PHYSICAL REVIEW E **82**, 057101 (2010)

External bias in the model of isolation of communities

Julian Sienkiewicz, Grzegorz Siudem, and Janusz A. Hołyst
Faculty of Physics, Center of Excellence for Complex Systems Research, Warsaw University of Technology, Koszykowa 75, PL-00-662 Warsaw, Poland

(Received 23 August 2010; published 15 November 2010)

We extend a model of community isolation in the *d*-dimensional lattice to a case with an imposed imbalance between the birth rates of competing communities. We provide analytical and numerical evidences that in the asymmetric two-species model there exists a well-defined value of the asymmetry parameter when the emergence of the isolated (blocked) subgroups is the fastest, i.e., the characteristic time t_c is minimal. The critical value of the parameter depends only on the lattice dimensionality and is independent of the system size. A similar phenomenon is observed in the multispecies case with a geometric distribution of the birth rates. We also show that blocked subgroups in the multispecies case are absent or very rare when either there is a strictly dominant species that outnumbers the others or there is a large diversity of species. The number of blocked species of different kinds decreases with the dimension of the multispecies system.

DOI: 10.1103/PhysRevE.82.057101 PACS number(s): 89.75.Hc, 05.50.+q, 02.50.−r

The question of *imbalance* or *asymmetry* is long known in many popular and fundamental phenomena in nonequilibrium statistical physics, such as gradient-induced transfer of particles, heat, or current [1]. In this Brief Report, we extend previously obtained results [2] by drawing attention to the issue of imbalance in a model of isolation of species (communities). As social imbalance has already been of interest to physicists [3,4], we shall try to show the impact that different methods of imbalance introduction have on the number of isolated species.

Recently, we introduced a simple model of community isolation [2] whose basic rules can be described in the following way: in each time step, one puts a representative species (↑) or (↓) in a random unoccupied node in a chain of N nodes. The species can be an individual belonging to a given community or a person characterized by a certain opinion. The probabilities of choosing either of the possible species (community birth rates) are equal to 0.5. If n nodes filled with identical species (e.g., ↓↓↓) are surrounded by individuals belonging to another community (e.g., ↑↓↓↓↑), then the nodes inside the cluster are called *blocked*, and they no longer interact with the rest of the system. This model can be easily extended to a case with multiple different species $m \geq 2$, where the type of the species is drawn from the uniform distribution $[1,m]$. The analytical treatment shows that the number of blocked nodes Z for a specific time of the simulation t, regardless of the number of different species introduced in the chain, grows roughly as t^3 [2]. It also enables the calculation of the *characteristic time* t_c, i.e., the time when the first isolated node appears.

In this Brief Report, we investigate the asymmetric case of the model [2]. Symmetry breaking is introduced as an external bias and is received by changing birth rates of different species occurring in the system, either by a simple imbalance in the two-species case or by setting a specific species probability distribution in the multispecies case. We shall consider the model of isolation of communities on a d-dimensional lattice where clusters of nodes sharing the same species will be isolated when neighbors (using von Neumann's sense of neighborhood [5]) of all the nodes in the cluster are connected with each other or with a different species.

First, we consider two species (↑ and ↓) that we put into d-dimensional hypercubic lattice with a total number of sites $N=L^d$. Occurrence probabilities $p_\uparrow$ and $p_\downarrow$ of both species are nonequal and given by $p_\uparrow=0.5+\epsilon$ and $p_\downarrow=0.5-\epsilon$, where $\epsilon\in[0,0.5]$ is the symmetry-breaking parameter. These probabilities are microscopic parameters that describe model evolution. Starting from a lattice of empty nodes, after t time steps the probabilities that a randomly picked node is occupied with the species (↑) or (↓) are $\text{Prob}(\uparrow)=p_\uparrow t/N$ and $\text{Prob}(\downarrow)=p_\downarrow t/N$.

Now, let us consider number of blocked nodes $Z^\uparrow$ and $Z^\downarrow$ of both species at time t. These numbers are well approximated by numbers of individual blocked nodes, i.e., by numbers of blocked clusters of size one of both species. When the total density of all blocked nodes is small ($Z^\uparrow+Z^\downarrow \ll t$), then

$$Z^\uparrow \approx Z_1^\uparrow = (L-2)^d \text{Prob}^{2d}(\downarrow)\text{Prob}(\uparrow), \quad (1)$$

$$Z^\downarrow \approx Z_1^\downarrow = (L-2)^d \text{Prob}^{2d}(\uparrow)\text{Prob}(\downarrow). \quad (2)$$

Substituting the values Prob(↑) and Prob(↓) and taking into account the form of $p_\uparrow$ and $p_\downarrow$, the above relations for $L \gg 2$ may be expressed as

$$Z^{\uparrow,\downarrow} \approx \frac{t^{2d+1}}{N^{2d}}\left(\frac{1}{4}-\epsilon^2\right)\left(\frac{1}{2}\mp\epsilon\right)^{2d-1}, \quad (3)$$

where $Z^\uparrow$ corresponds to the "−" sign in the second brackets. The above equations allow us to calculate characteristic times $t_c^\uparrow$ and $t_c^\downarrow$ when the first representative of either species emerges. Putting $Z^\uparrow=1$ (analogously for ↓), we obtain after a short algebra

$$t_c^{\uparrow,\downarrow} = \left[\frac{N^{2d}}{\left(\frac{1}{4}-\epsilon^2\right)\left(\frac{1}{2}\mp\epsilon\right)^{2d-1}}\right]^{1/(2d+1)}, \quad (4)$$

where $t_c^\uparrow$ corresponds to the "−" sign in the second brackets.

If we consider more than two species, then we can use probability distribution Prob(i), instead of the previously introduced probabilities Prob($\uparrow$) and Prob($\downarrow$). It requires changing microscopic probabilities $p_\uparrow$ and $p_\downarrow$ with the value of p_i. The connection between these two sets of variables is given by $\text{Prob}(i)=p_i t/N$, for $i=1,2,\ldots$ (the case p_i independent of i corresponds to the symmetric problem considered in [2]). Using the same approximation as is used in the two-species case, we can obtain

$$Z^{(i)} \approx Z_1^{(i)} \approx N \sum_{k=1,k\neq i}^{\infty} \text{Prob}(k)^{2d}\text{Prob}(i), \tag{5}$$

for $i=1,2,\ldots$, which becomes

$$Z^{(i)} = \frac{t^{2d+1}}{N^{2d}} p_i \left(\sum_{k=1}^{\infty} p_k^{2d} - p_i^{2d} \right). \tag{6}$$

Putting $Z^{(i)}=1$, one obtains a set of characteristic times for each species i:

$$t_c^{(i)} = \left[\frac{N^{2d}}{p_i \left(\sum_{k=1}^{\infty} p_k^{2d} - p_i^{2d} \right)} \right]^{1/(2d+1)}, \quad i = 1,2, \ldots . \tag{7}$$

Here, we are presenting the comparison between the analytical approach given above and the results obtained from the numerical simulations performed for topologies of d-dimensional lattices.

Investigation of $Z^\downarrow$ and $Z^\uparrow$ as functions of the parameter ϵ for a fixed value of time t reveals some surprising results. Figure 1(a) shows the dependence of the number of blocked nodes of each species versus the symmetry-breaking parameter ϵ obtained in the numerical simulations compared with the theoretical expectations given by Eq. (3). The discrepancies between the simulation results and the theoretical approach can be justified by the approximations used in obtaining $Z^\downarrow$ and $Z^\uparrow$. Still, one can immediately spot the main difference between those two quantities: $Z^\uparrow$ is monotonic while $Z^\downarrow$ first increases, reaches a prominent and well-defined maximum, and then drops down. This observation is backed with a simple analysis of Eq. (3); in fact, the derivative $\partial Z_\uparrow / \partial \epsilon < 0$ for the whole range $\epsilon \in [0, 0.5]$, and in the case of $Z^\downarrow$ there is a maximum value for

$$\epsilon^* = \frac{1}{2}\frac{2d-1}{2d+1}. \tag{8}$$

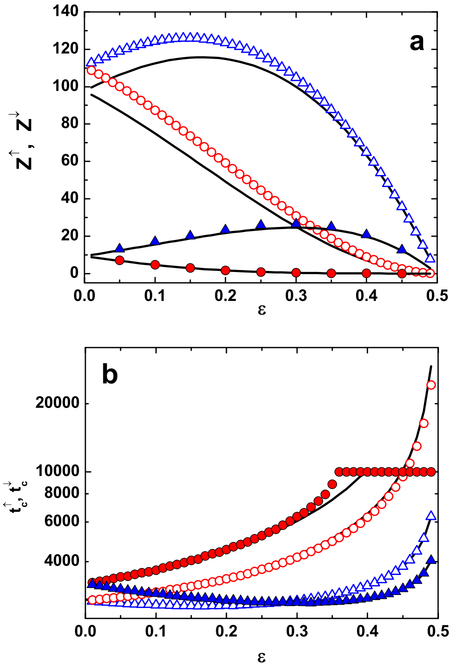

FIG. 1. (Color online) (a) Number of blocked nodes $Z_\uparrow$ and $Z_\downarrow$ versus symmetry-breaking parameter ϵ on a chain (open symbols, simulations performed for $t=N/4$) and on a square lattice (filled symbols, simulations performed for $t=N/2$); circles ($Z_\uparrow$) and triangles ($Z_\downarrow$) correspond to simulation results and lines are calculated using Eq. (3). (b) Logarithmic-linear plots of the characteristic times $t_c^{(\uparrow)}$ (circles) and $t_c^{(\downarrow)}$ (squares) for $d=1$ (filled symbols) and $d=2$ (open symbols). Data points are taken from numerical simulations, while solid lines come from Eq. (4). For $d=1$ simulations were performed for $N=5\times10^4$ nodes, averaged over 10^5 runs, while for $d=2$ it was $N=10^4$ nodes, averaged over 10^3 runs.

The decrease in $Z^\uparrow$ is rather obvious, as it is the dominant species—the higher is the number of its individuals that are introduced in the system, the smaller is the probability of this species being blocked. On the other hand, as it concerns $Z^\downarrow$, Eq. (8) suggests that there exists a specific value of ϵ for which the number of blocked individuals reaches the highest rate. It follows that this value depends only on the dimensionality of the system (i.e., the number of the nearest neighbors). This phenomenon is even more pronounced while examining the characteristic times $t_c^\uparrow$ and $t_c^\downarrow$ versus ϵ presented in Fig. 1(b). The first quantity exhibits a constant growth, and the second one possesses a clear minimum for $\epsilon=\epsilon^*$, which is consistent with the maximum values for $Z^\downarrow$ observed in Fig. 1(a).

For the asymmetric multispecies case, we used the geometric probability distribution $p_i=q^{i-1}(1-q)$, where $i=1,2,\ldots$. This very case of probability distribution has been chosen as an example because it easily yields the analytical approach. Moreover, it can be transformed directly into continuous exponential distribution assuming that $q=\exp(-\alpha)$. First, we shall discuss the issue of the global characteristic time t_c. It can be approximately calculated by assuming that the total number of blocked species at time t is equal to $Z(t)=\Sigma_i Z^{(i)}=1$. Thus, the global characteristic time can be expressed as

$$t_c = \left[\frac{N^{2d}}{\frac{(1-q)^{2d}}{1-q^{2d}} - \frac{(1-q)^{2d+1}}{1-q^{2d+1}}} \right]^{1/(1+2d)}. \tag{9}$$

Figure 2 presents the characteristic time versus parameter q given for different values of d and N. In each case the curve has a similar shape, exceeding $t_c=N$ for both small and large values of q (horizontal lines in Fig. 2) with a well-defined minimum in between. Moreover, for a specific value of N, the curves, regardless of the dimensionality of the system,

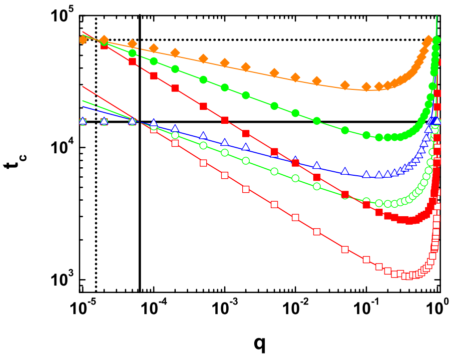

FIG. 2. (Color online) Log-log plot of the global characteristic times t_c versus parameter q of the geometric distribution for $N_1=15\,625$ (empty symbols) and $N_2=65\,536$ (filled symbols). Data points are taken from numerical simulations: squares are $d=1$, circles are $d=2$, triangles are $d=3$, and diamonds are $d=4$; all data points are averaged over 1000 runs. Solid lines come from Eq. (9) and the horizontal solid line is drawn for $t_c=N_1$, whereas the horizontal dotted one is for $t_c=N_2$. The vertical solid line marks $q_{min}=1/N_1$, while the vertical dotted line is drawn for $q_{min}=1/N_2$ (see description in text).

seem to intersect at one point. In fact, a closer analysis reveals that for $q \ll 1$, Eq. (9) takes a form of $t_c=(N^{2d}/q)^{1/(2d+1)}$, which results in the intersection point $q_{min}=1/N$, shown as vertical lines in Fig. 2. A heuristic explanation of this fact can be also derived in the following way: as long as the probability of drawing the second species ($i=2$) is above $1/N$, there is a statistical chance of it appearing in the system and thus being blocked by the overwhelming first species ($i=1$). As soon as the probability drops below that value, there is only one species and so the blocking is impossible. On the other hand, when q approaches 1, the number of different species is relatively high, and all the probabilities p_i are close to p_1. After crossing a certain value q_{max}, the number of the individuals of the first species is too low to make blocking on the lattice possible. The obvious necessary condition preventing this scenario is that $Np_1^{2d}(p_2+p_3+p_4+\cdots)=1$, which leads to $Np_1^{2d}\approx 1$ and eventually gives $q_{max}\approx 1-N^{-1/2d}$. Finally, Fig. 2 suggests that there is some specific value of q for which the characteristic time is the lowest; assuming that $q^{2d}\ll 1$, one can estimate this value with

$$q^* = \frac{1}{2d+1}. \tag{10}$$

We stress that the critical value of q^* given by Eq. (10) is independent of the system size N, which is similar to the critical value of the ϵ^* parameter given by Eq. (8).

One can also focus directly on the question of characteristic times $t_c^{(i)}$ of various species. Substituting Eq. (7) with the geometric distribution, one gets the general formula for the characteristic time of the ith species being blocked in the topology of the d-dimensional lattice:

$$t_c^{(i)} = \left[\frac{N^{2d}(1-q^{2d})q^{1-i}}{(1-q)^{2d+1}(1-q^{2d(i-1)}+q^{2di})}\right]^{1/(2d+1)}. \tag{11}$$

The comparison of the characteristic times $t_c^{(i)}$ versus parameter q obtained in numerical simulations and the theoretical expectations given by Eq. (11) is presented in Fig. 3. The plots for $t_c^{(i)}$ bear close resemblance to those obtained for t_c. Following a similar line of thought as in the case of the global characteristic time, it is possible to estimate the crucial points of these curves. First of all, Eq. (11) can be approximated as

$$t_c^{(1)} \approx \frac{(N^{2d}q^{-2d})^{1/(2d+1)}}{1-q},$$

$$t_c^{(i)} \approx \frac{(N^{2d}q^{1-i})^{1/(2d+1)}}{1-q}, \quad i=2,3,\ldots, \tag{12}$$

which is valid for all $q\in[0,1]$. However, in order to estimate $q\ll 1$ when the above function intersects with $t_c^{(i)}=N$,

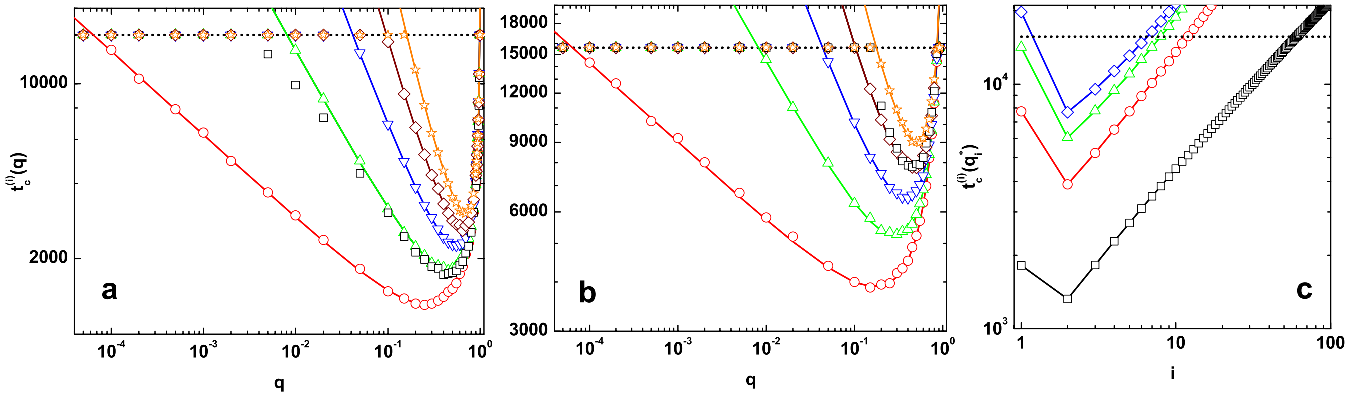

FIG. 3. (Color online) (a) and (b) Log-log plots of characteristic times $t_c^{(i)}$ versus parameter q of the geometric distribution for (a) $d=1$ and (b) $d=2$. Data points are taken from numerical simulations: squares represent species $i=1$, circles: $i=2$, upward triangles: $i=3$, downward triangles: $i=4$, diamonds: $i=5$, and stars: $i=6$; solid lines come from Eq. (11). Theoretical curves for $i=1$ and $i=3$ in one-dimensional case, as well as $i=1$ and $i=5$ in two-dimensional case overlap. In both cases $N=15\,625$ and the horizontal dotted line is drawn for $t_c^{(i)}=N$. Numerical data are averaged over 1000 runs. (c) Log-log plot of the characteristic time $t_c^{(i)}$ calculated from Eqs. (12) for $q=q_i^*$ against species number i. Squares are $d=1$, circles are $d=2$, triangles are $d=3$, and diamonds are $d=4$; in each case $N=15\,625$. Solid lines are for visual guidance and the horizontal dotted line is drawn for $t_c^{(i)}=N$.

one uses a further approximation, i.e., $t_c^{(i)}=N^{2d}q^{(1-i)/(2d+1)}$. It gives a straightforward formula for the intersection point $q_{min}=N^{1/(1-i)}$, which suggests that the intersection point should be dependent only on the number of nodes in the lattice, while its dimension d is irrelevant. Figure 3 also gives clear evidence that the characteristic times for the first species behave differently from those for $i\geq 2$. In both cases, it is the species $i=2$ which gets blocked first for smaller values of q. Moreover, in the case of $d=1$, the results from $t_c^{(1)}$ cover with $t_c^{(3)}$ and in the case of $d=2$ with $t_c^{(5)}$. This last fact is fully comprehensible after comparing both equations in Eqs. (12) where, after short algebra, one finds that the characteristic time for the first and $(2d+1)$th species are the same. One can also notice by comparing Eqs. (9) and (11) that it is, in fact, $t_c^{(2)}$ that plays the dominant role in creating the shape of the global characteristic time t_c and may be used as its good approximation.

Taking into account Eq. (12), it is possible to estimate the maximal species number that gets blocked in the environment for a fixed value of the system size N. The curves presented in Figs. 3(a) and 3(b) indicate that for each i there is a specific value q_i^* for which the function $t_c^{(i)}$ takes its minimum. It follows that as long as $t_c^{(i)}(q_i^*)<N$, the species will be blocked, at least for $q=q_i^*$. Closer analysis of Eq. (12) leads to

$$q_1^* = \frac{2d}{4d+1},$$

$$q_i^* = \frac{i-1}{i+2d}, \quad i=2,3,\ldots, \tag{13}$$

After substituting Eqs. (12) with the above values of q_i^*, one gets the value of characteristic time in the minimum. A corresponding plot for different values of dimension d is shown in Fig. 3(c). It immediately gives the idea of the fast restriction of the blocked species number with a system's dimension: whereas for $d=1$ the value of i can be as large as 70, in the case of $d=4$, it substantially drops down to 6. Furthermore, Fig. 3(c) underscores the specific role of the first species discussed in the previous paragraph.

In conclusion, we have extended the simple model of community isolation to the case with the symmetry breaking. Our calculation and numerical simulations show that even a simple way of introducing the external bias between species can lead to interesting and nontrivial results. We have found that in both the two-species case where a parameter ϵ governs the symmetry breaking and the multispecies case where the numbers of each entity are given by the geometric distribution, there exists some specific and well-defined value of the control parameter giving a minimum of the characteristic time t_c. In general the value of this parameter is dependent only on the dimensionality of the lattice. We have also shown that the requirement of a nonblocked system in the multispecies case leads to two somehow opposite conditions: either there has to be a strictly dominant species, outnumbering the others, or the diversity should be very large. In the end, we have also given evidence that the number of blocked species of different kinds decreases with the dimension of the system. The presented results can be easily further generalized to the cases of other topologies (e.g., complex networks) and other kinds of biases.

The authors acknowledge support from the European COST Action MP0801 Physics of Competition and Conflicts and from the Polish Ministry of Science Grant No. 578/N-COST/2009/0.

---

[1] C. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 2004).

[2] J. Sienkiewicz and J. A. Hołyst, Phys. Rev. E **80**, 036103 (2009).

[3] T. Antal, P. L. Krapivsky, and S. Redner, Phys. Rev. E **72**, 036121 (2005).

[4] F. Radicchi, D. Vilone, S. Yoon, and H. Meyer-Ortmanns, Phys. Rev. E **75**, 026106 (2007).

[5] J. von Neumann and A. W. Burks, *Theory of Self-Reproducing Automata* (University of Illinois Press, Urbana, 1966).