www.sciencemag.org/cgi/content/full/339/6124/1185/DC1

Supplementary Materials for

Avoiding a Spanning Cluster in Percolation Models

Y. S. Cho, S. Hwang, H. J. Herrmann, B. Kahng*

*To whom correspondence should be addressed. E-mail: bkahng@snu.ac.kr

Published 8 March 2013, *Science* **339**, 1185 (2013)
DOI: 10.1126/science.1230813

This PDF file includes:

Supplementary Text

Percolation transitions for $d < d_c$: Analytic solutions for t_{cm}

We consider the probability $q(\ell)$ that all randomly chosen m unoccupied bonds are bridge bonds at t, given by $\ell = tN_b$. This probability is obtained as

$$q(\ell) = \left[\frac{N_{BB}}{N_b(1-t)} \right]^m \sim N^{-m/m_c} \left[\frac{(t-t_c)^\zeta}{(1-t)} \right]^m , \tag{1}$$

where $q(\ell) = 0$ for $\ell \le t_c N_b$, and N_{BB} is the number of bridge bonds, given by

$$N_{BB} \sim \begin{cases} L^{1/\nu_0} & \text{at } t = t_c , \\ L^{d_{BB}} (t-t_c)^\zeta & \text{at } t > t_c . \end{cases} \tag{2}$$

The exponent ν_0 is the correlation length exponent for ordinary percolation transition (PT), d_{BB} is the fractal dimension of a pattern formed by bridge bonds (*20*), and the exponent ζ is an exponent describing the transition behavior of N_{BB}. As N_{BB} increases, bridge bonds form a chain or a surface in two and three dimensions, respectively, and thus the fractal dimension of bridge bonds is bounded as $d-1 < d_{BB} < d$. m_c is denoted as

$$m_c(d) = \frac{d}{d-d_{BB}} , \tag{3}$$

across which the order of the PT changes. Numerical values of $m_c(d)$ for various dimensions are listed in Table S1.

Next, we calculate the probability that the system reaches a percolating state at ℓ for the first time, which is given by

$$Q(\ell)=\left[\prod_{k=0}^{\ell-1}\left(1-q(k)\right)\right]q(\ell)\approx e^{-\sum_{k=0}^{\ell-1}q(k)}q(\ell). \tag{4}$$

We call ℓ_{cm} the value at which $Q(\ell)$ has its maximum. Then, the relation

$$q(t)=\frac{1}{N_b}\frac{\partial \ln q(t)}{\partial t} \tag{5}$$

holds at $t_{cm}=\ell_{cm}/N_b$. t_{cm} depends on the system size N. We find that $t_{cm}(N)$ behaves differently for the cases $m<m_c$ and $m>m_c$. When $1<m<m_c$, $t_{cm}(N)-t_c$ decreases as a power law $\sim N^{-1/\bar{\nu}_<}$, and when $m>m_c$, $1-t_{cm}(N)$ decreases as $\sim N^{-1/\bar{\nu}_>}$. The exponents are obtained by inserting Eq. (1) into Eq. (5) as

$$\frac{1}{\bar{\nu}_<}=\frac{1}{m\zeta+1}\left(1-\frac{m}{m_c}\right), \tag{6}$$

and

$$\frac{1}{\bar{\nu}_>}=\frac{1}{m-1}\left(\frac{m}{m_c}-1\right). \tag{7}$$

We remark that at m_c, $q(\ell)\sim N^{-1}\left[(t-t_c)^{\zeta}/(1-t)\right]^{m_c}$, which may be written as $q(\ell)=N^{-1}\tilde{q}(\ell)$, where $\tilde{q}(t)$ is independent of N. Then, $\tilde{q}(t)$ satisfies the equation $\tilde{q}(t)=\partial \ln \tilde{q}(t)/\partial t$ and its solution is t_{cm_c}. Thus, t_{cm_c} is finite in the thermodynamic limit, which is neither t_c nor unity. Moreover, $q(\ell)\sim O\left(1/N\right)$ at t_{cm_c}, and thus the probability to actually occupy a bridge bond is non-negligible.

Percolation transitions for $d\geq d_c$

When $d=d_{BB}$ for $d\geq d_c$, the probability $q(t)$ in Eq. (1) is independent of the system size N as

$$q(t) \sim \left[\frac{(t-t_c)^{\zeta}}{(1-t)} \right]^m \equiv f^m(t). \tag{8}$$

Then, for finite m, following a similar step as used to derive Eq. (6) from Eq. (5), we obtain that

$$t_{cm}(N) - t_c \sim N^{-1/(m\zeta+1)}. \tag{9}$$

Then a continuous PT occurs at t_c in the thermodynamic limit.

If m varies with N as $m \sim \ln N$, then $t_{cm} - t_c$ is independent of N. That is, t_{cm} is finite. Then, a discontinuous PT occurs at finite t_{cm}, which is neither t_c nor unity. Furthermore, when m is increased faster than $\ln N$ as N is increased, t_{cm} increases. In this case, we rewrite Eq. (5) as

$$\ln f = -\frac{\ln N}{m+1} + \frac{\ln f'}{m+1} + \text{const}, \tag{10}$$

at $t = t_{cm}$, where the prime denotes the first derivative with respect to t. For the case $\lim_{N\to\infty} \ln N/(m+1) \to 0$, we obtain a finite t_{cm} which satisfies the relation $(t_{cm} - t_c)^{\zeta} = C(1-t_{cm})$, where C is const. Actually, this t_{cm} is the percolation threshold of the bridge percolation transition (*20*).

Statistical fluctuations of the percolation threshold

Next, we consider the statistical fluctuations of the critical point t_{cm}. To quantify the fluctuations, we consider the percolating probability distribution function $Q(t)$ around the critical point t_{cm}. The standard deviation $\sigma(m,N)$ of t_{cm} is calculated using the relation

$$1/\sigma^2 \sim (\ln Q(t))'' \tag{11}$$

at t_{cm}, where the double prime denotes the second derivative with respect to t. Then, we obtain that

$$\sigma \sim \begin{cases} N^{-1/\bar{\nu}_<} & \text{for} \quad m < m_c, \\ N^{-1/\bar{\nu}_>} & \text{for} \quad m > m_c. \end{cases} \tag{12}$$

Thus, for $m < m_c$ ($m > m_c$), the fluctuations are relatively large near m_c and become smaller with decreasing m (increasing m) for a fixed system size. The standard deviations shrink with increasing N for a given m. However, when $m = m_c$, the standard deviation remains constant, independent of the system size. Thus, it generically appears that the explosive percolation threshold fluctuates heavily from sample to sample at the tricritical point m_c. The theoretical prediction is confirmed by simulation data, shown in Fig.S1. On the other hand, from Eq. (12), we can find that the exponent ν characterizes the scaling relation between length L and the occupation probability t.

Percolation transitions in the product rule model

Here we study the explosive PT for the product rule (PR) model under the best-of-*m* strategy in Euclidean space, in which the dynamics proceeds under the suppression against the formation of large cluster. At each time, *m* potential bonds are selected randomly, and a set of the products of sizes of the two clusters connected by each of these potential bonds are calculated. If a selected potential bond is an intra-bond, then the product is taken as the square of that cluster size. The potential bond producing the minimum of those products is actually occupied, and the other $m-1$ potential bonds are discarded. Previous studies focused on the case $m = 2$ in Euclidean space (*6, 13, 14*). Here, we want to see the behavior of the PT for general *m*, and compare it with the results for the SCA model.

First, in Fig. S3A, for a given system size, for example, $L = 200$ in two dimensions, we examine the behavior of the giant cluster size per node $G_m(t)$ as a function of time t for various m. The transition point $t_{cm}(L)$ is delayed as m is increased. Interestingly, the increase of $G_m(t)$ for relatively small m such as $m = 2-8$ is different from that for relatively large m such as $m \geq 20$: (i) For those small m, $G_m(t)$ increases sharply. However, the jump size decreases with increasing the system size (Fig. S3E). For this case, large clusters are not compact, but contain isolated small clusters inside at the onset of the abrupt change of the order parameter, denoted as t_{cm}^- (Fig. S4A). The cluster size distribution at t_{cm}^-, decays as a power-law in the small-cluster-size region, but shows a bump in the large-cluster-size region (Fig. S4B). (ii) For those large m, $G_m(t)$ increases irregularly, caused by the averaging over plateaus stemming from different configurations (Fig. S3B). Plateaus appear when intra-cluster bonds are added to the system. During the transient period, small clusters merge with surrounding large clusters, and this merging dynamics leads to the formation of compact clusters at t_{cm}^- as shown in Fig. S4C. The interface between two compact clusters is self-affine with the same fractal dimension as that of bridge bonds in the SCA model (Fig. S5). Cluster sizes of such compact clusters are almost monodisperse as shown in the cluster size distribution in Fig.S4D, because smaller (larger) clusters are more (less) likely to merge to others. We find that the

jump size of the order parameter remains finite as the system size is increased (Fig. S3F).

The number of distinct clusters $N_{cl}(t_{cm}^-)$ at t_{cm}^- exhibits distinct features. (i) When m is small such as $m=3$, the extensive relationship $N_{cl}(t_{cm}^-) \sim N$ holds, whereas (ii) when m is large, a crossover behavior occurs: For small N, the sub-extensive or the system-size independent relationship $N_{cl}(t_{cm}^-) \sim N^{\alpha}$ with $\alpha \ll 1$ or $\alpha \approx 0$ appears, whereas for large N, the extensive relationship $N_{cl}(t_{cm}^-) \sim N$ holds. This behavior in 2-dim is shown in Fig. S6A. The crossover point, denoted as N^*, scales as $N^* \sim e^{am}$, where a is const. When N is fixed, the crossover occurs at $m^* \sim \ln N$. Thus, as $N \to \infty$ we have $m^* \to \infty$. The phase diagram is shown in Fig. S6B. This result is intrinsic, independent on dimension. We argue that when the number of clusters linearly depends on (is sub-extensive of) system size, the jump size of the order parameter decreases (finite). Thus, for the PR model, the PT is continuous in the limit $N \to \infty$ for a finite fixed m even in low dimensional systems. However, if m varies with system size with the constraint $m \geq \ln N$, then the percolation transition could be discontinuous. This result is the same as the mean-field result above the upper critical dimension $d_u = 6$ for the SCA model. The crossover behavior is also observed on random graphs (Fig. S6C and D).

Here, we present a heuristic argument supporting the above numerical results. We consider the case (i) with relatively small m. We show snapshots of the system at several time steps to display how the system evolves, finding that small clusters still remain above the percolation threshold but large clusters merge, eventually making the largest cluster (Fig. S7). As m is increased, the bump size in the large-cluster-size region is increased, whereas the small-cluster-size region is reduced (Fig. S8). We denote the characteristic cluster size as s^*, which divides those two regions. We calculate the probability p_s (p_l) that a cluster of size $s < s^* \left(s > s^* \right)$ is selected randomly as

$$p_s(t) = \frac{\int_1^{s^*} ds\, n_s s f_s}{(1-t)} \quad \text{and} \quad p_l(t) = \frac{\int_{s^*}^{\infty} ds\, n_s s f_l}{(1-t)}, \tag{13}$$

where $n_s(t)$ is the cluster size distribution at t, f_s and f_l are the mean densities of unoccupied bonds in small and large clusters, respectively and $(1-t)$ is the density of unoccupied bonds in the system. The relation $p_s + p_l = 1$ holds. Then, the probability that all m potential bonds are selected in the bump region is given by $P_{l,m}(t) \equiv p_l^m(t)$. If the condition $P_{l,m} < N^{-1}$ is satisfied, then the probability that an unoccupied bond is actually added to a cluster of size $s > s^*$ is negligible. Thus, the bonds are actually occupied in small clusters, and as a result, clusters become compact and $N_{cl}(t_{cm}^-) \sim N^{\alpha}$ with $\alpha < 1$ at t_{cm}^-. This condition may be written as

$m > m_c \equiv \ln N / |\ln p_l|$ unless $p_l = 0$. For $m < m_c$, small clusters co-exist, and then $N_{cl}(t_{cm}^-) \sim N$ at t_{cm}^-. We note that such different cluster shapes for $m < m_c$ and $m > m_c$ can be seen in the SCA model in Fig. S9.

When $m > m_c$, clusters are compact, and they merge as time goes on. We consider the cluster merging dynamics (Fig. S9) between times t_{cm}^- and t_{cm}^+, at which a single macroscopic-scale cluster remains. Because the cluster size distribution has a narrow bell shape (Fig.S4D), we assume for simplicity that all cluster sizes are equal. Then, the linear size of each cluster is $L_{cl}(t) \sim \left(N / N_{cl}(t)\right)^{1/d}$ and the total length of the interfaces scales as $M_{BB}(t) \sim N_{cl}(t) L_{cl}^{d_{BB}}(t)$. Cluster merging process takes place when an inter-cluster bond is actually occupied, which occurs with probability,

$$\begin{aligned} Q^{(\mathrm{PR})}(t) &\sim \sum_{k=1} \frac{k}{m} \binom{m}{k} \left(\frac{M_{BB}}{N_b(1-t)} \right)^k \left(1 - \frac{M_{BB}}{N_b(1-t)} \right)^{m-k} \\ &\sim \frac{1}{1-t} \left(\frac{N_{cl}(t)}{N} \right)^{1-\frac{d_{BB}}{d}}. \end{aligned} \tag{14}$$

After two clusters are merged, we assumed that remaining clusters are reorganized and their sizes are again monodisperse with a larger size but their number is reduced. This assumption is viewed in ensemble average perspective. The process is repeated. The decreasing rate of the number of clusters is given by

$$\frac{1}{N_b} \frac{dN_{cl}}{dt} = -Q^{(\mathrm{PR})}(t). \tag{15}$$

Because the abrupt phase transition occurs in a short time interval, we may make the approximation $1/(1-t) \approx 1/(1-t_{cm}^-)$ for $t_{cm}^- < t < t_{cm}^+$. Using Eqs. (14) and (15), and the boundary conditions $N_{cl}(t_{cm}^-) \sim N^{\alpha}$ and $N_{cl}(t_{cm}^+) = 1$, we obtain that

$$t_{cm}^+ - t_{cm}^- \sim N^{-(1-\alpha)d_{BB}/d}. \tag{16}$$

We have confirmed the exponent numerically in the inset of Fig.S3F. Since the exponent is $\frac{(1-\alpha)d_{BB}}{d} > 0$ for $\alpha < 1$, $t_{cm}^+ - t_{cm}^-$ is reduced to zero as $N \to \infty$, whereas the jump size of the order parameter remains as $O(1)$. Thus, the transition becomes sharper as the system size is increased. However, when the system size exceeds N^*, $\alpha = 1$. Thus, the transition becomes continuous. Therefore, the criterion whether

$\alpha < 1$ or $\alpha = 1$ in the thermodynamic limit is a good indicator to determine the type of PT.

It is noteworthy that the PT for the PR model is reduced to the mean-field result for the SCA model, which is due to the following reason. In the SCA model, the cluster merging dynamics is mainly determined by bridge bonds, of which the number is smaller than $O(N)$ below the upper critical dimension, but is of order of $O(N)$ in the mean-field limit. In the PR model, however, the cluster merging dynamics is determined not only by the selection of unoccupied bonds on the interface, but also of intra-cluster bonds, and their total number is $O(N)$. Thus, the PT for the PR model is reduced to the mean-field behavior of the SCA model even in low dimensions.

Percolation transitions in the Gaussian model

The Gaussian model (*18*) was introduced to search a discontinuous percolation transition, which is defined as follows: At each time step, an empty bond is selected randomly, and is occupied with probability,

$$p \sim \exp\left[-w\left(\frac{s - \bar{s}}{\bar{s}} \right)^2 \right], \tag{17}$$

where s is the cluster size created by occupying the bond and $\bar{s}$ the average cluster size in the system that would be formed after an empty bond is occupied. w is a control parameter defined in the region $w > 0$. When $w = 0$, the model is reduced to the ordinary percolation model. When $w = 1$, this model displays compact clusters at the onset of the PT, and the transition is discontinuous, independent on spatial dimension (*18*).

Here, we examine the number of clusters $N_{cl}(t_{cm}^-)$ as a function of the system size N at the onset of the PT. It is expected that for a fixed w, $N_{cl}(t_{cm}^-)$ increases linearly with N for small system sizes, and sub-extensive for large system sizes. Indeed, we find that such crossover behavior appear in two dimensions (Fig. S11). Thus, for finite w, the percolation transition is discontinuous in the thermodynamic limit for the Gaussian model.

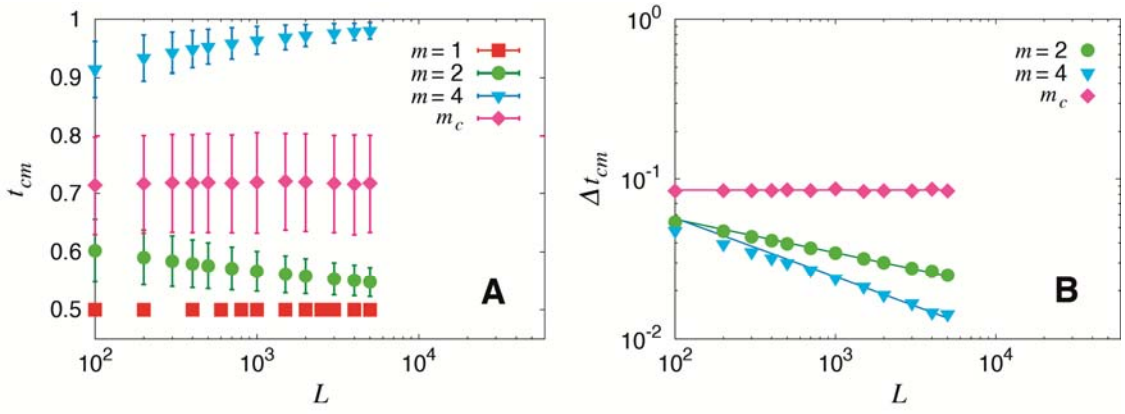

Figure S1.

A. Plot of the percolation thresholds with variance versus system size L in two dimensions for the SCA model. **B.** Plot of the standard deviations of t_{cm} versus system size L in two dimensions. They decay in a power-law manner for $m \neq m_c$, and are flat for $m = m_c$. Solid lines are guidelines with theoretically obtained slopes. All data are obtained after averaging over 10^4 configurations.

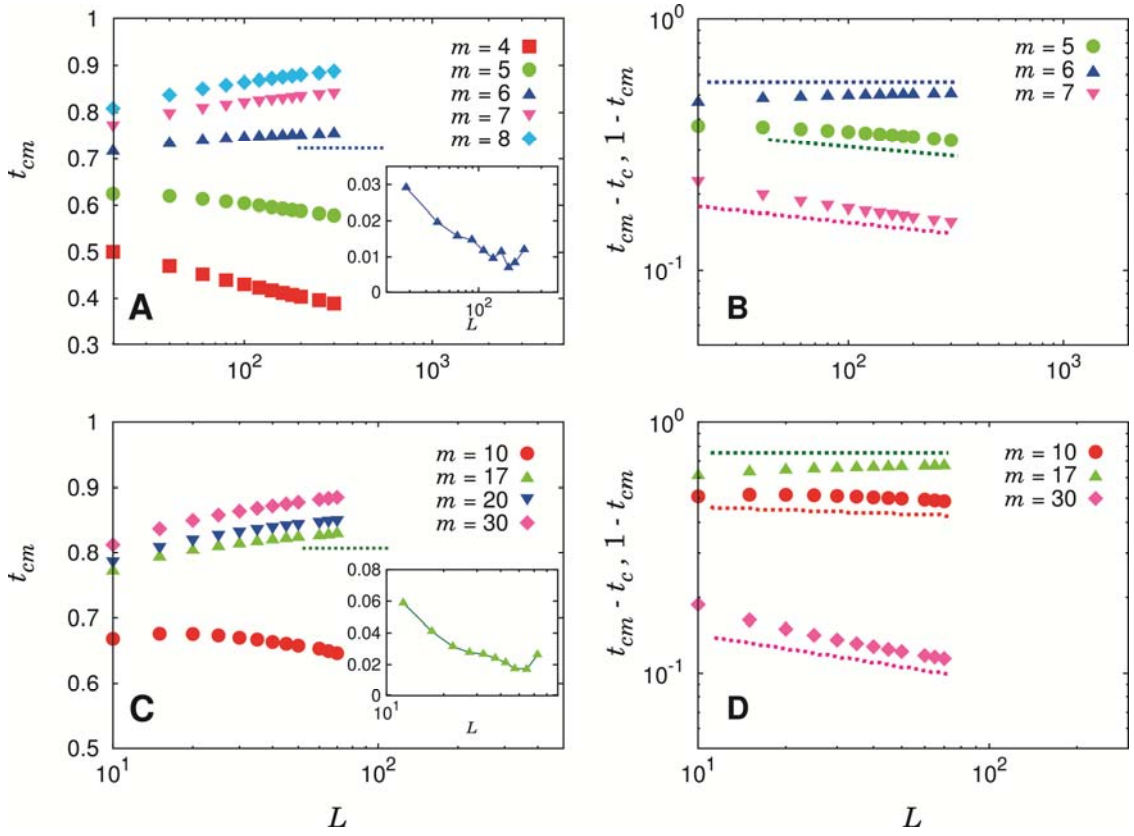

Figure S2.

A. and **C**. Plot of the percolation threshold t_{cm} versus system size L for various numbers of potential bonds m in three (**A**) and four (**C**) dimensions for the SCA model. Dotted lines are straight lines. Inset: Successive slopes of t_{cm} vs. L for $m_c \approx 6$ (**A**) and $m_c \approx 17$ (**C**). They tend to decrease to zero, which implies that t_{cm} converges to a finite value. We also checked that the successive slopes of $1-t_{cm}$ and $t_{cm}-t_c$ decrease to zero at m_c. This indicates that t_{cm} is finite, neither t_c nor 1. **B**. and **D**. Plot of scaling relations, $t_{cm}-t_c$ vs. L for $m < m_c$ and $1-t_{cm}$ vs. L for $m > m_c$ in three (**B**) and four (**D**) dimensions. Dotted lines are guidelines with theoretically obtained slopes. All data are obtained after averaging over 10^4 configurations.

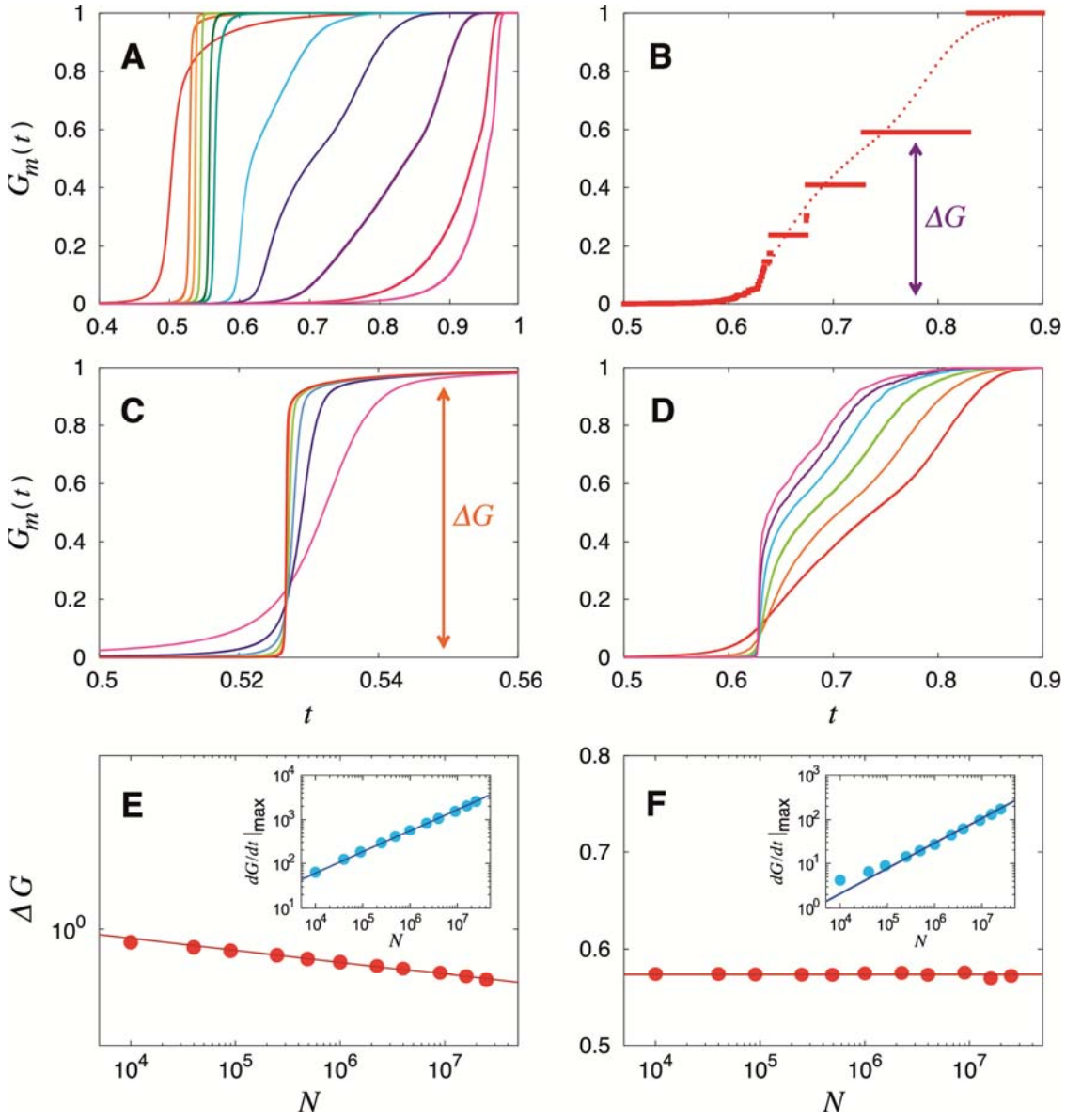

Figure S3.

A. Plot of the giant cluster size $G_m(t)$ vs. time t after ensemble average for various m for the PR model in two dimensions. Data are obtained for $m = 1$, 2, 3, 5, 8, 10, 20, 30, 50, 100, and 300 from the left to the right with a fixed size $L = 300$. **B.** Similar plot of $G_m(t)$ for $m = 30$ and $L = 200$ after ensemble average (dotted line) and for a single configuration (solid lines). **C.** and **D.** Plot of the giant cluster $G_m(t)$ vs. time t for fixed $m = 2$ (**C**) and $m = 30$ (**D**), but various system sizes $L = 100$, 300, 700, 1500, 3000, and 5000 from the right to the left. As the system size increases, a giant cluster grows more drastically. **E**. Plot of the jump size vs. system size N for $m = 2$. For $m < m_c$, jump size

is measured as the height of $G_m(t)$ just after the abrupt growth. Solid line is a guideline with slope -0.02. Inset: Plot of the maximum value of $dG_m(t)/dt$ vs. system size N. Solid line is a guideline with slope 0.48. **F.** Plot of the jump size vs. system size N for $m = 30$. In this case ($m > m_c$), jump size is measured as the height when a large-scale plateau takes place. The estimated jump size is independent of N. Inset: Plot of the maximum value of $dG_m(t)/dt$ vs. system size N. Solid line has slope 0.57, consistent with the theoretical value $(1-\alpha)d_{BB}/d$. Data for $G_m(t)$ and ΔG are obtained after averaging over 10^4 configurations.

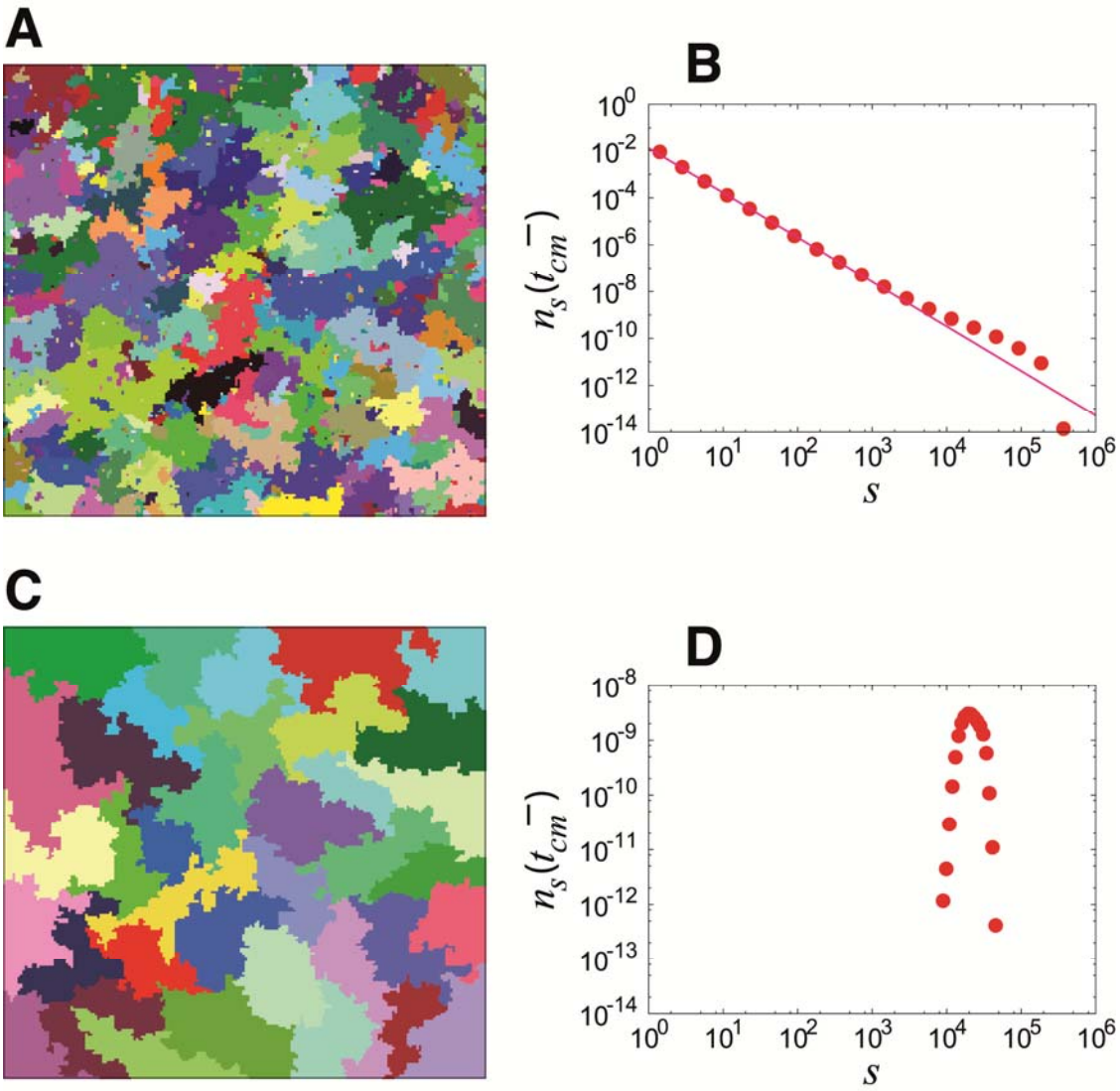

Figure S4.

Snapshots of clusters for the PR model at time t_{cm}^- at which the order parameter begins to increase drastically. The data are obtained from two dimensional systems with linear size $L = 200$, but for different m values. For small $m = 2$ (**A**), clusters contain small-sized clusters within them. For large $m = 30$ (**C**), clusters are compact. Plot of the cluster size distributions for $m = 2$ (**B**) and $m = 30$ (**D**) at t_{cm}^-. Data in **B** and **D** are obtained after averaging over 10^3 configurations in two dimensions with linear size $L = 10^3$. Solid line in (**B**) has slope -1.95.

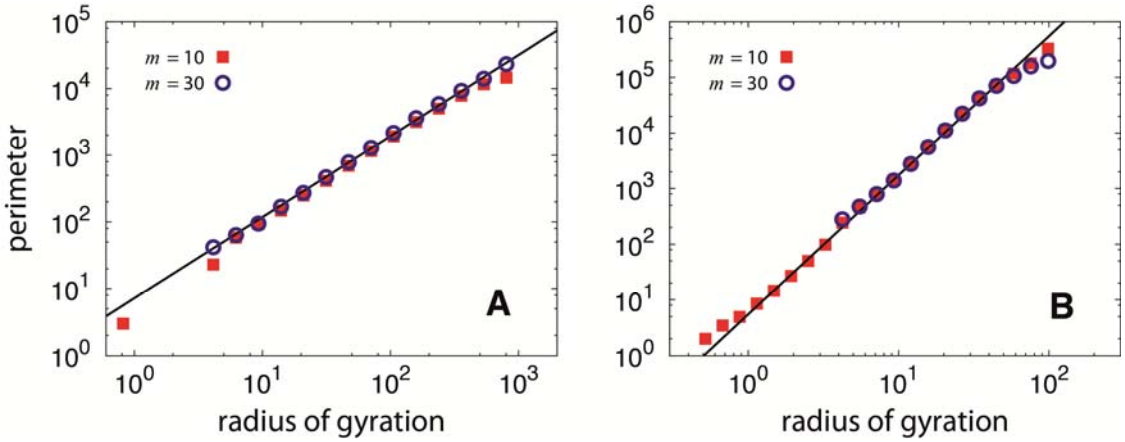

Figure S5.
Plot of the number of unoccupied bonds which inter-connect one cluster and its neighbor if occupied (perimeter) vs. radius of gyration of each cluster for $m = 10$ and 30 in two (**A**) and three (**B**) dimensions at t_{cm}^{-} for the PR model. Thus, the slope means the fractal dimension of perimeter of clusters. Solid lines are guidelines with slopes 1.22 (**A**) and 2.5 (**B**), respectively. Simulation results are obtained after averaging over 10^3 configurations.

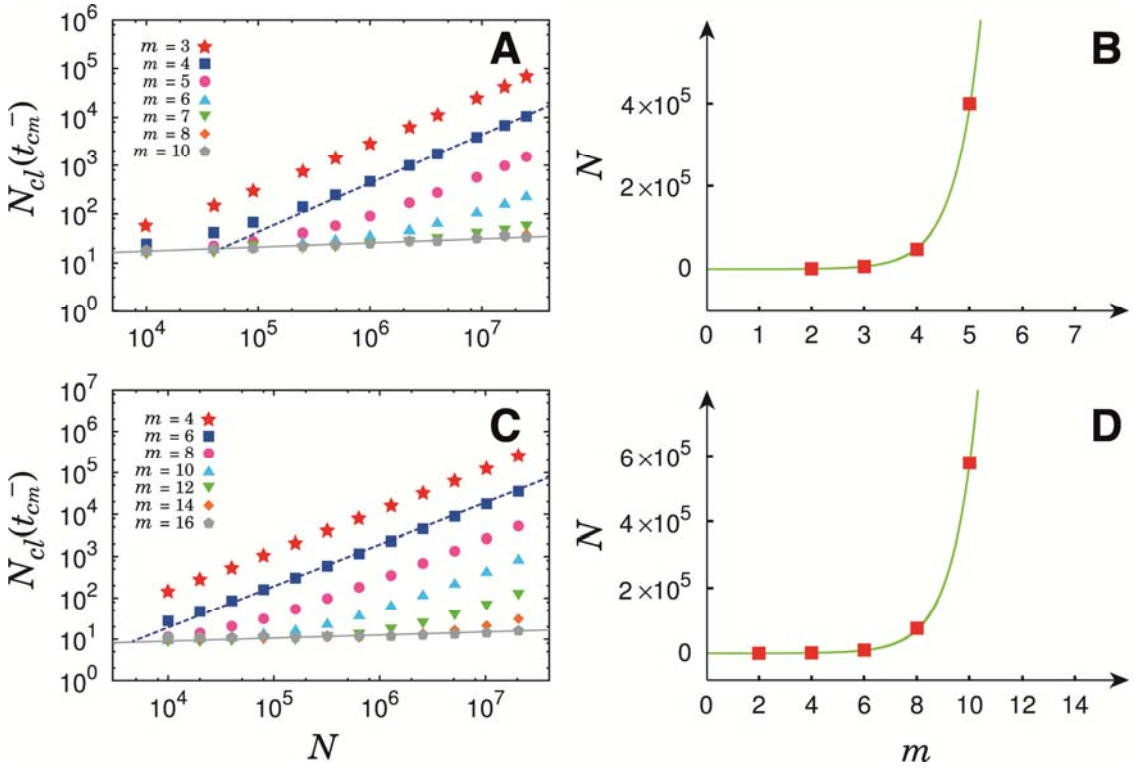

Figure S6. A. Plot of the number of clusters at $t = t_{cm}^-$ vs. system size N for several values of *m* for the PR model in two dimensions (**A**) and random graphs (**C**). Dashed and solid lines are guidelines with slopes one and 0.05, respectively. A crossover behavior occurs between the behaviors $N_{cl} \sim N^{\alpha}$ $(\alpha \approx 0.05)$ for $N < N^*$ and $N_{cl} \sim N$ for $N > N^*$, at the point $N^* \sim e^{am}$ ($a = \text{const.}$), which are represented by the solid curves in **B** and **D**. Red squares in **B** and **D** are estimated values of N^* from the numerical data in **A** and **C**. All data are obtained after averaging over 10^4 configurations.

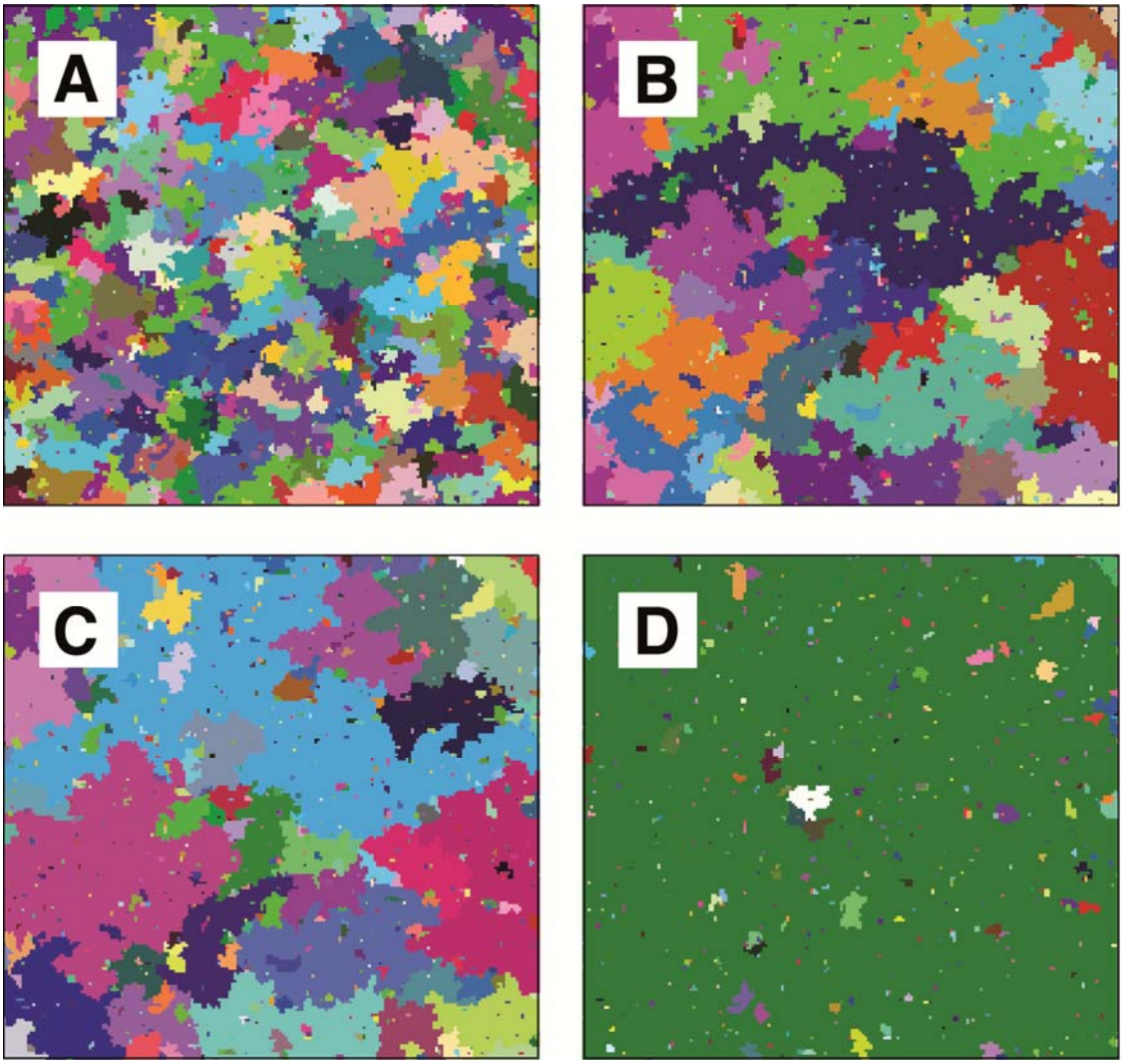

Figure S7. Snapshots of clusters in the system with $L = 200$ at several time steps for $m = 2 < m_c$ for the PR model at t_x in ($t_c < t_x < t_{cm}^-$) (**A**), t_{cm}^- (**B**), t_{cm} (**C**), and t_{cm}^+ (**D**). Even after the giant cluster is formed, small-sized clusters still remain in the system.

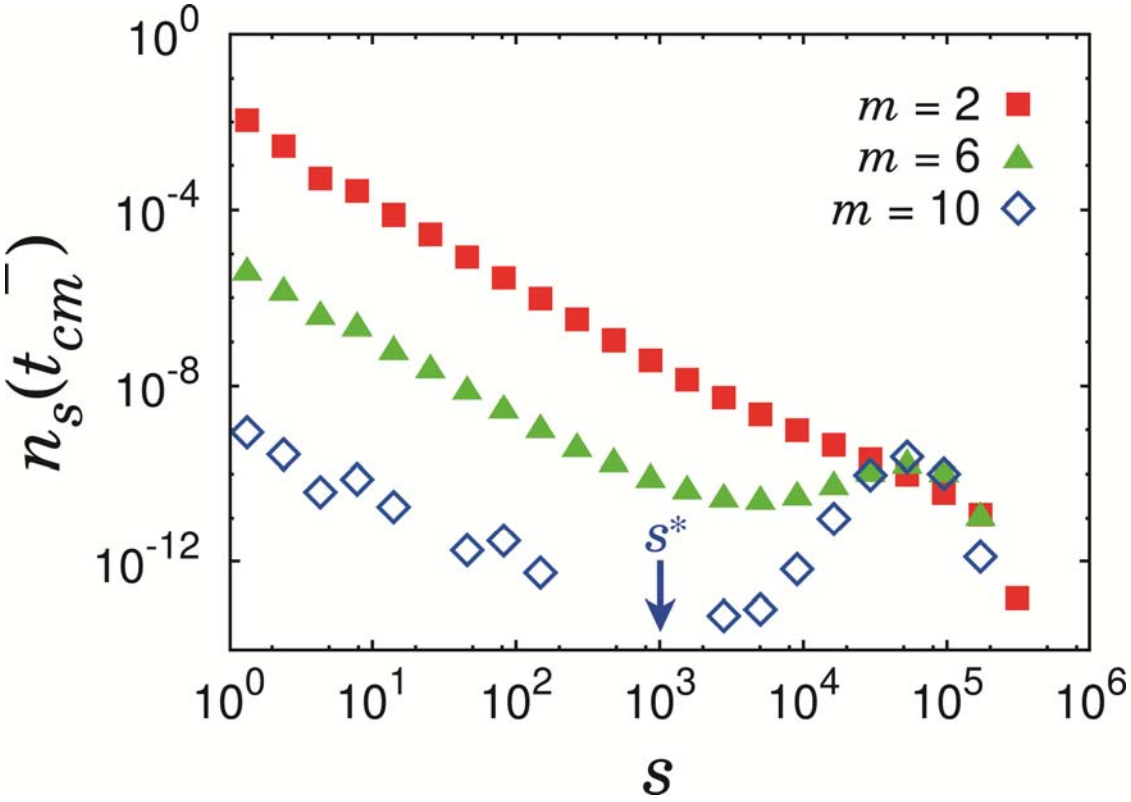

Figure S8. Plot of the density of *s*-sized clusters at t_{cm}^- vs. cluster size *s* in 2-dim with system size $L = 10^3$ for several *m*. As *m* is increased, bump size is increased, and the small-cluster-size region is reduced. Here s^* is the characteristic size beyond which the bump is distinct. All data points are obtained after averaging over 10^4 configurations.

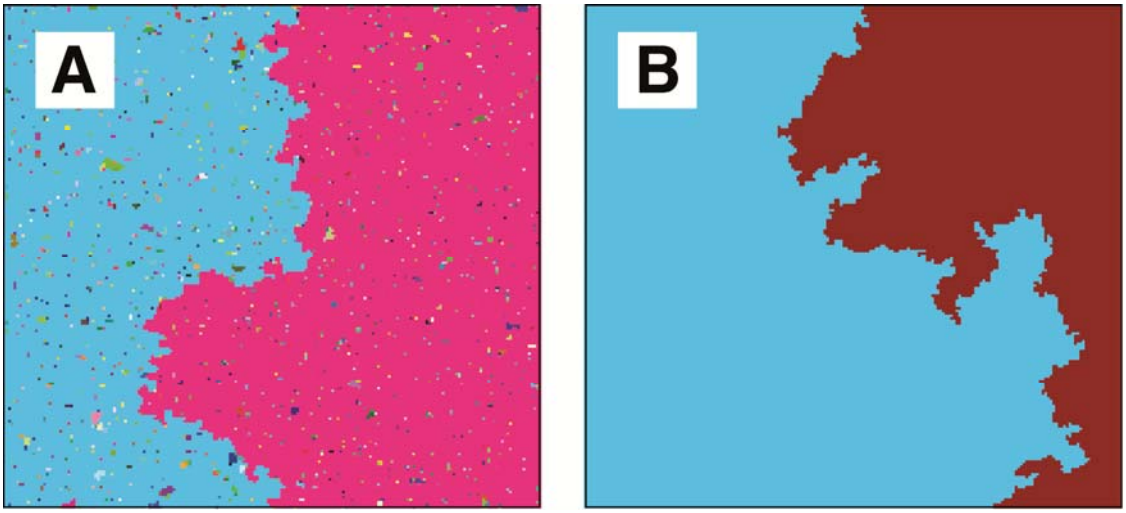

Figure S9. Snapshots of clusters for the SCA model when $m = 2$ (**A**) and 5 (**B**) in the system with $L = 200$ at t_{cm}^{-}. Two large clusters are not compact but include small clusters in **A**, but compact in **B**.

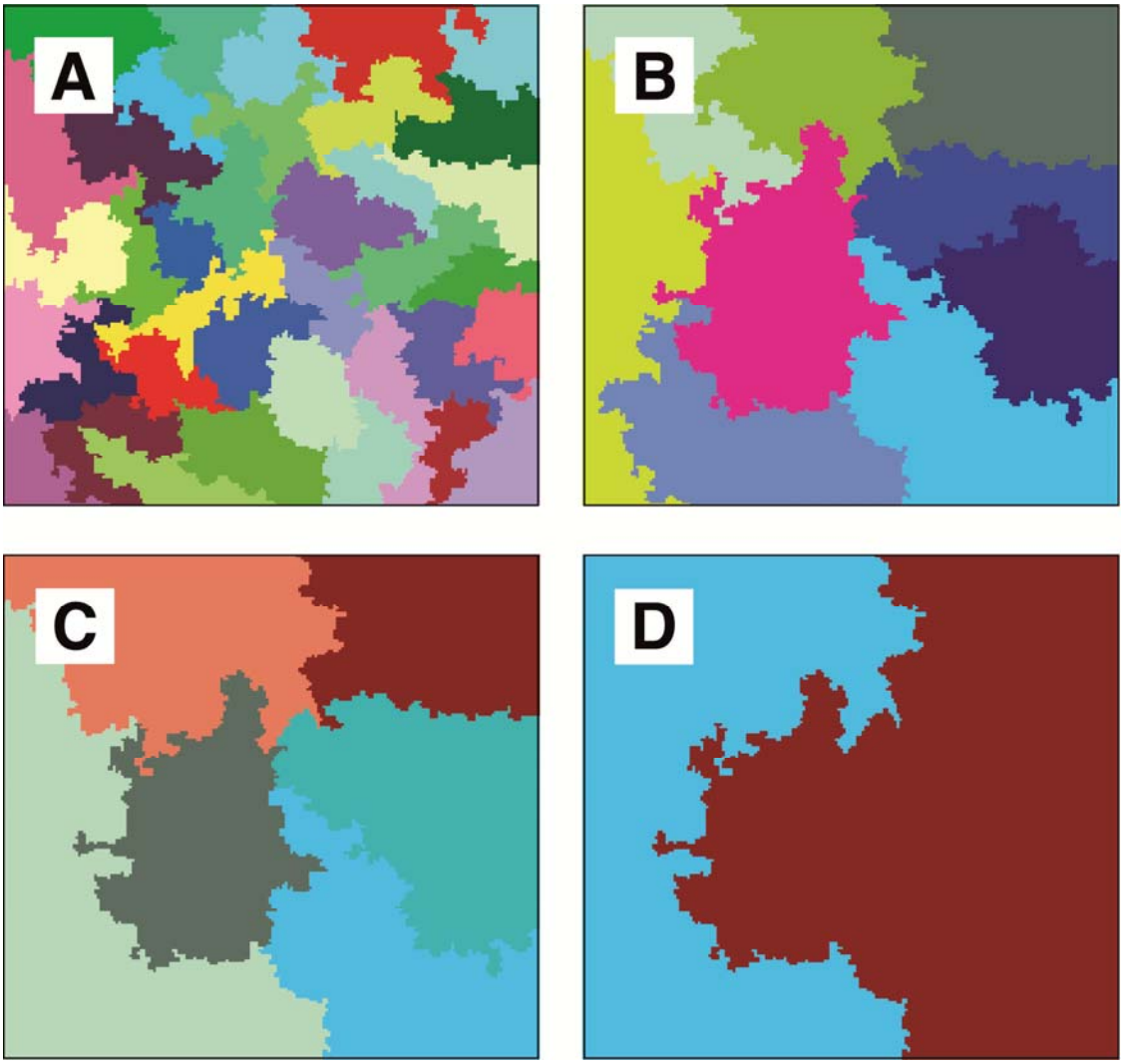

Figure S10. Snapshots of clusters for $m = 30 > m_c$ in a system with $L = 200$ for the PR model at several time steps when jumps arise as shown in Fig. S3B.

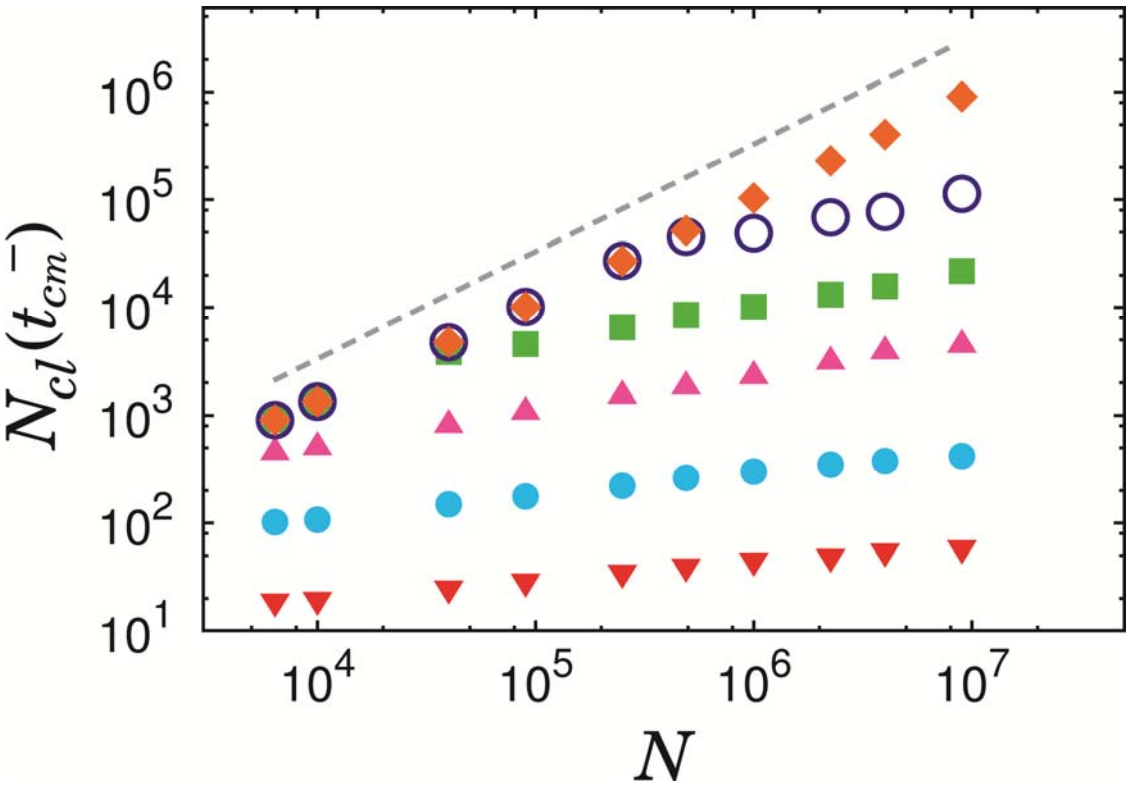

Figure S11. Plot of the number of clusters N_{cl} at t_{cm}^- as a function of system size N for several w values for the Gaussian model in two dimensions. Data are for $w = 0, 10^{-8}, 10^{-6}, 10^{-4}, 10^{-2}$, and 1 from the top. Dashed line is a guideline with slope 1. For $w = 10^{-8}$ (open circle), a crossover can be seen from the behavior $N_{cl} \sim N$ in small-N region to the one $N_{cl} \sim N^{\alpha}$ with $\alpha < 1$ in large-N region. All data are obtained after averaging over 10^3 configurations.

Table S1. List of numerical values of $m_c(d)$ for dimensions $d = 2-6$. The values for ζ and d_{BB} were measured in (*16*). * The standard deviation becomes larger than the mean value m_c.

d	ζ	d_{BB}	m_c
2	0.50 ± 0.03	1.215 ± 0.002	2.55 ± 0.01
3	1.0 ± 0.1	2.498 ± 0.005	5.98 ± 0.07
4	1.3 ± 0.5	3.74 ± 0.08	16.99 ± 5.23
5	1.4 ± 0.6	4.9 ± 0.2	50^*
6	1.5 ± 0.7	6.0 ± 0.1	∞