Evolution of autocatalytic sets in a competitive percolation model

Renquan Zhang\(^1\), Sen Pei\(^2\), Wei Wei\(^2\) and Zhiming Zheng\(^2\)

\(^1\) School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, People’s Republic of China
\(^2\) LMIB and School of Mathematics and Systems Science, Beihang University, Beijing 100191, People’s Republic of China
E-mail: peisen@buaa.edu.cn and weiw@buaa.edu.cn

Received 12 August 2014
Accepted for publication 9 October 2014
Published 10 November 2014

Online at stacks.iop.org/JSTAT/2014/P11018
doi:10.1088/1742-5468/2014/11/P11018

Abstract. The evolution of autocatalytic sets (ACSs) is a widespread process in biological, chemical, and ecological systems and is of great significance in many applications such as the evolution of new species or complex chemical organization. In this paper, we propose a competitive model with an \(m\)-selection rule in which an abrupt emergence of a macroscopic independent ACS is observed. By numerical simulations, we find that the maximal increase of the size grows linearly with the system size. We analytically derive the threshold \(t_\alpha\) where the explosive transition occurs and verify it by simulations. Moreover, our analysis explains how this giant independent ACS grows and reveals that, as the selection rule becomes stricter, the phase transition is dramatically postponed, and the number of the largest independent ACSs coexisting in the system increases accordingly. Our result indicates that suppression during evolution could lead to the abrupt appearance of giant ACSs.

Keywords: network dynamics, random graphs, networks, nonlinear dynamics
1. Introduction

Complex networks are often used to describe a variety of chemical, biological, and social systems [1–5]. For instance, the metabolism of a cell is a network of substrates and enzymes interacting via chemical reactions [6, 7]. Ecosystems are networks of biological organisms with predator-prey, competitive, or symbiotic interactions [8, 9]. In real-world systems, these networks are by no means static. On the contrary, biological or chemical networks often evolve into certain structures to adapt to their environment. It has been shown that the small-world structure of metabolic networks may have evolved to enable a cell to react rapidly to perturbations [6]. Similarly, the visual cortex may have evolved into a small-world architecture since it would aid the synchronization of neuron firing patterns [10]. Therefore, understanding the mechanisms responsible for the evolution of such networks is an important issue.

To explore the mechanisms underlying the network evolution, a set of models based on artificial chemistry of catalyzed reactions are proposed [11–14]. An artificial chemistry is a system whose components react with each other in a way analogous to molecules participating in chemical reactions. This kind of system is widespread in biological research, including works on the protein or enzymes within a cell [6, 7] and some organic molecules in a pool on the prebiotic Earth [15]. The networks evolve over time as mutations occur [16–18], and the structure of networks in turn affects subsequent evolutions. With these models, questions about self-organization, the origin of life, and other evolvability issues are explored in works on artificial chemistries [11–14].

Based on this framework, Bak and Sneppen [16] introduced a simple and robust model of biological evolution of an ecology of interacting species with the feature that the least fit species mutated. The model self-organized into a critical, steady state with intermittent coevolutionary avalanches of all sizes [16]. Later on, inspired by the Bak–Sneppen model,
Jain and Krishna [17, 18] proposed a similar model in which the mutation of a species also changed its links to other species. They investigated how the network of interactions among the species evolved over a longer time scale and the growth of the autocatalytic set (ACS) [17]. It was shown that, starting from a sparse random graph, an ACS inevitably appeared and triggered a cascade of exponentially increasing connectivity until it spanned the entire graph.

The concept of an ACS is introduced in the context of a set of catalytically interacting molecules. It is defined to be a set of molecular species that contains, within itself, a catalyst for each of its member species [19–22]. Mathematically, in a graph of interacting agents, an ACS is defined as a subgraph whose every node has at least one incoming link from a node that belongs to the same subgraph [17]. This definition is meant to capture the property that an ACS has ‘catalytic closure’ [11], i.e. it contains the catalysts for all its members. Therefore, autocatalytic sets might be more stable to perturbations because of their ability to self-replicate. On the prebiotic Earth, autocatalytic sets have been suggested as one of the possible means by which a complex chemical organization could have evolved [23]. Due to their property of self-replication, ACSs play an important role in the overall dynamics in chemical or biological networks. As defined, an ACS may have several disconnected components. These components are independent units that possess the property of self-replication. We define each component as an independent ACS and focus on the largest one in this paper, which contains the largest number of species.

The definition of the largest independent ACS is somewhat analogous to that of the giant component (GC) in percolation transition [24]. Percolation can be interpreted as the formation of a giant component in networks. One important model that shows this process is the classic Erdős and Rényi (ER) [25] model. In the ER model, the evolution proceeds as follows: Beginning with \( N \) isolated nodes, an edge is connected between a randomly selected unconnected pair of nodes at each time step. Then, as the number of connected edges increases, a macroscopic cluster, i.e. the giant component, appears at the percolation threshold, and its size grows continuously. Recently, based on the ER model, an explosive percolation (EP) [26] model was introduced. In this model, the ER model was modified by additionally imposing a so-called product rule or sum rule, which suppresses the formation of a large cluster [26]. Because of this suppressive bias, the percolation threshold is delayed. When the giant component eventually emerges, it does so explosively. Initially, this explosive phase transition was regarded as a discontinuous transition. However, it was recently found that the transition is continuous in the thermodynamic limit [27], followed by a mathematical proof [28] and extensive supporting simulations [29–31]. Even though this transition is continuous, the explosive percolation in this model is clearly different from the phase transition in the classical ER model. More importantly, several models, including the Gaussian model [32], the avoiding a spanning cluster model [33], and the BFW model [34, 35], are known to show discontinuous transitions. These results have attracted much interest [27–37].

Inspired by the EP model, we propose an evolving network model that presents an abrupt emergence of the largest independent ACS. By imposing a selection rule [33–35], the formation of the largest component of ACS is suppressed. In the next section, we introduce the definition of this competitive model and then discuss its properties. After, we discuss the analysis of the threshold where a macroscopic independent ACS appears.
Evolution of autocatalytic sets in a competitive percolation model

Figure 1. Schematic illustrations of network evolution. (a) This shows the structure of the graph at some time step. There are three independent ACSs (marked in red) of size 2, 3 and 4 respectively. The pink nodes represent the nodes that were randomly selected from the least fitness. (b) Update process for node 16. Remove the link \(\{16 \rightarrow 10\}\) and replace it by two randomly selected links. (c) Update process for node 3. After this update, node 3 will become a member of an independent ACS. According to the selection rule, since the update of node 16 does not increase the size of maximal independent ACS, node 16 should be chosen to be updated.

The theoretical results are verified by simulations. In the last section, we conclude our findings and provide a brief discussion.

2. Model

The system is described by a directed graph in which the \(N\) nodes represent the species or chemicals and the directed links stand for the catalytic interactions between them. The graph can be completely described by an adjacency matrix \(C = \{c_{ij}\}_{N \times N}\), where \(c_{ij} = 1\) if there exists a link from node \(j\) to \(i\), and zero otherwise. A directed link from node \(j\) to \(i\) means that \(i\) is catalyzed by \(j\). Specifically, we exclude self-replicating species, i.e. \(c_{ii} = 0\) for all \(i = 1, 2, \ldots, N\).

According to the definition, an ACS may consist of several disjointed, smaller ACSs, which have no intersections with each other and form as independent catalytic systems. To describe this, we introduce the concept of an independent ACS. Concretely, for the nodes in ACSs, an independent ACS is the maximal, weakly-connected subgraph that is composed only of ACS-nodes. For example, in figure 1(a), there are three independent ACSs marked in red. Notice that node 11 and 16 do not belong to the independent ACS since they are not ACS-nodes, even though they have links connecting to independent ACSs. Denote the size of independent ACSs as \(S_1, S_2, \ldots\) in descending order. In particular, since there may be several independent ACSs with the same size \(S_k\), we denote them as \(S_{k,1}, S_{k,2}, \ldots, S_{k,n_k}\). Here, \(n_k\) represents the number of independent ACSs with size \(S_k\).

For convenience, we simply denote \(S'_k\) as the set of independent ACSs with size \(S_k\). Based on the definition of ACS, \(S'_k\) is still an ACS with \(n_kS_k\) nodes. In this paper, we focus on the largest independent ACSs \((S'_1)\), which generally play the most important role of self-replicating in the evolution process.

doi:10.1088/1742-5468/2014/11/P11018
At the beginning time $t = 0$, the initial graph is a random graph with average indegree (or out-degree) $d$ (i.e. the probability of linking a directed edge is $p = d/N$). For every discrete time step $t = 1, 2, \cdots$, we denote $n_k(t)$ as the number of the independent ACSs in $S_k'$ at time $t$. The graph is updated as follows:

**Step 1.** Select $m$ ‘most mutating’ nodes in the network. This procedure is performed based on a special set named least fitness [17], which is explained later.

**Step 2.** For each of the selected nodes $i$, determine the size of the largest independent ACS in the new network if node $i$ is updated. The update associated with node $i$ is processed as follows: for node $i$, remove all the incoming and outgoing links attached to it. Then, among the $2(N-1)$ directed node pairs between the selected node $i$ and the rest of the nodes, we assign links with probability $p$. On average, $2(N-1)p$ directed links would be added to the network. Therefore, for each selected node $i$, we get a new adjacency matrix $C(i)$ and a new largest size of the independent ACS $S_1(i)$.

**Step 3.** Randomly select a node $i$ with $S_1(i) \leq S_1$, where $S_1$ is the size of the largest independent ACS before the update. If all the $S_1(i)$ are larger than the original $S_1$, we simply choose one node $i$ uniformly from the $m$ selected nodes. Then, the network is updated to $C(i)$.

At each time step, we need to select $m$ nodes with the ‘least fitness’. To achieve this, each node $i$ is assigned a population $y_i \geq 0$ and a relative population $x_i = y_i/Y$, where $Y = \sum_i y_i$. We assume that the total population $Y$ is fixed during the evolution and that there is no competition among different species. Between two successive graph updates, the evolution of population is given by [17]

$$\frac{dy_i}{dt} = \sum_{j=1}^{N} c_{ij} y_j - \phi y_i,$$

where $\phi > 0$ is a constant. In fact, the $x_i$ dynamics depends only on $C$ and not on $\phi$. Replacing $y_i$ with $x_iY$ in equation (1), we obtain

$$\frac{dx_i}{dt} = \sum_{j=1}^{N} c_{ij} x_j - \phi x_i.$$  \hspace{1cm} (2)

Since the relative population is normalized, $\sum_{i=1}^{N} x_k = 1$. We can calculate the value of $\phi$ through

$$0 = \frac{d}{dt} \sum_{k=1}^{N} x_k = \sum_{k=1}^{N} \frac{dx_k}{dt} = \sum_{k,j=1}^{N} c_{kj} x_j - \phi.$$  \hspace{1cm} (3)

From the above equation, we have $\phi = \sum_{k,j=1}^{N} c_{kj} x_j$. Therefore, $x_i$ has the dynamics

$$\frac{dx_i}{dt} = \sum_{j=1}^{N} c_{ij} x_j - x_i \sum_{k,j=1}^{N} c_{kj} x_j.$$  \hspace{1cm} (4)

For the initial condition of $x$, we first generate a random vector where $x_i$ is uniformly distributed in $[0, 1]$ and then normalize this vector. In this model, we can use $x_i$ to measure the fitness of the species $i$ in the environment defined by the graph. The larger $x_i$ is, the more fit species $i$ is. Such an idealized assumption has been justified in previous
studies [16, 17]. In a molecular context, the species with the least population is the most likely to be lost in fluctuation in a hostile environment. In addition, in an ecological context, low fitness might be correlated with a smaller barrier to mutation [16]. The elimination of the worst fit is an extreme idealized case of this correlation. In our research, we also adopt this assumption. Thus, the \( m \) nodes with smallest values of \( x_i \) are picked as the ‘most mutating’ nodes. Moreover, when \( x_i \) reaches a stable solution, the value of \( x_i \) is zero for almost all the nodes outside the ACS. Therefore, before an ACS spans across the whole graph, the selected nodes are almost outside the ACS.

In our model, the selection rule can be viewed as a suppression of the largest independent ACS. In each time step, we only adopt the update that cannot expand the size of the largest independent ACS unless the update of all the selected nodes could lead to its growth. Due to this effect, the size of the largest independent ACS is suppressed during the evolution. This kind of selection rule is applied in several percolation models in physics such as the EP model, the BFW model, and the avoiding a spanning cluster model. In these models, previous studies have confirmed that the suppression could bring about an explosive percolation transition. Consequently, we expect our model may also exhibit an explosive transition.

Figure 1 illustrates the rule of the model with \( m = 2 \). At some step, two nodes 16 and 3 are selected from the set least fitness and one will be updated. In the case of figure 1(b), the update of node 16 will not affect the size of the largest independent ACS. In contrast, in figure 1(c), node 3 joins the largest independent ACS, causing its size to increase by 1. According to the rule, node 16 will be picked as the updating node since it will not increase the size of the largest independent ACS.

When \( m = 1 \), our model degenerates to the classic ACS evolutionary model [17] in which the size of ACS grows exponentially. However, if \( m \) is larger than 1, the evolution process will present an essential difference. Specifically, the size of the largest independent ACS \( S_1 \) will ‘jump’ by size \( O(N) \) in one single step. This explosive increase stems from the coexistence of several independent ACSs with maximal size during the evolution. We are interested in the number of the largest independent ACSs \( n_1 \) that the system can maintain in the process. In the next section, we discuss this phenomenon in more detail.

3. Properties

Since the evolution process of the model with \( m = 1 \) was investigated in [17], we focus on the situation where \( m \) is larger than 1. Figure 2 shows the evolution process of the model with \( N = 100, m = 2 \), and \( d = 0.25 \). Before the first ACS appears, almost all the nodes are in the least fitness set. Based on this fact, every graph can be roughly assumed as a sample of the classical ER directed random network. In the evolution, the first ACS is typically generated in the form of a cycle, which is also the simplest pattern of ACS. When we update a single node, according to our assumption, a cycle will appear with the probability \( q \equiv \sum_{i=0}^{\infty} p^{i+2} N^{i+1} = \frac{Np}{1-Np} \), where \( p^{i+2} N^{i+1} \) is the probability of forming a cycle of size \( i + 2 \). As we select \( m \) nodes at each step, according to the selection rule, the probability of forming a cycle is \( q^m \). If taking \( t_0 \) as the time that the first ACS
appears, the distribution of $t_0$ can be approximated by a geometric distribution with $p(t_0 = k) = q^m(1 - q^m)^{k-1}$, and the expectation of $t_0$ is $1/q^m$. Although $t_0$ is very large, the appearance of an ACS is inevitable. For convenience, take $t_0$ as the beginning time 0 in figure 2. Before $t_0$, both $S_1$ and the maximal eigenvalue are always zero, and the number of links fluctuates around the expectation value $dN = 25$.

In figure 2(a), an abrupt jump of $S_1$ can be observed at time $t_\alpha$. This is different from the continuous growth of ACS in the classical model. We explain this in more detail later. In figure 2(b), we present the evolution of the largest eigenvalue of $C$. Based on the graph theory, we can draw the following conclusions [17]: (1) An ACS always contains a cycle. (2) If a graph has no ACS, then the largest eigenvalue is zero. (3) If a graph has an ACS, then the largest eigenvalue is no smaller than 1. There is no ACS before $t_0$, so the largest eigenvalue remains zero in this period. As shown in figure 2(c), the number of links changes dramatically when $S_1$ grows to the entire graph or collapses down as in figure 2(a). Both the curves in figures 2(a) and (c) have the same tendency and also have a strong relationship with the least fitness set.

Here, we explain the variation in the number of links shown in figure 2(c). In fact, nodes’ degree, defined as the number of both incoming and outgoing links and population, are not totally unrelated. We investigate the relationship between nodes’ population and
Figure 3. The relationship between nodes’ degree and relative population. For a network with $N = 100$ nodes, we present nodes’ relative population and degree at four different time steps. The $x$-axis shows nodes’ labels in descending order of their relative population. The left and right $y$-axis represent the relative population and degree, respectively. (a) and (b) are results before the transition time $t_\alpha$, while (c) is after the transition. (d) shows their relation when the largest independent ACS spans the entire network.

degree. We first rank nodes’ labels according to their relative population in descending order, then display their degree and relative population. In figure 3, the relative population and degree in a network with $N = 100$ at four different time steps are displayed. It can be seen that, before the largest independent ACS spans the entire network, small population nodes usually have smaller degree. Therefore, the $m$ selected nodes generally have fewer links. At the beginning of the evolution, the link number of small population nodes is below $2(N - 1)p$, which is the average number of directed links added to the network in one update. If we remove these links and add an average $2(N - 1)p$ new links, the number of links will increase. As the link number in the network grows, the degree of small population nodes will become larger than $2(N - 1)p$. At this time, if we remove nodes and form new directed links with probability $p$, the link number will decrease. This explains why the number of links has large fluctuation during the evolution.

The nature of the explosive jump of $S_1$ is better revealed in figure 3. The maximal increase of the largest independent ACS $\Delta S_{\text{max}}$ is linear with the system size $N$ for $m \geq 2$. For $m = 1, 2, 3, 4$, we fit the data with a linear function $\Delta S_{\text{max}} = aN + b$ by the least squares method. Then, we plot $(\Delta S_{\text{max}} - b)/N$ versus $N$, shown in figure 4. We observe that the value $(\Delta S_{\text{max}} - b)/N$ keeps a small value near 0 for $m = 1$.
Figure 4. Maximal jump of the $S_1$. $\Delta S_{\text{max}}$ represents the maximal increase of $S_1$ for an update of a single node before the independent ACS spanning across the entire graph. The data points are averaged on 100 random instances, and the error bars are the standard deviations. This shows that the $\Delta S_{\text{max}}$ is linear with the system size $N$ for $m = 2, 3$ and 4. The slopes of the fitting lines (black lines) are 0.018, 0.29, 0.42 and 0.45 for $m = 1, 2, 3$, and 4, respectively.

In fact, it should be $(\Delta S_{\text{max}} - b)/N \sim 1/N$ since $\Delta S_{\text{max}}$ is at least 1. On the other hand, $(\Delta S_{\text{max}} - b)/N$ remains a positive constant for $m \geq 2$. With this trend, we have $\lim_{N \to \infty} \frac{\Delta S_{\text{max}}}{N} = \lim_{N \to \infty} \frac{\Delta S_{\text{max}} - b}{N} > 0$. This evidence from simulation shows that our model can lead to an abrupt emergence of a macroscopic independent ACS, irrelevant of the network size $N$.

In the evolution process, there are several large independent ACSs coexisting and finally merging together to overtake the previous largest ACS. With the increase of $m$, it is more difficult for $S_1$ to grow. As the suppression strengthens, the system can maintain more large independent ACSs, and there will be fewer nodes outside the ACS. Moreover, it is nearly impossible to break up a large ACS into smaller ones since the updating nodes mostly do not belong to ACSs. For large $m$, the increase of $S_1$ will mainly depend on merging two giant independent ACSs rather than adding isolated nodes. In the next section, we analyze this process theoretically and perform simulations to verify the results.

4. Theoretical analysis

Based on the investigation above, it is clear that the models with $m = 1$ and $m \geq 2$ are very different. In [17, 23], the properties of the model with $m = 1$ were discussed with analytical methods so in this section, we analyze the model with $m \geq 2$. First, we consider
Figure 5. The relationship between the value \((\Delta S_{\text{max}} - b)/N\) and \(N\). \(\Delta S_{\text{max}}\) is the maximal increase of \(S_1\) for an update of a single node before the independent ACS spanning across the entire graph. \(b\) is obtained by fitting the data with a linear function \(\Delta S_{\text{max}} = aN + b\). The data points are averaged on 100 random instances, and the error bars are the standard deviations.

the probability of adding a node to an ACS that has \(S\) nodes (denoted as \(P_{\text{add}}\)). Based on the definition of ACS, a node could be added to ACS if and only if there is an incoming link from the ACS to this node. Therefore, the probability \(P_{\text{add}}\) is:

\[
P_{\text{add}} = 1 - (1 - p)^S.
\]

Here, \(p = d/N\) is the probability of linking an edge in an update. As our model is applied to sparse graphs in which the linking probability \(p\) is very small, equation (5) can be approximated by

\[
P_{\text{add}} \approx pS.
\]

At each time step, we update only one node out of the \(m\) selected nodes. The size \(S_1\) would change if all the \(m\) nodes had at least one incoming link from \(S'_1\). The size of \(S'_1\) is \(n_1(t)S_1\) and based on equation (6), and we can calculate the average change of \(S_1\) for the sparse graph

\[
\Delta S_1 = (n_1(t)ps_1)^m \Delta t.
\]

If we take the first time that an independent ACS appears as the beginning step \(t_0\), equation (7) can be integrated from \(t_0\) to \(t_1\) as

\[
\int_{S_1(t_0)}^{S_1(t_1)} \frac{1}{S_1^n} dS_1 = p^m \int_{t_0}^{t_1} n_1^m(t) dt.
\]

Taking \(S_0\) as the initial size of \(S_1(t_0)\), the equation above can be written as

\[
\frac{1}{S_0^{m-1}} - \frac{1}{S_1(t)^{m-1}} = (m - 1)p^m \int_{t_0}^{t_1} n_1^m(t) dt.
\]
For the last part of the equation (9), we define a quantity \( n_\beta \) such that \( \int_{t_0}^{t_1} n_\beta^m(t) dt = n_\beta^m(t_1 - t_0) \). In fact, \( n_\beta \) can be viewed as the average tolerance of the largest independent ACSs during the evolution. In other words, the larger \( n_\beta \) is, the more largest independent ACSs the system can maintain. With this definition, we get the function of \( S_1(t) \) for different values of \( m \) as

\[
S_1(t) = \frac{1}{1/S_0^{m-1} - n_\beta^m (m-1)p^m(t-t_0)}^{\frac{1}{m-1}}. \tag{10}
\]

The evolution function for \( m = 1 \) should be \( S_1(t) \sim S_1(t_0)e^{p(t-t_0)} \), which presents an exponential and continuous increase of the ACS size. This is quite different from the function of \( m \geq 2 \) above. For the parameter \( m \geq 2 \), when \( t \) takes a certain value, the denominator in equation (10) will become zero. Therefore, we identify a phase transition at some time step \( t_\alpha \). The threshold \( t_\alpha \) is

\[
\begin{align*}
t_\alpha - t_0 &= \frac{1}{n_\beta^m (m-1)p^m S_0^{m-1}}, \tag{11} \\
\frac{t_\alpha - t_0}{N^m} &= \frac{1}{n_\beta^m (m-1)d^m S_0^{m-1}}. \tag{12}
\end{align*}
\]

Figures 6 and 7 are the numerical results with \( m = 2 \) and \( m = 3 \), respectively. Because of the assumption of the sparse graph, we choose the average in-degree \( d \) from 0.2 to 0.3. For the fixed value of \( S_0 \), we performed the simulations starting from an initial ACS with size \( S_0 \) on random graphs. To determine the threshold, in simulations we took the time step where the increase of the largest independent ACS’s size \( S_1 \) exceeds 10% of the system size \( N \) as \( t_\alpha \). Since the starting time \( t_0 \) was set to 0, we rearranged equation (12) as follows

\[
\frac{t_\alpha}{N^m} = \frac{1}{n_\beta^m (m-1)S_0^{m-1}} \cdot \frac{1}{d^m}. \tag{13}
\]

Therefore, there is a linear relation between the quantity \( t_\alpha/N^m \) and \( 1/d^m \). In figures 6 and 7, there is clearly a linear relation between these two quantities, and the numerical results agree with the fitting line very well. To check the fitting errors for different system sizes, we define the fitting error as the average distance from the data points to the corresponding fitting line. For both cases, the fitting errors decrease as the system size increases.

Moreover, from the slope (denoted as \( k \)) of the fitting line, we can obtain \( n_\beta \) by relation

\[
k = \frac{1}{(m-1)S_0^{m-1}n_\beta^m}. \tag{14}
\]

To compare the simulations with the actual number of the largest independent ACSs coexisting in the system during the evolution, we recorded this number in each time step during simulations. Then, we took the average of these values as the real \( n_\beta \). Table 1 displays the results of \( n_\beta \) from both the theoretical analysis and simulations. As \( m \) increases from 2 to 3, \( n_\beta \) grows significantly, which means stricter suppression will allow the system to maintain more large independent ACSs during the evolution process. This further leads to the result that the jump of \( S_1 \) is enhanced significantly when \( m \) increases, as shown in figure 3. Furthermore, as the system size \( N \) grows, \( n_\beta \) also increases slightly.
Evolution of autocatalytic sets in a competitive percolation model

Figure 6. The threshold $t_\alpha$ with $m = 2$ versus average degree $d$ for different $S_0$. Different colors represent distinct system sizes. The data are obtained by averaging 100 instances in each case. The black lines are the linear fits with least-square regression. (a)–(c) Results with different initial size $S_0$ from 2, 3 and 4. (d) Errors for different system size $N$. Error is defined as the average distance between the data points and the theoretical line.

This is a natural result of the increase of system size. Therefore, the choice of $m$ will dramatically affect the formation of the giant independent ACS, both the emergence time and the jump of size.

5. Conclusions

ACS is an important concept in the evolution dynamics of biological, chemical, and ecological systems. The emergence of an ACS is often used to explain the mechanism by which a complex chemical organization or species could have evolved. In this paper, by imposing an $m$-selection rule, we propose a competitive model to investigate the evolution process of ACSs under suppression. In this model, we observe an explosive phase transition where a microscopic independent ACS appears abruptly. By simulations, its size is found to grow linearly with the system size. We derive the threshold $t_\alpha$ analytically and verify our result through numerical simulations on different system sizes and various choices of $m$. As the suppression increases, the phase transition is dramatically deferred. Furthermore, we
**Figure 7.** The threshold $t_\alpha$ with $m = 3$ versus average degree $d$ for different $S_0$. Different sizes are marked with distinct colors and symbols. Each data point is obtained by averaging 100 simulations. The black lines are the fitting lines. (a)–(c) Results with different initial size $S_0$ from 2, 3 and 4. (d) Fitting errors versus $N$.

**Table 1.** The comparison of $n_\beta$ between the simulations and the theoretical results in figures 6 and 7(a)–(c) with $m = 2$ and $m = 3$. The first line shows the theoretical results calculated from equation (13) and the second line represents results from simulations. Each simulation result is obtained by averaging values from 100 simulations.

| $N$ | 100 | 150 | 200 | 250 | 300 |
|-----|-----|-----|-----|-----|-----|
| $S_0 = 2$ | 2.71 | 3.20 | 3.46 | 3.67 | 3.97 |
| $m = 2$ | 2.72 | 3.11 | 3.41 | 3.64 | 3.85 |
| $S_0 = 3$ | 2.65 | 2.97 | 3.31 | 3.52 | 3.71 |
| $m = 2$ | 2.50 | 2.84 | 3.10 | 3.40 | 3.56 |
| $S_0 = 4$ | 2.47 | 2.70 | 3.09 | 3.29 | 3.56 |
| $m = 2$ | 2.27 | 2.59 | 2.84 | 3.08 | 3.30 |
| $S_0 = 2$ | 7.39 | 9.13 | 10.89 | 11.74 | 12.98 |
| $m = 3$ | 7.24 | 8.73 | 10.34 | 11.35 | 12.50 |
| $S_0 = 3$ | 6.50 | 7.91 | 9.20 | 10.33 | 11.12 |
| $m = 3$ | 6.16 | 7.42 | 8.64 | 9.81 | 10.50 |
| $S_0 = 4$ | 5.55 | 6.96 | 8.15 | 9.01 | 9.96 |
| $m = 3$ | 5.26 | 6.43 | 7.50 | 8.37 | 9.20 |

doi:10.1088/1742-5468/2014/11/P11018
explore the evolution process of the largest independent ACS. To quantify the tolerance of the system to the emergence of a microscopic independent ACS, we define a quantity $n_B$ that describes the average number of large independent ACSs during the evolution process. We find that $n_B$ increases as the selection rule becomes stricter. Therefore, on average, a system with larger $m$ would contain more large independent ACSs during its evolution. Our model provides a possible explanation for the sudden appearance of a class of species or chemical organizations in specific situations. By only introducing a selection rule, the evolution of ACS presents qualitative differences from the classical model. Our study sheds light on the research of the evolutionary process of ACS and provides helpful instructions to design effective strategies to control the appearance of ACS in practice.

Acknowledgments

We wish to thank L A N Amaral for useful discussions. This work is supported by National Natural Science Foundation of China (11290141, 11290143, 11201018), International Cooperation Project No.2010DFR00700, Fundamental Research of Civil Aircraft No. MJ-F-2012-04, Foundation of Dalian University of Technology No.3004852038, and Innovation Foundation of BUAA for PhD Graduates.

References

[1] Albert R and Barabási A-L 2002 Rev. Mod. Phys. 74 47–97
[2] Newman M E J 2003 SIAM Rev. 45 167–256
[3] Castellano C, Fortunato S and Loreto V 2009 Rev. Mod. Phys. 81 591–646
[4] Pei S, Munchik L, Andrade J S Jr, Zheng Z and Malse H A 2014 Sci. Rep. 4 5547
[5] Pei S, Tang S, Yan S, Jiang S, Zhang X and Zheng Z 2012 Phys. Rev. E 86 021909
[6] Fell D A and Wagner A 2000 Nat. Biotechnol. 18 1121–2
[7] Jeong H, Tombar B, Albert A, Oltvai Z N and Barabási A-L 2000 Nature 407 651–4
[8] Williams R J and Martinez N D 2000 Nature 407 651–4
[9] Camacho J, Guimerà R and Amaral L A N 2002 Phys. Rev. Lett. 88 228102
[10] Watts D J and Strogatz S H 1998 Nature 393 440–2
[11] Kauffman S A 1986 J. Theor. Biol. 119 1–24
[12] Farmer J D, Kauffman S A and Packard N H 1986 Physica D 22 50–67
[13] Fontana W and Buss L W 1994 Bull. Math. Biol. 56 1–64
[14] Stadler P F, Fontana W and Miller J H 1993 Physica D 63 378–92
[15] Joyce G F 1989 Nature 338 217–23
[16] Bak P and Sneppen K 1993 Phys. Rev. Lett. 71 4083–6
[17] Jain S and Krishna S 1998 Phys. Rev. Lett. 81 5684–7
[18] Jain S and Krishna S 2002 Phys. Rev. E 65 026103
[19] Eigen M 1971 Naturwissenschaften 58 465–523
[20] Kauffman S A 1971 J. Cybern. 1 71–96
[21] Rossler O E 1971 Z. Naturf, b 26 741–6
[22] Zhang R, Wei W, Pei S, Guo B, Zhang Y and Zheng Z 2014 Adv. Electr. Electron. Eng. 92 109
[23] Jain S and Krishna S 2001 Proc. Natl Acad. Sci. USA 98 543–7
[24] Stauffer D and Aharony A 1994 Introduction to Percolation Theory 2nd edn (London: Taylor and Francis)
[25] Erdős P and Rényi A 1960 Publ. Math. Inst. Hungarian Acad. Sci. 5 17
[26] Achlioptas D, D’Souza R M and Spencer J 2009 Science 323 1453
[27] da Costa R A, Dorogovtsev S N, Goltsev A V and Mendes J F F 2010 Phys. Rev. Lett. 105 255701

doi:10.1088/1742-5468/2014/11/P11018
Evolution of autocatalytic sets in a competitive percolation model

[28] Riordan O and Warnke L 2011 Science 333 322–4
[29] Grassberger P, Christensen C, Bizhani G, Son S and Paczuski M 2011 Phys. Rev. Lett. 106 255701
[30] Lee H K, Kim B J and Park H 2011 Phys. Rev. E 84 020101
[31] Bastas N, Kosmidis K and Argyakis P 2011 Phys. Rev. E 84 066112
[32] Araújo N A M and Herrmann H J 2010 Phys. Rev. Lett. 105 035701
[33] Cho Y S, Hwang S, Herrmann H J and Kahng B 2013 Science 339 1185–7
[34] Chen W and D’Souza R M 2011 Phys. Rev. Lett. 106 115701
[35] Zhang R, Wei W, Guo B, Zhang Y and Zheng Z 2013 Physica A 392 1232–45
[36] Radicchi F and Fortunato S 2009 Phys. Rev. Lett. 103 168701
[37] Ziff R M 2009 Phys. Rev. Lett. 103 045701