoul 151-742, Korea*

(Received 29 January 2016; revised manuscript received 5 May 2016; published 23 May 2016)

Human beings develop the land and transform land use patterns, constructing artificial structures. Among them, the city is a representative system and its morphology has attracted much attention. While most existing studies have been devoted to individual dynamics and focused on the proximity of specific areas of a city, we here pay attention to the city as a complex system, where interactions between individuals give rise to emergent properties. Specifically, analyzing the big data on every building in Seoul City, we specify the relevant interactions among constituents and probe the emergence of complex land use patterns. In particular, based on the empirical observations, we illustrate that interactions between land uses are frustrated, which serves as a basic postulate of the theory of urban morphology. We examine this conjecture with the help of a layered Ising-type model and disclose that the actual land use pattern emerges at the criticality of the system in the presence of heterogeneously distributed fields. It is also remarked that our model, allowing quantitative predictions, can easily be applied to other cities around the world.

DOI: 10.1103/PhysRevE.93.052309

I. INTRODUCTION

Cities are highly organized spatial structures that have been developed by human beings. As a result of the organization, each city forms a distinctive pattern depending on its own topographical, historical, economical, and political environments. Nevertheless, substructures composing urban areas should be organized in a proper way for the cities to function properly. Specifically, each city should provide the bases and infrastructures for urban dwellers including shelters, job opportunities, transportation, and so on. To elucidate the pattern of the urbanized area, called urban morphology, various studies have been conducted [1–7].

Meanwhile, the city is one of the most representative complex phenomena among social systems. Recently, complex system perspectives begin to find general acceptance in studies of social systems [8–10]. Many social phenomena are now regarded as emergent and collective properties, exhibiting such characteristics as criticality and scaling [10,11], and have been widely probed by means of statistical mechanics. For instance, studies of urban growth [12,13] and urban structures including street networks [7,14] and public transportation systems [11,15] have been conducted extensively from this viewpoint. Still current understanding as to the fundamental rules governing the formation and emergence of cities remains in its infancy due to the inherent complex nature of the urban system. Specifically in the case of urban morphology, most existing studies, focusing on the properties of individuals, are descriptive rather than theoretical and there still lacks an integrated framework for urban morphology.

Furthermore, insufficiency of real data has also prevented proper understanding of urban formation and the land use pattern. In principle, most detailed measurements of a social system including a city would be possible if the locations and movements of every individual are specified. Indeed individual dynamics has proved to be traceable mostly by analyzing mobile phone data [16,17]; nevertheless, complete enumeration of a whole urban system remains to be achieved. Moreover, a city involves various scales, ranging from individuals [16] to urban substructures such as the central business district (CBD) [5,13]. In addition, there might exist strong fluctuations when individual dynamics is taken into account in detail. Therefore, the most relevant measure should be determined carefully.

Related to the land use pattern and urban morphology, the distribution of buildings provides an appropriate measure [18,19]. The building distribution typically characterizes unique features of each city and provides an accurate description of the detailed configuration of the urban area. Undoubtedly, buildings are not distributed randomly in space; rather, they tend to aggregate in a city, accompanying economic activities. To measure the tendency toward aggregation of urban economic activities, distance-based methods have been developed [20–22]. Movements in a city take place under the constraint guided by such inhomogeneous building distribution, and the city operates accordingly. Dealing with the building distribution data of the urbanized area rather than individual dynamics, one can also bypass the technical difficulty and make use of the relatively static nature of the building.

In this direction, we examine the building distribution of Seoul City, which includes information on the location, area, and functional characteristics of *every* building. Following the viewpoint of statistical mechanics and complex systems, we regard urban morphology as a collective behavior of constituents coupled to each other. To capture distinguished features of emergent phenomena, we adopt a coarse-grained description of Seoul City, placing its urban areas on a lattice. We not only provide a description of the observed phenomena but also infer the underlying principles of urban land use patterns, constructing a model Hamiltonian with relevant components: The frustrated interlayer coupling is proposed as a fundamental ingredient for the formation of urban morphology and the criticality as a necessary condition

*mychoi@snu.ac.kr
†kslee@sungshin.ac.kr

for the emergence of complex land use patterns. The role of heterogeneous fields is also addressed.

II. FRACTAL NATURE OF URBANIZED REGIONS IN SEOUL

Before proceeding further, we describe briefly the Seoul building data set [23]. As addressed already, it contains quantitative information on *every* building located in the administrative area of Seoul City in the year 2011. The attributes include the number of floors, structure, gross floor area, position or address, zip code, use, and type of the roof of the building. There were 678 594 buildings in Seoul at the time of the data gathering, the total area of which amounted to 429.34 km^2 (cf. the area of Seoul, 605.25 km^2). In particular, uses of the buildings are classified into 29 types, e.g., apartments, offices, neighborhood facilities, and so on. As the use of the building represents the land use directly, it serves as an indicator of the urban morphology which should comprise differentiated patterns as well as shapes of the urbanized areas. Representative attribute types and their ratios in terms of the gross floor area are displayed in Table I. Among various attributes of buildings, longitude and latitude coordinates and gross floor areas turn out to be relevant information for specifying the configurations of coarse-grained areas, as discussed in this and the next sections. (According to the Korean Enforcement Decree of the Building Act, the gross floor area of a building is defined to be the sum of the floor areas of all the floors of the building.)

We now extract the urbanized area of Seoul City from the building data set. It is well known that the urbanized areas constitute fractals [24,25], as to which the diffusion-limited aggregation model was studied extensively [12,24–26]. We also verify the fractal nature of the urbanized area and compute the fractal dimension from the building distribution data, which, unlike the data analyzed in previous studies, provide direct measurement. In particular, the dimension needs to be specified accurately since it plays a crucial role in the collective behavior.

However, there exist severe drawbacks in the bare description where the use or location of a certain building is fixed at a particular attribute or value. For example, shopping facilities around a residential area, categorized as commercial in the bare description, should be considered a residential-commercial complex. Moreover, although the surface areas occupied by buildings are different from one another, each building is put on an equal footing if regarded as a point element. In particular, such a fine-grained description on a length scale smaller than the typical size of a single building is obviously superfluous, for it would generate fictitious empty spaces, which are in reality occupied by buildings, and in consequence introduce heterogeneity, which is an artifact. This may make an undesirable alteration on the behavior of system, unless there exist infinitely many buildings so that the size can be ignored.

In this paper, we follow the process suggested in a recent study [11] and perform coarse graining of the urbanized areas into small boxes to circumvent the redundancy in the fine-grained description. Constructing a coarse-grained lattice by putting a site on the face of each box, we can also utilize methods developed for lattice systems. The lattice constant and the total number of boxes or sites on the lattice are denoted by a and N, respectively. A typical example of the coarse-grained lattice is shown in Fig. 1(a).

To compute the fractal dimension of the lattice, we first examine the coordination number, which stands for the number of nearest neighbors, and observe that its value varies from site to site. Labeling the site dependence, we write z_i to denote the coordination number of site i. In fact, the presence of empty spaces and consequent variations in the number of the nearest-neighboring sites apparently reflect the fractal nature of the lattice. Namely, whereas $z_i = z = 2d$ at every site on a d-dimensional regular lattice, in the case of a fractal lattice, the coordination number has a noninteger mean value z and there should thus exist empty sites. In Fig. 1(b), which exhibits half the (mean) coordination number depending on the lattice constant, we obtain directly the fractal dimension of the coarse-grained lattice of Seoul City: $d = 1.82 \pm 0.01$.

TABLE I. Classification of buildings and the ratio in terms of the gross floor area of each attribute. The attribute of each building is provided in the raw data set. We further classify the use of building into three categories: residential (R), workplace (W), and commercial (C) buildings. There are 58 079 buildings (6.56% in terms of the area) for which the uses are not categorized. Among them, 55 146 buildings (5.8%) have no attribute in the raw data, while attributes such as farming and fishing houses, transportation facilities, hazardous material storage facilities, or vehicle maintenance facilities in the data (2933 buildings, 0.68%) are, in consideration of uncertainty, not categorized.

Attribute	Ratio	Category	Attribute	Ratio	Category
Apartment	0.3136	R	Multifamily house	0.1496	R
Multiplex house	0.0681	R	Detached house	0.0488	R
Row house	0.0181	R	Employee apartment	0.0012	R
Office	0.0828	W	Education/research facility	0.0450	W
Manufacturing facility	0.0240	W	Neighborhood facility	0.1286	C
Large store facility	0.0194	C	Religious facility	0.0114	C
Hotel	0.0070	C	Accommodation	0.0045	C
Cultural/broadcast facility	0.0035	C	Sports facility	0.0031	C
Food sanitation facility	0.0031	C	Marketplace	0.0009	C
Medical facility	0.0009	C	Public sanitation facility	0.0005	C
Amusement facility	0.0005	C	Inn	0.0004	C
Ritual facility	0.0003	C	Recreational facility	0.0001	C

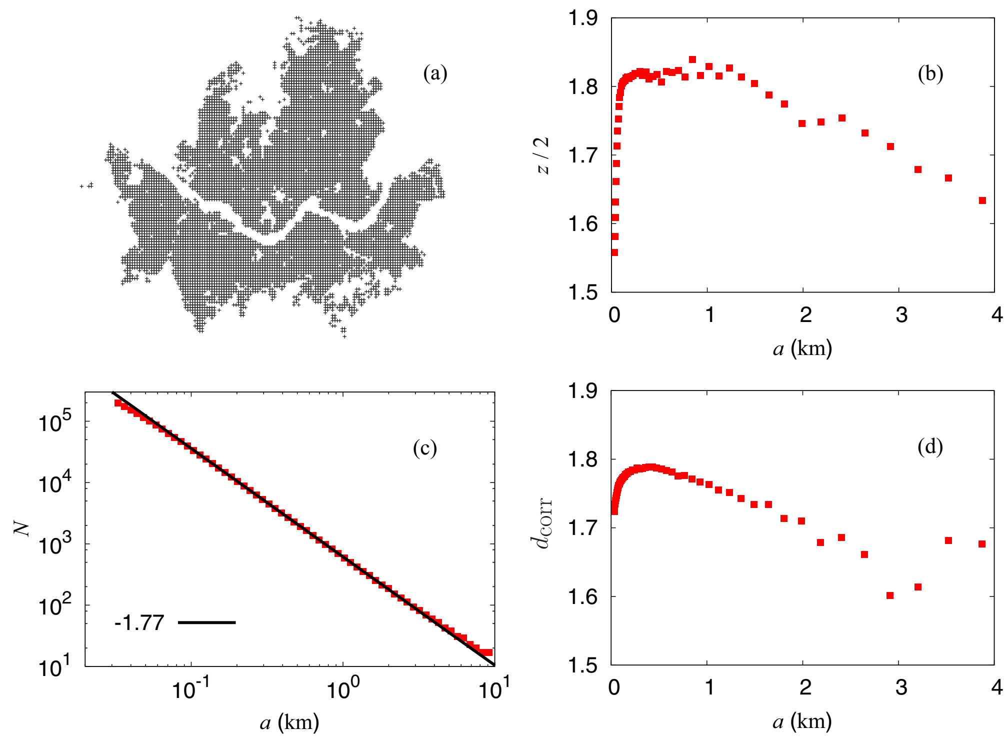

FIG. 1. Lattice structure of the urbanized area in Seoul City and its fractal dimension. (a) A coarse-grained lattice with lattice constant $a = 250$ m. (b) Half the coordination number $z/2$ versus lattice constant a, which gives the dimension $d = 1.82 \pm 0.01$ in the range $100 < a < 1500$ m. Note the strong finite-size effects for $a < 100$ m and strong boundary effects for $a > 1500$ m. (c) Number N of sites versus lattice constant a, disclosing the scaling relation between them. Fitted to the solid line of slope -1.77, it leads to the fractal dimension $d = 1.77 \pm 0.01$. (d) Correlation dimension $d_{\rm corr}$ versus lattice constant a, displaying a pattern similar to that in (b). In the range $100 < a < 1000$ m, we obtain $d = 1.78 \pm 0.01$.

The presence of empty spaces, however, does not guarantee that the lattice is a fractal. For example, when the sites are removed randomly, the average value of z becomes a noninteger; nevertheless, the dimension of the lattice remains the same as that of the original lattice. Furthermore, the average value of z could also be reduced when the sites are removed in an orderly manner; such a reduction, which is exemplified by the Han (river) in the case of Seoul, brings on the boundary effects rather than the fractal nature. Therefore, the fractal character of the lattice should be confirmed in a more concrete way. Examining how the number of sites changes with the lattice constant, we verify the scaling relation between them: $N \sim a^{-d}$, which confirms a fractal with dimension $d = 1.77 \pm 0.01$ [see Fig. 1(c)]. In addition, one can also consider the correlation sum $C(r)$ [27] and probe its scaling relation. Computing $C(r)$, we indeed observe the scaling relation $C(r) \sim r^{d_{\rm corr}}$ and plot in Fig. 1(d) the obtained values of the correlation dimension $d_{\rm corr}$ as the lattice constant is varied. The fractal character of the lattice is manifested again, with dimension $d = d_{\rm corr} = 1.78 \pm 0.01$.

Note here the appropriate length scale for the coarse-grained description. First, we observe a sharp decrease in the fractal dimension for small values of the lattice constant, $a < 100$ m, including the correlation dimension computed directly from the point element description ($d_{\rm corr} \approx 1.67$). This has its origin in the artificial heterogeneity introduced in the fine-grained description. Second, in the opposite case of large values, $a > 1000$ m, boundary effects are evident. Gathering the results from three different approaches, we finally conclude that the urbanized area of Seoul City is indeed a fractal of dimension $d \approx 1.8$, which is consistent with the results in previous studies [25,28]. In the following, we restrict our analysis to the coarse-grained system satisfying these conditions.

III. STATE VECTOR DESCRIPTION OF URBAN AREAS

In this section, we specify the state of each site on the fractal lattice, constructed of the coarse-grained areas in Seoul City, and analyze the distribution of configurations of the whole system. For simplicity, one may define the state of a site according to, e.g., whether the area of the site is active or inactive. As remarked, however, cities are highly patterned systems consisting of functionally organized modules, such as CBDs, industrial complexes, and residential areas. Even the actual distribution of modules differs from one city to another, and coexistence of them is a key ingredient which defines and enables cities, especially metropolises such as Seoul, to work as self-sufficient spatial structures for human life. It is therefore desirable to categorize the state of each site in detail, so as to reflect the land use. In previous studies [1,29,30], urban land use has been categorized as three distinct allocations. Also urban life has been classified into three representative activities, based on the eigenmode analysis of mobile phone usage data [31]. Following these, we classify the building into three distinct categories: residential, workplace, and commercial buildings (denoted by r, w, and c categories hereafter), as summarized in Table I. Note that one may subdivide the categories in a more detailed way, appropriate for the specific purpose of the study [32,33].

Finally, we consider the total gross floor area of each category in an area, which is denoted by $\sigma_{i,x}^{\mathrm{raw}}$ with i and $x\,(=r,w,c)$ being lattice site and category indices, respectively, and regard it as a component of the state vector defined in coarse-grained area i. Namely, the state of each site is described by the state vector, specifying the total areas of the buildings of three categories: $\vec{\sigma}_i^{\mathrm{raw}} \equiv (\sigma_{i,r}^{\mathrm{raw}},\sigma_{i,w}^{\mathrm{raw}},\sigma_{i,c}^{\mathrm{raw}})$. In other words, the total gross floor areas and uses of buildings in each coarse-grained area quantify the amplitude and the direction of the state vector at the site. It is also conceivable to adopt the number of buildings as a measure of the amplitude. However, there are no correlations between the numbers and the gross floor areas of buildings (results not shown) due to the large variations in the size of buildings. Therefore, the analysis based on the number of buildings could give an inaccurate description of the reality.

Until this stage, the state of each site is represented by continuous variables $\sigma_{i,x}^{\mathrm{raw}}$. Henceforth, for further simplicity, we adopt a two-state description, characterizing the total gross floor area of each category as a binary variable. Setting an appropriate threshold area, we assign $\sigma_{i,x} = \pm 1$ as the x component of the state of lattice site i if the sum of the gross floor areas of buildings of category x in the corresponding area is larger or smaller than the threshold. The state of site i is then specified by $\vec{\sigma}_i \equiv (\sigma_{i,r},\sigma_{i,w},\sigma_{i,c})$, where each component $\sigma_{i,x}$ can be regarded as an Ising spin. Accordingly, each site can take eight states: $\vec{\sigma}_i = (-1,-1,-1)$, $(1,-1,-1)$, $(-1,1,-1)$, $(-1,-1,1)$, $(1,1,-1)$, $(1,-1,1)$, $(-1,1,1)$, and $(1,1,1)$, which are, for convenience, abbreviated as 0, r, w, c, rw, cr, wc, and rwc states, respectively.

Then the remaining issue is how to determine the threshold of the total gross floor area. In search of the appropriate value, we examine the rank distribution of the total gross floor areas and plot the results in Fig. 2. Interestingly, residential and commercial building distributions follow power-law distributions in the high-rank regimes. Specifically, the power-law behavior persists approximately down to the top 30th percentile. We therefore consider this value as a lower limit for the threshold. In contrast, the rank distribution of workplace buildings decays exponentially rather than algebraically. Furthermore, the number of boxes in which at least a single workplace building is located is smaller than those of residential and commercial buildings. In the case of lattice constant $a = 138$ m, for example, only 55% of boxes contain workplace buildings and we adopt it as an upper limit. Note that these behaviors are consistently observed irrespective of the lattice constant. One may then take any percentile value for the threshold, and assign $\sigma_{i,x} = \pm 1$ to site i if the total gross floor area of category x is larger or smaller than the specific percentile value. It is pleasing that threshold values taken from criteria between 30% and 55% turn out to give consistent results, and in this paper we show only the results of the median (50%) criterion for the threshold.

The site classified accordingly in the three-category description, shown in Fig. 3, manifests characteristic nonuniform distributions of sites in different states. First of all, sites in the

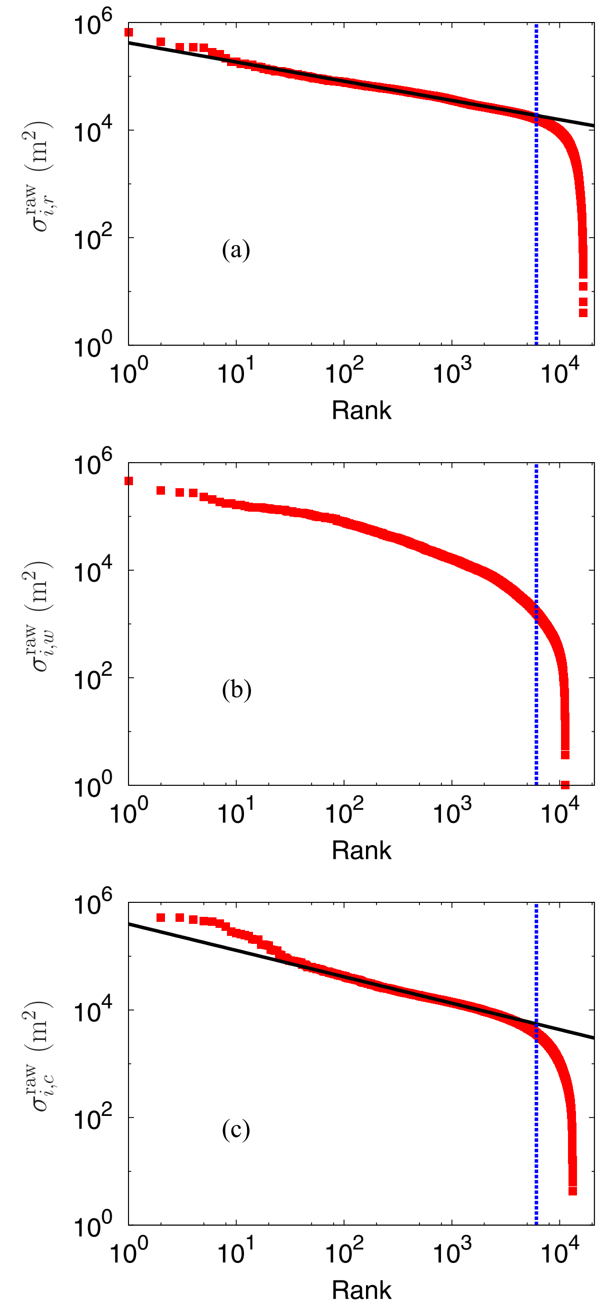

FIG. 2. Rank distributions (red squares) of total gross floor areas for (a) residential, (b) workplace, and (c) commercial buildings. The lattice constant is taken to be $a = 138$ m. Power-law behaviors are observed in (a) and (c), down to the 30th percentile indicated by vertical blue lines (6093 among the total of 20 309 boxes in this coarse-grained scale). The black solid lines describe the power-law distributions with exponents (a) 0.36 and (c) 0.49.

r, w, and c states are primarily located close to the boundary, as shown in Fig. 3(a). On the other hand, rwc-state sites exhibiting mixed land use are relatively centered, forming clusters from place to place. It is also observed in Fig. 3(b) that the CBD areas of Seoul are represented mostly by the wc state. One can further confirm that sites in the r or w state are located mainly near the boundary and easily distinguished from the mixed use sites (rwc state) or the CBD (wc state) at the center.

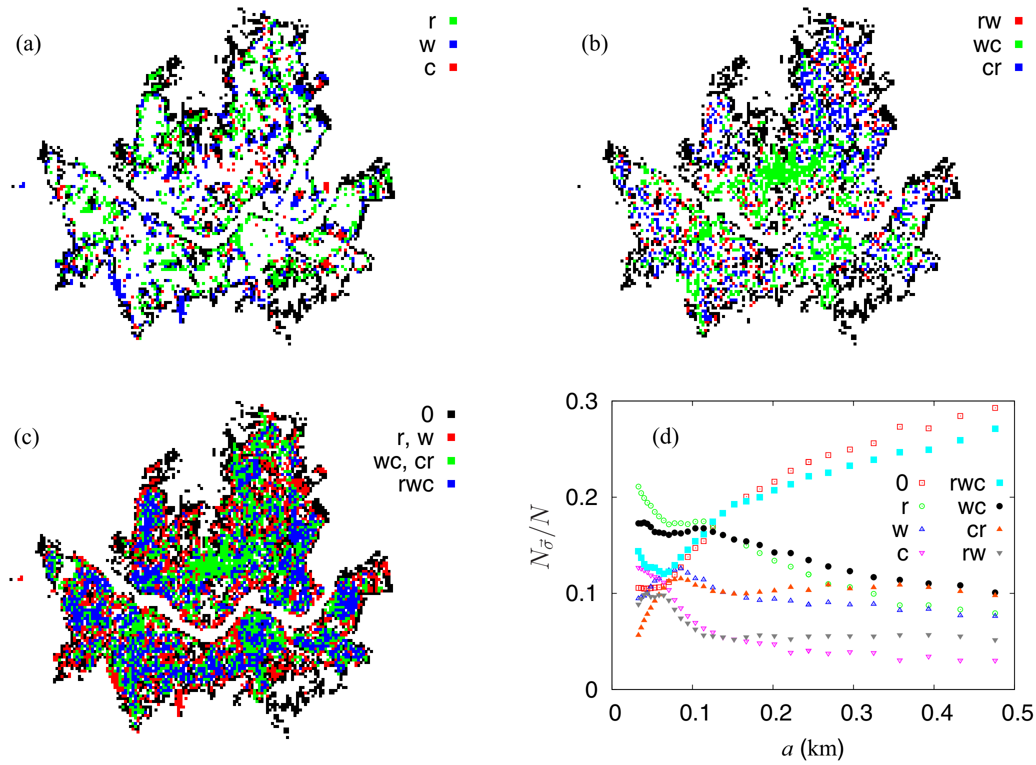

FIG. 3. Three-category Ising-type description of Seoul City under the median criterion. (a) Sites in the r, w, and c states are depicted. It is observed that sites near the boundary of Seoul City are mostly in these states. (b) Sites in the rw, wc, and cr states are displayed. Note that downtown areas are populated mainly by the wc-state sites; in effect buildings of the residential category are expelled from the downtown, giving rise to an urban doughnut pattern. (c) Sites in the rwc state as well as in the 0 state. In contrast to r-, w-, and c-state sites, rwc-state sites are located away from the boundary. (d) Fraction $N_{\vec{\sigma}}/N$ of sites in state $\vec{\sigma}$ are presented versus the lattice constant a. Notably, the fractions of c- and rw-state sites are smaller than those of other state sites.

These discriminative features reflect the spatially modularized pattern of the urbanized area. Finally, we compute the fraction of states $N_{\vec{\sigma}}/N$, where $N_{\vec{\sigma}}$ denotes the number of lattice sites in state $\vec{\sigma}$, and plot in Fig. 3(d) the results as the lattice constant a is varied. With the finite-size effects for $a \lesssim 0.1$ km excepted, there exist fewer c-state sites than r- and w-state ones and fewer rw-state sites than wc- and cr-state ones. These characteristics are expected to originate from the coupling between the buildings of different categories. In particular, the commercial category appears to play a distinctive role in the pattern formation, subordinate to the principal components of human life, habitation and production.

IV. MODELING URBAN MORPHOLOGY

A. Model description

Equipped with the analysis in the previous section, we construct a model describing the major characteristics of the land use pattern. Based on the Ising-spin description, we introduce three layers of the lattice, each of which hosts the state component corresponding to each category. Specifically, we place the Ising spin $\sigma_{i,x}$ at site i on layer x, describing the x-category state of the coarse-grained area i. The Hamiltonian of the urban morphology system is then proposed in the following way: We first assume that there exists a tendency toward aggregation [34–36] between buildings of the same category, which may be taken into account by the "ferromagnetic" nearest-neighbor interactions on a layer, i.e., between the same category components of the state vector. The Hamiltonian describing such intralayer coupling or interactions between the buildings of category x then reads

$$\mathcal{H}_x(\{\sigma_{i,x}\}) = -\sum_i h_{i,x}\sigma_{i,x} - J_x \sum_{\langle i,j\rangle} \sigma_{i,x}\sigma_{j,x}, \qquad (1)$$

where the temperature has been absorbed in defining the Hamiltonian and the nonuniform external field $h_{i,x}$ has been introduced to measure advantages and disadvantages in constructing the building of category x in area i. For example, if the slope of area i is steep, we may assign a negative value of the field $[h_{i,x} < 0]$ to mimic the difficulty to construct buildings in that area. History of the city, infrastructures such as the road and subway networks, or even the price of land can also be good candidates for the field. We may also take into account the urban development planning by assigning positive external fields and probe the consequence of the planning.

Leaving the detailed verification for further studies, we in this paper assign negative fields only at the boundary as a simple but significant example. Such negative fields reflect the minimal constraint, imposed mostly by the mountains and rivers located at the boundary of Seoul City. When specifying the strength of the field at the boundary, we consider the following two factors: (1) if a lattice site is surrounded by many outer points, the negative field should be strong, and (2) if the nearest-neighbor coupling is strong, the effects of

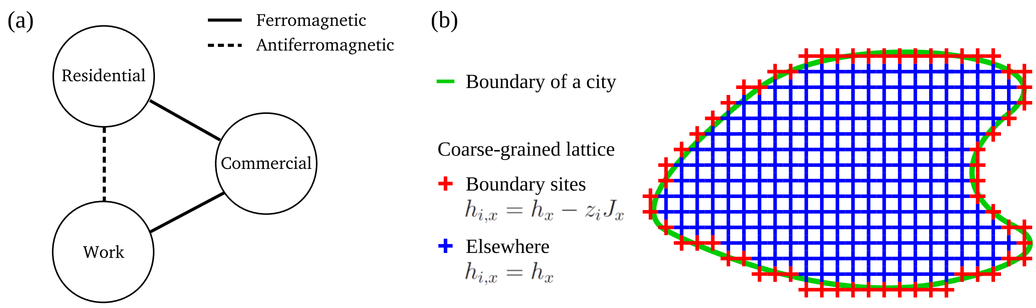

FIG. 4. Model description. (a) A schematic diagram of the interlayer coupling, displaying frustration due to the antiferromagnetic interaction between the residential layer and the workplace layer. (b) External fields applied in the city and boundary conditions imposed at the boundary.

the boundary should also be strong. In addition, we also apply weak fields on the whole lattice to simulate the efforts of urban settlers to maintain the city. Accordingly, the external fields are specified by

$$h_{i,x} = \begin{cases} h_x - z_i J_x & \text{at the boundary } \Omega, \\ h_x & \text{elsewhere.} \end{cases}$$

Now the strength of h_x controls the activity of the city. For example, if the city is in its developmental phase, h_x should have large values. To model a sustaining city like Seoul, we assume that the total field strength vanishes:

$$\sum_i h_{i,x} = -\sum_{i\in\Omega} z_i J_x + N h_x = 0, \tag{2}$$

which yields the uniform field

$$h_x = \frac{1}{N}\sum_{i\in\Omega} z_i J_x \tag{3}$$

assigned on the whole lattice.

Next we turn to the interactions between the buildings of distinct categories in an area. Such on-site interlayer coupling terms are written in the most general form:

$$\begin{aligned}\mathcal{H}_{\text{int}} = &- K_{rw}\sum_i \sigma_{i,r}\sigma_{i,w} - K_{wc}\sum_i \sigma_{i,w}\sigma_{i,c} \\ &- K_{cr}\sum_i \sigma_{i,c}\sigma_{i,r}.\end{aligned} \tag{4}$$

According to the analysis in the previous section, both residential and workplace buildings tend to attract commercial buildings. We thus assume ferromagnetic interactions between the commercial layer and other layers, i.e., $K_{wc} > 0$ and $K_{cr} > 0$. This assumption is indeed consistent with the observed phenomenon that commercial land uses tend to locate near residential or working areas having a high level of floating populations [37]. Because profits generated by commercial activities can stem from distributing products and services, one of the most important factors for the location of the commercial property is the geographical proximity to its demands, consuming products or services. On the other hand, residential and workplace buildings appear mutually exclusive, implying that the coupling between residential and workplace layers is antiferromagnetic: $K_{rw} < 0$. As is well known, residences tend to avoid locating near industrial uses since industrial facilities generate negative externality, causing noise, pollution, and traffic jams. Thus, even though there exists a certain level of housing demands (e.g., factory workers may prefer to live close to their working places), people are hesitant to reside near these facilities in general due to their negative impact on the environment [38]. The schematic diagram of the on-site interlayer couplings is given in Fig. 4(a). It is remarkable that the interlayer coupling is frustrated; otherwise (i.e., if all couplings are ferromagnetic), the system should consist mostly of the rwc state (at the center) and the 0 state (at the boundary). Then the modular structure of the urban area could not be explained.

Finally, we arrive at the total Hamiltonian in the form

$$\mathcal{H}(\{\sigma_{i,x}\}) = \mathcal{H}_r + \mathcal{H}_w + \mathcal{H}_c + \mathcal{H}_{\text{int}}, \tag{5}$$

which, with the simplifying assumption $-K_{rw} = K_{wc} = K_{cr} \equiv K$ and $J_x \equiv J$, takes the simple form

$$\begin{aligned}\mathcal{H}(\{\sigma_{i,x}\}) = &-h\sum_i(\sigma_{i,r} + \sigma_{i,w} + \sigma_{i,c}) \\ &- J\sum_{\langle i,j\rangle}(\sigma_{i,r}\sigma_{j,r} + \sigma_{i,w}\sigma_{j,w} + \sigma_{i,c}\sigma_{j,c}) \\ &- K\sum_i(-\sigma_{i,r}\sigma_{i,w} + \sigma_{i,w}\sigma_{i,c} + \sigma_{i,c}\sigma_{i,r}).\end{aligned} \tag{6}$$

Here the effects of the external field at the boundary have been considered by fixing boundary conditions: $\sigma_{i,x} = -1$ at the boundary, as illustrated in Fig. 4(b). Then Eq. (3) determines completely the uniform field h, and only the coupling constants J and K remain as control parameters.

Remarkably, the frustrated interlayer coupling explains directly the rarity of c and rw states observed in the building data [see Fig. 3(d)]. For convenience of notation, we define $\tilde{\sigma}^{\pm}_{i,x} \equiv (1 \pm \sigma_{i,x})/2$, which takes values 0 or 1 depending on the state of the site, and compute the total number of sites of state $\vec{\sigma}$ through the use of the relation $N_{\vec{\sigma}} = \sum_i \tilde{\sigma}^{\pm}_{i,r}\tilde{\sigma}^{\pm}_{i,w}\tilde{\sigma}^{\pm}_{i,c}$ (where the signs are chosen in accord with the state $\vec{\sigma}$). Expressing the numbers of rw and c states in terms of the spin states, we then have

$$\begin{aligned}N_{rw} + N_c &= \sum_i(\tilde{\sigma}^{+}_{i,r}\tilde{\sigma}^{+}_{i,w}\tilde{\sigma}^{-}_{i,c} + \tilde{\sigma}^{-}_{i,r}\tilde{\sigma}^{-}_{i,w}\tilde{\sigma}^{+}_{i,c}) \\ &= \frac{N}{4} + \frac{1}{4}\sum_i(\sigma_{i,r}\sigma_{i,w} - \sigma_{i,w}\sigma_{i,c} - \sigma_{i,c}\sigma_{i,r}).\end{aligned} \tag{7}$$

Accordingly, the third term, depending on the interlayer coupling constant K, of the Hamiltonian in Eq. (6) can be rewritten in the form $4K(N_{rw} + N_c)$, with the constant term

$4KN$ neglected. This manifests that the coupling constant K controls the fraction of c and rw states; in particular, $N_{rw}, N_c \to 0$ as $K \to \infty$.

Note here that the eight states of the system are connected to each other via symmetries of the system. In the thermodynamic limit ($N \to \infty$), the boundary effects become negligible, and the uniform field r also becomes vanishingly small. In consequence the Hamiltonian given by Eq. (6) carries $Z_2 \times Z_3$ symmetry: The three states rwc, r, and w, as well as the three states 0, wc, and cr, are connected by Z_3 symmetry while the two states in each of the three pairs (rwc and 0 states, r and wc states, and w and cr states) are connected by Z_2 symmetry. Depending on the values of J and K, the system may thus exhibit ordered or disordered states together with the intermediate critical state.

B. Monte Carlo simulations and mean-field approximation

To probe the emergent phases of the model, we first carry out Monte Carlo simulations under the fixed boundary conditions and weak external fields. Fractal lattices with lattice constant $a = 138$, 183, 244, 325, and 432 m are used; 1×10^5 Monte Carlo steps are performed to equilibrate the system and another 1×10^5 steps performed to compute the ensemble averages.

To make use of the $Z_2 \times Z_3$ symmetry inherent in the Hamiltonian, we introduce q-fold symmetry order parameters m_q ($q = 2,3,6$) and their fluctuations χ_q. In view of the symmetry among the three categories ($x = r$, w, and c), we take into account the state $\vec{\sigma}_i = (1,1,1)$ and its symmetric counterparts and define the order parameters as

$$m_2 \equiv \left\langle \left| \frac{1}{N} \sum_i (\tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^+ - \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^-) \right| \right\rangle, \tag{8}$$

$$m_3 \equiv \left\langle \left| \frac{1}{N} \sum_i (\tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^+ + e^{2\pi i/3} \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^- + e^{4\pi i/3} \tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^-) \right| \right\rangle, \tag{9}$$

$$m_6 \equiv \left\langle \left| \frac{1}{N} \sum_i (\tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^+ + e^{\pi i/3} \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^+ + e^{2\pi i/3} \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^- + e^{\pi i} \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^- + e^{4\pi i/3} \tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^- + e^{5\pi i/3} \tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^+) \right| \right\rangle, \tag{10}$$

where $\langle \cdots \rangle$ stands for the ensemble average. Then their fluctuations are measured by

$$\chi_2 \equiv \left\langle \left| \frac{1}{N} \sum_i (\tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^+ - \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^-) \right|^2 \right\rangle - \left\langle \left| \frac{1}{N} \sum_i (\tilde{\sigma}_{i,r}^+ \tilde{\sigma}_{i,w}^+ \tilde{\sigma}_{i,c}^+ - \tilde{\sigma}_{i,r}^- \tilde{\sigma}_{i,w}^- \tilde{\sigma}_{i,c}^-) \right| \right\rangle^2 \tag{11}$$

and by χ_3 and χ_6 defined in the same manner. Figures 5(a) and 5(b) display the obtained results of m_q and χ_q as the intralayer coupling constant J is varied. Interestingly, we observe two distinctive phase transitions at $J = 0.24$ and at $J = 0.30$ when the interlayer coupling is taken to be $K = 0.2$. We have also performed simulations of the two-dimensional regular lattice with periodic boundary conditions, varying K, up to the linear size of $L = 1024$, and obtained similar behaviors (results not shown). Rather than computing the rigorous phase diagram employing the finite-size scaling, we focus on the accompanying phenomena applicable to urban morphology, computing the order parameter m_x and its fluctuations χ_x:

$$m_x = \left\langle \frac{1}{N} \sum_i \sigma_{i,x} \right\rangle, \quad \chi_x = \left\langle \left(\frac{1}{N} \sum_i \sigma_{i,x} \right)^2 \right\rangle - \left\langle \frac{1}{N} \sum_i \sigma_{i,x} \right\rangle^2. \tag{12}$$

As shown in Figs. 5(c) and 5(d), layer c corresponding to the commercial category displays rather distinctive behavior compared with layers r (residential category) and w (workplace category): The symmetry is broken earlier at a lower value of J, giving rise to larger values of m_c than m_r and m_w in the ordered phase. As expected, fluctuations exhibit peaks at the critical points estimated from the behavior of χ_q.

To elucidate this observation, we take the mean-field approximation: Neglecting fluctuations around the mean value m_x, we write the mean-field Hamiltonian in the form

$$\begin{aligned} \mathcal{H}_{\mathrm{MF}} = &-\left(Jm_r - \frac{K}{2} m_w + \frac{K}{2} m_c + h \right) \sum_i \sigma_{i,r} \\ &- \left(Jm_w - \frac{K}{2} m_r + \frac{K}{2} m_c + h \right) \sum_i \sigma_{i,w} \\ &- \left(Jm_c + \frac{K}{2} m_r + \frac{K}{2} m_w + h \right) \sum_i \sigma_{i,c}. \end{aligned} \tag{13}$$

It is then easy to obtain the self-consistent equations for the order parameters:

$$\begin{aligned} m_r &= \tanh \left(2dJm_r - \frac{K}{2} m_w + \frac{K}{2} m_c + h \right), \\ m_w &= \tanh \left(2dJm_w - \frac{K}{2} m_r + \frac{K}{2} m_c + h \right), \\ m_c &= \tanh \left(2dJm_c + \frac{K}{2} m_r + \frac{K}{2} m_w + h \right). \end{aligned} \tag{14}$$

In the thermodynamic limit, we set $h = 0$ and take only positive solutions, i.e., $m_r, m_w, m_c \geqslant 0$, observable in simulations. In the ordered phase, we find $m_r = m_w \equiv m_s$ and $m_c \equiv m_l$ with

$$\begin{aligned} m_s &= \tanh \left[2dJm_s + \frac{K}{2}(m_l - m_s) \right] \\ &= \tanh \left[2dJm_s - \frac{K}{2} m_s + \frac{K}{2} m_l \right], \\ m_l &= \tanh[2dJm_l + Km_s] \\ &= \tanh \left[2dJm_l - \frac{K}{2} m_l + \frac{K}{2} m_l + Km_s \right], \end{aligned} \tag{15}$$

where $m_l > m_s$. Solving the equations numerically with the fractal dimension $d = 1.8$, we obtain the critical value

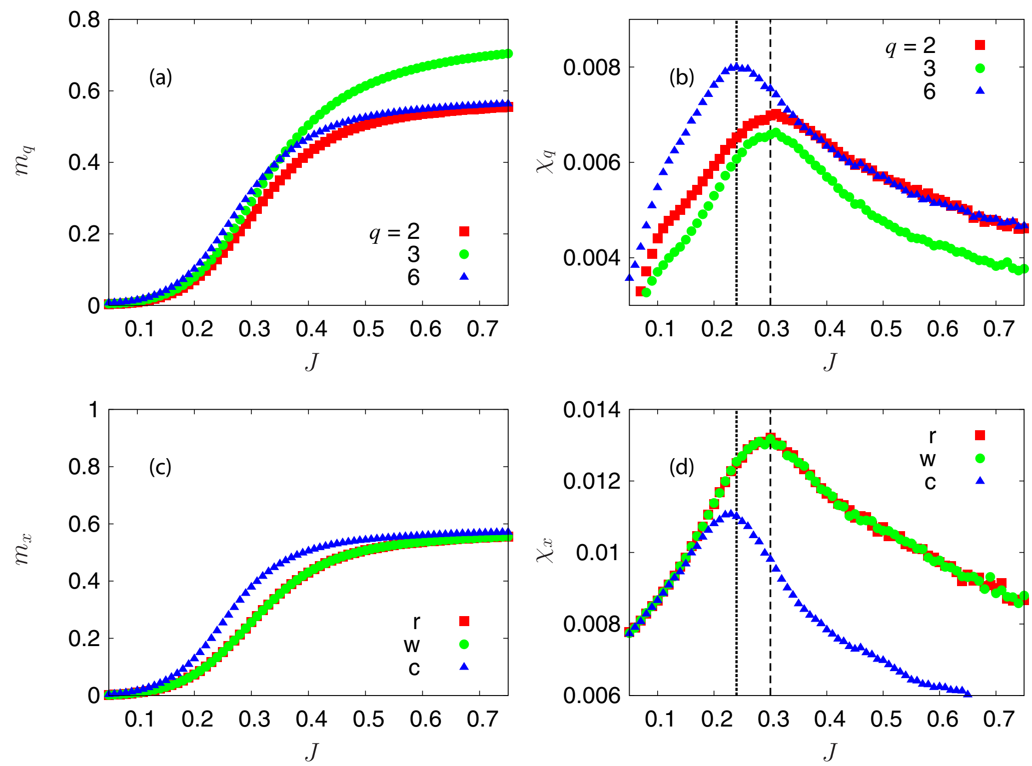

FIG. 5. Order parameters (a) m_q and (c) m_x and their fluctuations (b) χ_q and (d) χ_x. The dotted and dashed vertical lines in (b) and (d) indicate the values of J at which χ_q reaches the peak: $J = 0.24$ for $q = 6$ and $J = 0.30$ for $q = 2,3$. It is evident that m_r and m_w (similarly, χ_r and χ_w) exhibit the same behavior while χ_c peaks at a smaller value of J than χ_r and χ_w. The lattice constant and the interlayer coupling constant are taken to be $a = 138$ m and $K = 0.2$, respectively.

$J_c = 0.250$ for $K = 0.2$, which coincides with the Z_6-symmetry-breaking point $J = 0.24$ in Monte Carlo simulations. Note, however, that the intermediate phase emerging for $0.24 < J < 0.30$ is not verified in the mean-field approximation and still remains to be investigated.

C. Criticality in the urban morphology of Seoul

Finally, in this section the model predictions are compared with the results from data analysis. We first introduce the "macroscopic state" of the system described by the eight-component state vector

$$\mathbf{S} \equiv \frac{1}{N_S}(N_0, N_r, \ldots, N_{rwc}), \tag{16}$$

where $N_S \equiv [N_0^2 + \cdots + N_{rwc}^2]^{1/2}$ is the normalization factor. In other words, while the microscopic state details the state of every site and is thus characterized by N (microscopic) state vectors $\vec{\sigma}_i$ identifying the state of site i (=1,2, ..., N), the macroscopic state specifies just the number of sites in each of the eight states $\vec{\sigma}$. Comparison of the macroscopic state vectors $\mathbf{S}_\mathrm{M}$ and $\mathbf{S}_\mathrm{D}$, obtained via simulations of the model and from the building data, respectively, should expose the discrepancy between the simulation results and the real data. In particular, the inner product of the two state vectors provides a convenient measure of how faithful the model description is, and we thus define the accuracy factor $Q \equiv \mathbf{S}_\mathrm{M} \cdot \mathbf{S}_\mathrm{D}$, quantifying the accuracy of the simulation results. Here we encounter again the threshold issue. To treat the simulation results and data on an equal footing, we use the order parameters from simulations to fix the threshold when coarse-graining the building data. Specifically, the positive state is assigned to the first $\frac{m_x+1}{2}N$ sites in descending order of area.

Note that the accuracy factor also manifests the criticality of the system. As shown in Fig. 6, the accuracy factor reaches the maximum in the range $0.24 < J < 0.30$. In other words, simulations reproduce the building data best in the intermediate phase between breaking of Z_6 symmetry and that of Z_2 or Z_3 symmetry. In the critical region, the accuracy

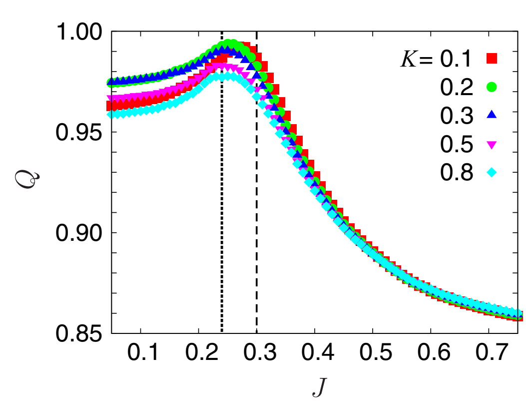

FIG. 6. Accuracy factor Q versus intralayer coupling constant J for various values of the interlayer coupling constant K. Q reaches the peak in the vicinity of the Z_6 symmetry breaking, revealing the criticality of the city. The lattice constant is given by $a = 138$ m.

TABLE II. Fractions of states $N_{\vec{\sigma}}/N$ for several values of the lattice constant, extracted from the Seoul building data and from numerical simulations of the model with coupling constants J and K at which Q is maximum.

	Building Data					Simulations				
State $\vec{\sigma}$	138 m	183 m	244 m	325 m	432 m	138 m	183 m	244 m	325 m	432 m
$rwc\,(1,1,1)$	0.276	0.409	0.502	0.563	0.623	0.270	0.402	0.488	0.556	0.622
$r\,(1,-1,-1)$	0.126	0.092	0.070	0.054	0.042	0.101	0.067	0.054	0.047	0.038
$w\,(-1,1,-1)$	0.083	0.066	0.056	0.050	0.036	0.101	0.067	0.054	0.047	0.038
$0\,(-1,-1,-1)$	0.123	0.106	0.107	0.116	0.125	0.123	0.116	0.119	0.127	0.132
$wc\,(-1,1,1)$	0.154	0.121	0.095	0.071	0.056	0.157	0.125	0.104	0.081	0.062
$cr\,(1,-1,1)$	0.133	0.107	0.081	0.068	0.051	0.157	0.125	0.104	0.081	0.062
$c\,(-1,-1,1)$	0.064	0.058	0.050	0.043	0.035	0.043	0.042	0.032	0.026	0.020
$rw\,(1,1,-1)$	0.041	0.041	0.039	0.035	0.032	0.048	0.056	0.045	0.035	0.026

factor becomes almost unity ($Q \approx 1$). Table II summarizes the distribution ratios of the eight different states, obtained from the building data and from the model, for various values of the lattice constant a. The coupling constants J and K are adjusted to the values at which Q is a maximum for given a. Note the remarkable agreement in the state distributions between simulations and real data.

Encouraged by the successful results of the model, we are now in a position to examine the urban morphology generated by numerical simulations. For reference, we first analyze the Seoul building data set by applying the threshold value obtained from the simulation results with the lattice constant $a = 138$ m. The extracted urban morphology of Seoul City is displayed in Fig. 7(a). Note that the results are consistent with those presented in Fig. 3(c) even though the threshold value and the lattice constant are different. We have also confirmed the consistency of urban morphology across the threshold values and coarse-graining scales in proper ranges.

When analyzing simulation results, on the other hand, one should ponder over social fluctuations and collective dynamics operating in the urban area because via Monte Carlo simulations we can probe only equilibrium fluctuations rather than the actual dynamics of the urban system. Approximately, one may regard individual choices as noise fluctuating on short time scales. Observed phenomena, however, would change rather gradually, exhibiting temporal correlations as we probe

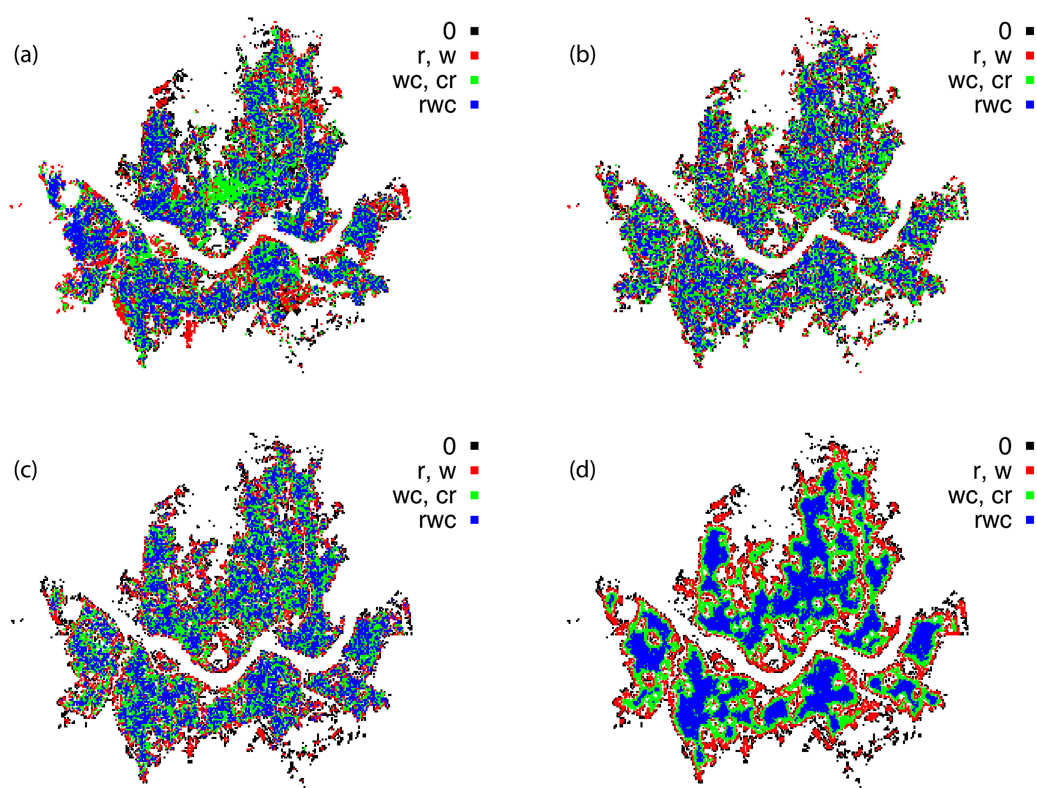

FIG. 7. Urban morphology of Seoul City, extracted from (a) the Seoul building data set and (b)–(d) Monte Carlo simulation results. The number of ensembles used in the averaging process is (b) 1 (snapshot, short-time limit), (c) 20 (intermediate time), and (d) 1000 (long-time limit). Fluctuations are mainly observed in the short-time limit while a rather regular pattern is extracted in the long-time limit. Both results are not quite appropriate due to the quasiequilibrium nature of the system.

the macroscopic properties or the morphology of the city, originating from the interactions between individuals. As a result, macroscopic dynamics of the system is an emergent property which reflects the averaged effects of microscopic fluctuations. Here we assume that the urban system is in a quasiequilibrium state and the individuals may feel mean-field-like couplings for the duration of time during which the system resides near equilibrium. Furthermore, it is obvious that there exists a time delay in the administrative procedure of data gathering and planning. We therefore conclude that the observed urban morphology is a consequence of averages over the time associated with decision making and executing as well as observations. However, the time scale should not be infinitely long as is usually assumed in equilibrium statistical mechanics. According as technology, relations to other cities, and natural environment evolve in time, the city should develop and come out of the current quasiequilibrium state, manifesting its nonequilibrium nature. If the actual trajectory of the system should be followed in the phase space, the relevant averaging time scale could be specified. Leaving this for further study, we here take the average of the results of Monte Carlo simulations, varying the number of ensembles, and thus probe the macroscopic patterns explicitly.

The detailed procedure taken to prepare the figure is summarized as follows: We first evaluate the frequency of each state appearing in the system and choose the lattice sites where the frequency of the given state is high. In some sites, there could be more than one state whose frequencies are high enough to be depicted while there also exist sites for which no states display high frequencies. To assign a single state to each site, we first choose the highest one among overlapped states in the former case, and leave the states of sites undetermined in the latter case. Then the above frequency comparison process is repeated only with the remaining sites whose states are not determined until the state of every site is specified. Note that we enforce the number of sites in a given state, i.e., the fraction of a state, to equal the ensemble average value computed from the simulation data displayed in Table II. Accordingly, if the fraction of a state is once fulfilled, the state is also excluded in the repeated procedure. In this manner, we specify the most probable state in every site while constraining the fractions of states to be the ensemble average values.

Figure 7 presents the obtained results, where the number of ensembles used in the averaging process is given by 1 [Fig. 7(b), snapshot, short-time limit], 20 [Fig. 7(c), intermediate time], and 1000 [Fig. 7(d), long-time limit]. As shown, it is successfully simulated to locate the sites of r and w states which are distributed mainly around the boundary, irrespective of the averaging time. These results reveal unambiguously the interplay between boundary conditions and overall external fields. Also observed are clustering patterns, especially of the rwc state, which become clearer as the averaging time is increased. As expected, the clustering patterns resemble the results from the data better in the intermediate time scale.

On the other hand, emergence of the CBD clusters consisting mostly of the wc state is not observed in the simulation. This is presumably due to the fact that the CBD of Seoul involves historical formation and land price issues. To observe the emergence of the wc-state cluster, we may thus need to apply an additional field, taking into consideration such issues in the CBD area. Furthermore, working areas are apparently categorized into further details. Specifically, the working areas may be split into two representative subcategories: Most of the working areas located on the edge of Seoul are either factories or warehouses. As is well known, many industrial facilities tend to locate farther from the center of the city, since they prefer to consume large spaces of the land at cheaper prices [39]. In contrast, office facilities tend to cluster at certain areas in a city and contribute to formation of the CBD cluster, which is generated by agglomeration economies. Agglomeration economies facilitate more face-to-face interactions, sharing inputs, and integrating the labor market; these factors reintensify agglomeration economies, resulting in the clustering of office firms [40]. With the external field issues put together, the historically quenched area of the city may serve as a seed for cluster formation.

To conclude, properly averaged simulation results, e.g., Fig. 7(c), successfully describe the overall differentiated patterns and fluctuating features (except for the formation of the CBD cluster due to the aforementioned issues): The explicitly broken translational symmetry due to the boundary conditions contributes to the formation of the spatial structures while the criticality of the system accounts for the statistics and fluctuating patterns. We also expect that these simulation results may shed light on the future land use pattern of Seoul unless the system is perturbed dramatically. For example, we observed recently evidence for reformation and population growth in downtown areas. Also the mixed land use issue attracts much attention in the context of urban planning. As shown in Fig. 7(d), the clustered mixed use (rwc-state) areas are observed ubiquitously in the central regions of cities and in fact are coincident with the actual downtown areas of Seoul as well.

V. DISCUSSION

We have analyzed the building data set of Seoul City to extract the land use pattern. Based on the interactions between buildings, we have constructed a three-layer Ising model, through which urban morphology has been studied. To be specific, we have performed coarse-graining of the system, which is described not only by the urbanized area but by three categories of building. The resulting coarse-grained system has been probed and the emergence of distribution patterns capturing urban morphology has been attributed to the inter- and intralayer couplings between the buildings. In particular, the coupling between residential areas and workplaces has turned out to be mutually repulsive or antiferromagnetic. Furthermore, it has been confirmed that boundary conditions are the key ingredients for generating the urban spatial structure. With minimal postulates, as summarized above, we have indeed reproduced successfully the urban morphology of Seoul.

As the model requires only the social or natural boundary of the city, one can simulate the morphology of other cities around the world and examine the applicability of the model. We further stress here the distinctive characteristics of commercial land uses. Related to the recent urban planning issue to improve the proximity of jobs to housing, the locations and distributions of commercial land uses may play an important role in urban formation, minimizing the distances between the work and residential areas.

Meanwhile, there still remain several issues to resolve or improve. First of all, we point out that the detailed phase diagram is not obtained. Even though the behavior of the model at criticality indeed resembles that observed in the real data set, we have not evaluated rigorously the critical behavior. Comparing the exponents of the model with the building data of various cities, one may probe the possibility of the universality class among different cities. Furthermore, if the correlation function, for example, could be evaluated accurately, one may predict responses of a city to external perturbations such as natural and/or social disasters.

Second, note that the fractal lattice constructed of the Seoul building data set is not fully connected in general. Most of all, the Han (river), divides the urbanized area of Seoul into two separate parts. As a consequence, each part of the lattice is decoupled and completely independent of each other if the coupling is assumed to exist between nearest-neighboring sites. In fact it is well known that the two parts of Seoul, usually called Gangbuk (northern part) and Gangnam (southern part), indeed exhibit endemic characteristics. The Han, approximately 1–2 km wide estimated by the length of the bridges, is rather wider than most rivers in other cities and is therefore likely to generate boundary effects. Obviously, however, Seoul should be regarded as a single city. A simple alternative is to modify the intralayer coupling to be long ranged, e.g., $J(r) \sim 1/r^{\alpha}$ with appropriate exponent α. On the other hand, one may introduce disordered coupling J_{ij}, reflecting the transportation infrastructure such as the subway or road network. For example, we may assign a positive coupling constant J_{ab} between the areas a and b connected by a bridge.

Furthermore, one may assume asymmetric interlayer couplings, e.g., $K_{wc} \neq K_{cr}$, to elucidate the difference between the distributions of r and w states (or wc and cr states). As shown in Fig. 3(d), the states are not evenly distributed among six other states as well as c and rw states. Since the relation between the interlayer coupling terms and the average number of a state could be obtained as in Eq. (7), one can extend the Hamiltonian to model the state distributions in more detail, according to the specific objective of the study. Such asymmetric couplings may indeed arise in the discrete spin description of the system, since the spin state $\sigma_{i,x} = 1$ does not necessarily represent the same amount of buildings or, more essentially, of population, irrespective of category x. For instance, commercial buildings are generally larger in floor areas than residential buildings and are expected to attract more people, which suggests that $K_{wc} > K_{rw}$. In this regard, one may consider an extended description [41] to take such heterogeneity into account.

Finally, the origin of the criticality still remains unanswered. In fact the coupling via transportation which enables the urban movements may play a crucial role in controlling fluctuations and/or the social temperature of the urban morphology. We speculate that, if the urban morphology corresponds to an ordered state, there would be too heavy a load on the transportation system. It may cause redistribution of the building to minimize the transportation cost. If, on the other hand, the urban morphology remains in the disordered state, the urban system may not function properly. To verify this hypothesis is left for further study.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea through NRF Grant No. 2013S1A5A2A01017764 (K.L.) and through the Basic Science Research Program [Grants No. 2014R1A6A3A01059467 (S.G.), No. 2011-0012331 and No. 2015K2A1A2070452 (M.Y.C.), and No. 0434-20140013 (K.K.)].

[1] R. White and G. Engelen, Environ. Plann. A **25**, 1175 (1993).
[2] H. A. Makse, J. S. Andrade, M. Batty, S. Havlin, and H. E. Stanley, Phys. Rev. E **58**, 7054 (1998).
[3] M. Batty, Environ. Plann. B **28**, 123 (2001).
[4] C. Andersson, K. Lindgren, S. Rasmussen, and R. White, Phys. Rev. E **66**, 026204 (2002).
[5] A. Bitner, R. Holyst, and M. Fialkowski, Phys. Rev. E **80**, 037102 (2009).
[6] G. F. Frasco, J. Sun, H. D. Rozenfeld, and D. ben-Avraham, Phys. Rev. X **4**, 011008 (2014).
[7] A. P. Masucci, E. Arcaute, E. Hatna, K. Stanilov, and M. Batty, J. R. Soc. Interface **12**, 20150763 (2015).
[8] J.-P. Bouchaud, Nature (London) **455**, 1181 (2008).
[9] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).
[10] L. M. Bettencourt, Science **340**, 1438 (2013).
[11] S. Goh, K. Lee, M. Choi, and J.-Y. Fortin, PLoS One **9**, e89980 (2014).
[12] Y. Chen, Chaos Solitons Fractals **45**, 115 (2012).
[13] R. Louf and M. Barthelemy, Phys. Rev. Lett. **111**, 198702 (2013).
[14] A. Gudmundsson and N. Mohajeri, Sci. Rep. **3**, 3324 (2013).
[15] R. Louf, C. Roth, and M. Barthelemy, PLoS One **9**, e102007 (2014).
[16] J. Candia, M. C. González, P. Wang, T. Schoenharl, G. Madey, and A.-L. Barabási, J. Phys. A **41**, 224015 (2008).
[17] C. Kang, X. Ma, D. Tong, and Y. Liu, Physica A **391**, 1702 (2012).
[18] M. Batty, R. Carvalho, A. Hudson-Smith, R. Milton, D. Smith, and P. Steadman, Eur. Phys. J. B **63**, 303 (2008).
[19] M. Schläpfer, J. Lee, and L. Bettencourt, arXiv:1512.00946.
[20] G. Duranton and H. G. Overman, Rev. Econ. Stud. **72**, 1077 (2005).
[21] P. Jensen, Phys. Rev. E **74**, 035101 (2006).
[22] E. Marcon and F. Puech, J. Econ. Geogr. **10**, 745 (2009).
[23] *2011 Building Register Data, Electronic Architectural Information System* (Ministry of Land, Infrastructure, and Transport, Republic of Korea, 2011).
[24] S. Fotheringham, Environ. Plann. A **21**, 1447 (1989).
[25] M. Batty and Y. Xie, Discrete Dyn. Nat. Soc. **3**, 109 (1999).
[26] T. A. Witten, Jr., and L. M. Sander, Phys. Rev. Lett. **47**, 1400 (1981).

[27] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[28] G. Shen, Int. J. Geogr. Inf. Sci. **16**, 419 (2002).
[29] J. Decraene, C. Monterola, G. K. K. Lee, and T. G. G. Hung, Int. J. Mod. Phys. C **24**, 1250092 (2013).
[30] J. Decraene, C. Monterola, G. K. K. Lee, T. G. G. Hung, and M. Batty, PLoS One **8**, e80309 (2013).
[31] J. Park, D.-S. Lee, and M. C. González, J. Stat. Mech. (2010) P11021.
[32] M. Lee and P. Holme, PLoS One **10**, e0140152 (2015).
[33] M. L. Daggitt, A. Noulas, B. Shaw, and C. Mascolo, R. Soc. Open Sci. **3**, 150688 (2016).
[34] M. Batty and Y. Xie, Int. J. Geogr. Inf. Syst. **8**, 291 (1994).
[35] F. Schweitzer, *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences* (Springer-Verlag, Berlin, 2003).
[36] G. Iannone, A. Troisi, C. Guarnaccia, P. D'Agostino, and J. Quartieri, Int. J. Mod. Phys. C **22**, 543 (2011).
[37] D. Mushinski and S. Weiler, J. Reg. Sci. **42**, 75 (2002).
[38] J. D. Burnell, Urban Stud. **22**, 399 (1985).
[39] P. A. Wood, *Suburban Growth: Geographical Processes at the Edge of the Western City*, edited by J. H. Johnson (John Wiley, London, 1974), pp. 129–154.
[40] A. O'Sullivan, *Urban Economics* (McGraw-Hill, New York, 2011).
[41] R. Louf and M. Barthelemy, arXiv:1511.04268.