PHYSICAL REVIEW E **75**, 066114 (2007)

Unifying evolutionary and network dynamics

Samarth Swarup[*]
Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Les Gasser
Graduate School of Library and Information Science and Department of Computer Science,
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

(Received 15 October 2006; published 28 June 2007)

Many important real-world networks manifest small-world properties such as scale-free degree distributions, small diameters, and clustering. The most common model of growth for these networks is preferential attachment, where nodes acquire new links with probability proportional to the number of links they already have. We show that preferential attachment is a special case of the process of molecular evolution. We present a single-parameter model of network growth that unifies varieties of preferential attachment with the quasispecies equation (which models molecular evolution), and also with the Erdős-Rényi random graph model. We suggest some properties of evolutionary models that might be applied to the study of networks. We also derive the form of the degree distribution resulting from our algorithm, and we show through simulations that the process also models aspects of network growth. The unification allows mathematical machinery developed for evolutionary dynamics to be applied in the study of network dynamics, and vice versa.

DOI: 10.1103/PhysRevE.75.066114 PACS number(s): 89.75.Hc, 89.75.Da, 87.23.Kg

I. INTRODUCTION

The study of networks has become a very active area of research since the discovery of small-world networks [1,2]. Small-world networks are characterized by scale-free degree distributions, small diameters, and high clustering coefficients. Many real networks, such as neuronal networks [2], power grids [3], the world wide web [4], and human language [5], have been shown to be small world. The small-world character has important consequences. For example, such networks are found to be resistant to random attacks, but susceptible to targeted attacks, because of the power-law nature of the degree distribution.

The process most commonly invoked for the generation of such networks is called preferential attachment [6,7]. Briefly, new links attach preferentially to nodes with more existing links. Simon analyzed this stochastic process, and derived the resulting distribution [8]. This simple process has been shown to generate networks with many of the characteristics of small-world networks, and has largely replaced the Erdős-Rényi random graph model [9] in modeling and simulation work.

Another major area of research in recent years has been the consolidation of evolutionary dynamics [10], and its application to alternate areas of research, such as language [11]. This work rests on the foundation of quasispecies theory [12,13], which forms the basis of much subsequent mathematical modeling in theoretical biology.

In this paper we bring together network generation models and evolutionary dynamics models (and particularly quasispecies theory) by showing that they have a common underlying probabilistic model. This unified model relates both processes through a single parameter, called a *transfer matrix*. The unification allows mathematical machinery developed for evolutionary dynamics to be applied in the study of network dynamics, and vice versa. The rest of this paper is organized as follows. First we describe the preferential attachment algorithm and the quasispecies model of evolutionary dynamics. Then we show that we can describe both of these with a single probabilistic model. This is followed by a brief analysis, and some simulations, which show that power-law degree distributions can be generated by the model, and that the process can also be used to model some aspects of network growth, such as densification power laws and shrinking diameters.

II. PREFERENTIAL ATTACHMENT

The preferential attachment algorithm specifies a process of network growth in which the addition of new (in-)links to nodes is random, but biased according to the number of (in-)links the node already has. We identify each node by a unique type i, and let x_i indicate the proportion of the total number of links in the graph that is already assigned to node i. Then Eq. (1) gives the probability $P(i)$ of adding a new link to node i [6]:

$$P(i) = \alpha x_i^{\gamma}, \tag{1}$$

where α is a normalizing term, and γ is a constant. As γ approaches 0 the preference bias disappears; $\gamma > 1$ causes exponentially greater bias from the existing in-degree of the node.

III. EVOLUTIONARY DYNAMICS AND QUASISPECIES

Evolutionary dynamics describes a population of *types* (species, for example) undergoing change through replication, mutation, and selection. Suppose there are N possible types, and let $s_{i,t}$ denote the number of individuals of type i

---

[*] Electronic address: swarup@uiuc.edu

in the population at time t. Each type has a fitness f_i, which determines its probability of reproduction. At each time step, we select, with probability proportional to fitness, one individual for reproduction. Reproduction is noisy, however, and there is a probability q_{ij} that an individual of type j will generate an individual of type i. The expected value of the change in the number of individuals of type i at time t is given by,

$$\Delta s_{i,t} = \frac{\sum_j f_j s_j q_{ij}}{\sum_j f_j s_j}. \tag{2}$$

This is known as the quasispecies equation [13]. The fitness f_i is a constant for each i. Fitness can also be *frequency dependent*, i.e., it can depend on which other types are present in the population. In this case the above equation is known as the replicator-mutator equation [10,14].

IV. A GENERALIZED POLYA URN MODEL THAT DESCRIBES BOTH PROCESSES

Urn models have been used to describe both preferential attachment [15] and evolutionary processes [16]. Here we describe an urn process derived from the quasispecies equation that also gives a model of network generation. In addition, this model of network generation will be seen to unify the Erdős-Rényi random graph model [9] with the preferential attachment model.

Our urn process is as follows.

(a) We have a set of n urns, which are all initially empty except for one, which has one ball in it.

(b) We add balls one by one, and a ball goes into urn i with probability proportional to $f_i m_i$, where f_i is the "fitness" of urn i, and m_i is the number of balls already in urn i.

(c) If the ball is put into urn j, then a ball is taken out of urn j, and moved to urn k with probability q_{kj}.

The matrix $Q=[q_{ij}]$, which we call the transfer matrix, is the same as the mutation matrix in the quasispecies equation.

This process describes the preferential attachment model if we set the fitness f_i to be proportional to $m_i^{\gamma-1}$, where γ is a constant [as in Eq. (1)]. Now we get a network generation algorithm in much the same way as Chung *et al.* did [15], where each ball corresponds to a half edge, and each urn corresponds to a node. Placing a ball in an urn corresponds to linking to a node, and moving a ball from one urn to another corresponds to rewiring. We call this algorithm "noisy preferential attachment" (NPA). If the transfer matrix is set to be the identity matrix, noisy preferential attachment reduces to pure preferential attachment.

In the NPA algorithm, just as in the preferential attachment algorithm, the probability of linking to a node depends only on the number of in-links to that node. The "from" node for a new edge is chosen uniformly randomly. In keeping with standard practice, the graphs in the next section show only the in-degree distribution. However, since the from nodes are chosen uniformly randomly, the total degree distribution has the same form. Consider the case where the transfer matrix is almost diagonal, i.e., q_{ii} is close to 1, and the same $\forall i$, and all the q_{ij} are small and equal, $\forall i \neq j$. Let $q_{ii}=p$ and

$$q_{ij} = \frac{1-p}{n-1} = q, \quad \forall \ i \neq j. \tag{3}$$

Then, the probability of the new ball being placed in bin i is

$$P(i) = \alpha m_i^{\gamma} p + (1 - \alpha m_i^{\gamma}) q, \tag{4}$$

where α is a normalizing constant. That is, the ball could be placed in bin i with probability αm_i^{γ} and then replaced in bin i with probability p, or it could be placed in some other bin with probability $(1-\alpha m_i^{\gamma})$, and then transferred to bin i with probability q. Rearranging, we get

$$P(i) = \alpha m_i^{\gamma}(p - q) + q. \tag{5}$$

In this case, NPA reduces to preferential attachment with initial attractiveness [17], where the initial attractiveness (q, here) is the same for each node. We can get different values of initial attractiveness by setting the transfer matrix to be nonuniform. We can get the Erdős-Rényi model by setting the transfer matrix to be entirely uniform, i.e., $q_{ij} = 1/n, \forall i,j$. Thus the Erdős-Rényi model and the preferential attachment model are seen as two extremes of the same process, which differ in the transfer matrix Q.

This process also obviously describes the evolutionary process when $\gamma=1$. In this case, we can assume that at each step we first select a ball from among all the balls in all the urns with probability proportional to the fitness of the ball (assuming that the fitness of a ball is the same as the fitness of the urn in which it is). The probability that we will choose a ball from urn i is proportional to $f_i m_i$. We then replace this ball and add another ball to the same urn. This is the replication step. This is followed by a mutation step as before, where we choose a ball from the urn and either replace it in the urn with with probability p or move it to any one of the remaining urns. If we assume that all urns (i.e. all types or species) have the same *intrinsic* fitness, then this process reduces to the preferential attachment process.

Having developed the unified NPA model, we can now point toward several concepts in quasispecies theory that are missing from the study of networks, that NPA makes it possible to investigate.

(a) Quasispecies theory assumes a genome, a bit string for example. This allows the use of a distance measure on the space of types.

(b) Mutations are often assumed to be point mutations, i.e., they can flip one bit. This means that a mutation cannot result in just *any* type being introduced into the population, only a neighbor of the type that gets mutated.

(c) This leads to the notion of a quasispecies, which is a cloud of mutants that are close to the most-fit type in genome space.

(d) Quasispecies theory also assumes a fitness landscape. This may in fact be flat, leading to neutral evolution [18]. Another (toy) fitness landscape is the sharply peaked landscape, which has only one peak and therefore does not suffer from problems of local optima. In general, though, fitness landscapes have many peaks, and the ruggedness of the landscape (and how to evaluate it) is an important concept in evolutionary theory. The notion of (node) fitness is largely missing from network theory (with a couple of exceptions [19,20]), though the study of networks might benefit greatly from it.

(e) The event of a new type entering the population and "taking over" is known as fixation. This means that the entire population eventually consists of this new type. Typically, we speak of gene fixation, i.e., the probability that a single new gene gets incorporated into all genomes present in the population. Fixation can occur due to drift (neutral evolution) as well as due to selection.

V. ANALYSIS AND SIMULATIONS

We next derive the degree distribution of the network. Since there is no "link death" in the NPA algorithm and the number of nodes is finite, the limiting behavior in our model is not the same as that of the preferential attachment model (which allows introduction of new nodes). This means that we cannot reuse Simon's result [8] directly to derive the degree distribution of the network that results from NPA.[1]

A. Derivation of the degree distribution

Suppose there are N urns and n balls at time t. Let $x_{i,t}$ denote the fraction of urns with i balls at time t. We choose a ball uniformly at random and "replicate" it, i.e., we add a new ball (and replace the chosen ball) into the same urn. Uniformly random choice corresponds to a model where all the urns have equal intrinsic fitness. We follow this up by drawing another ball from this urn and moving it to a uniformly randomly chosen urn (from the $N-1$ other urns) with probability $q=(1-p)/(N-1)$, where p is the probability of putting it back in the same urn. Let $P_1(i)$ be the probability that the ball to be replicated is chosen from an urn with i balls. Let $P_2(i)$ be the probability that the new ball is placed in an urn with i balls. The net probability that the new ball ends up in an urn with i balls is

$$P(i) = P_1(i) \text{ and } P_2(i) \quad \text{or} \quad \bar{P}_1(i) \text{ and } P_2(i). \tag{6}$$

The probability of selecting a ball from an urn with i balls is

$$P_1(i) = \frac{Nx_{i,t}i}{n_0+t},$$

where n_0 is the number of balls in the urns initially. $P_2(i)$ depends on the outcome of the first step.

$$P_2(i) = \begin{cases} p+(Nx_{i,t}-1)q & \text{when step 1 is ``successful,''} \\ Nx_{i,t}q & \text{when step 1 is a ``failure.''} \end{cases}$$

Putting these together, we get

$$\begin{aligned} P(i) &= \frac{Nx_{i,t}i}{n_0+t}[p+(Nx_{i,t}-1)q] + \left(1-\frac{Nx_{i,t}i}{n_0+t}\right)Nx_{i,t}q \\ &= \frac{Nx_{i,t}i}{n_0+t}(p-q) + Nx_{i,t}q. \end{aligned}$$

Now we calculate the expected value of $x_{i,t+1}$. $x_{i,t}$ will increase if the ball goes into an urn with $i-1$ balls. Similarly, it will decrease if the ball ends up in an urn with i balls. Otherwise it will remain unchanged. Remembering that $x_{i,t}$ is the *fraction* of urns with i balls at time t, we write

$$Nx_{i,t+1} = \begin{cases} Nx_{i,t}+1 \text{ with probability } \dfrac{Nx_{i-1,t}(i-1)}{n_0+t}(p-q)+Nx_{i-1,t}q, \\ Nx_{i,t}-1 \text{ with probability } \dfrac{Nx_{i,t}i}{n_0+t}(p-q)+Nx_{i,t}q, \\ Nx_{i,t} \quad \text{otherwise}. \end{cases}$$

From this, the expected value of $x_{i,t+1}$ works out to be

$$x_{i,t+1} = \left(1-\frac{i(p-q)}{n_0+t}-q\right)x_{i,t} + \left(\frac{(i-1)(p-q)}{n_0+t}+q\right)x_{i-1,t}. \tag{7}$$

We can show the approximate solution for $x_{i,t}$ to be,

$$x_{i,t} = \frac{1-p}{N}\frac{r^{i-1}\Gamma(i)}{\prod_{k=1}^{i}(kr+1)}(t+1)(1-q)^{t-1}, \tag{8}$$

where $r=(p-q)/(1-q)$. This approximation is valid when $t \ll N$. See Appendix A for details. For any particular i, the shape of this curve is given by $t(1-q)^t$. An example curve is shown in Fig. 1. This matches our intuition. Initially, $x_{i,t}=0$ for $i>1$. As t increases, $x_{i,t}$ increases through mutations. However, since N is finite and we keep adding balls, eventually the number of bins with i balls must go to zero for any particular i. Thus $x_{i,t}$ must eventually start decreasing, which is what we see in Fig. 1. The middle term can be simplified further as

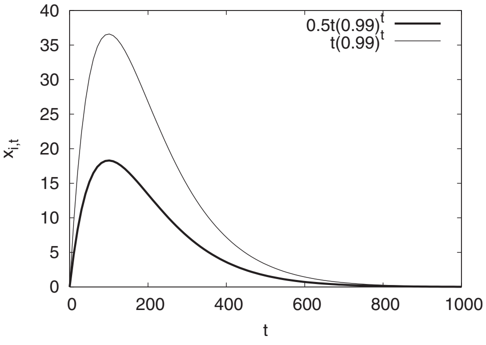

FIG. 1. Example $x_{i,t}$ curves.

---

[1] Simon (and Yule [27] before him) applied their stochastic model to the estimation of numbers of species within genera, but the notion of quasispecies was unknown at the time, and it addresses a much wider range of issues than species frequency.

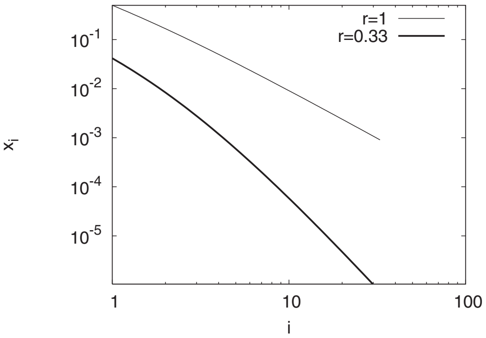

FIG. 2. Form of the degree distribution.

$$\frac{r^{i-1}}{\prod_{k=1}^{i}(kr+1)} = \frac{r^{i-1}}{\prod_{k=1+1/r}^{i+1/r}(kr)} = \frac{1}{r\prod_{k=1+1/r}^{i+1/r}(k)} = \frac{\Gamma(1/r)}{r^2\Gamma(i+1+1/r)}.$$

Therefore, in terms of i, equation (8) can be written as (for fixed t),

$$x_i = C\frac{\Gamma(i)}{\Gamma\left(i+1+\frac{1}{r}\right)}, \tag{9}$$

where C is a constant. This is the form of the degree distribution. This is a power law, because as $i \to \infty$ Eq. (9) tends to $i^{-(1+1/r)}$ [see the discussion of Eq. (1.4) in [8, p. 426]]. This is also demonstrated in the sample plots in Fig. 2.

These results are confirmed through simulation. We did an experiment where the number of possible nodes was set to 100 000, and 10000 links were added. The experiment was repeated for values of p ranging from 0.01 to 0.99, in steps of 0.01. Figure 3 shows a plot of *coherence* ϕ, which is defined as

$$\phi = \sum_i x_i^2. \tag{10}$$

Coherence is a measure of the nonuniformity of the degree distribution. It is 1 when a single node has all the links. When all nodes have one link each, coherence has its lowest value, $1/N$. We see that as p increases (i.e., mutation rate decreases), coherence also increases. This is borne out by the degree distribution plots (Figs. 4–6). The degree distribution is steeper for lower values of p.

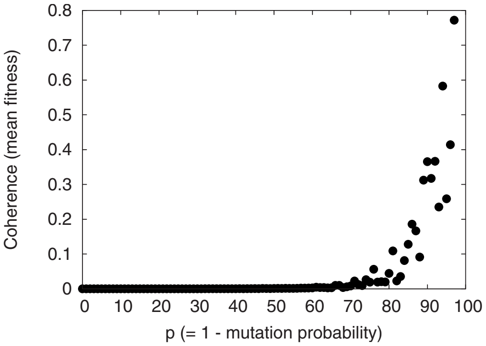

FIG. 3. N=100 000, number of edges=10 000.

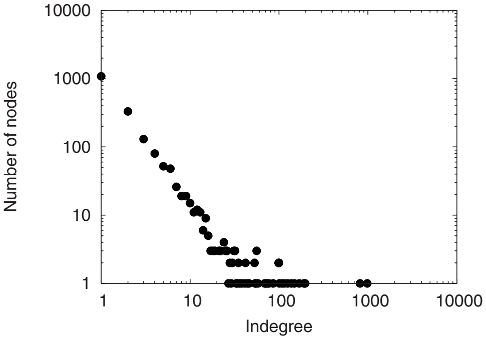

FIG. 4. p=0.8, N=100 000, number of edges=10 000.

B. Stability

We can rewrite Eq. (2) as

$$\Delta s_i = \frac{1}{\sum_j f_j s_j}\left(f_i s_i q_{ii} + \sum_{j\neq i} f_j s_j q_{ij}\right). \tag{11}$$

The first term in the parentheses represents the change in s_i due to selection. Some of the copies of type i are lost due to mutation. The fraction that are retained are given by the product $f_i q_{ii}$. If this product is greater than 1, the proportion of type i will increase due to selection, otherwise it will decrease. The second term represents the contribution to type i due to mutation from all the other types in the population. Thus, if s_i decreases toward zero due to a selective disadvantage, it will be maintained in the population at "noise" level due to mutations.

This leads to the notion of an error threshold. Suppose that the fitness landscape has only one peak. This is known as the sharply peaked landscape. Suppose further that mutations alter only one position on the genome at a time. Then it can be shown that, if the mutation rate is small enough, the population will be closely clustered about the fittest type. The fittest type keeps getting regenerated due to selection, and mutations generate a cloud of individuals with genomes very close to the genome of the fittest type. This cloud is known as a *quasispecies* [21].

If, on the other hand, the mutation rate is above a certain threshold (essentially $1/f_i$, where i is the fittest type) then all types will persist in the population in equal proportions. This threshold is known as the error threshold.

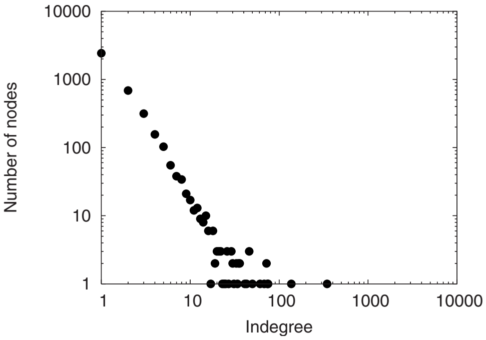

FIG. 5. p=0.6, N=100 000, number of edges=10 000.

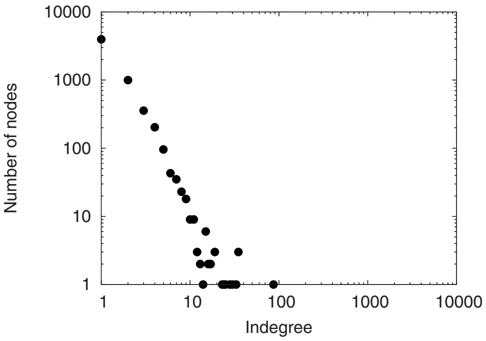

FIG. 6. $p=0.4$, $N=100\,000$, number of edges$=10\,000$.

VI. FITNESS LANDSCAPES AND NEUTRAL EVOLUTION

We have seen above that noisy preferential attachment is equivalent to molecular evolution where all intrinsic fitnesses are equal. If node fitnesses are allowed to be different, we get standard quasispecies behavior. If the mutation rate is low enough, the fittest node dominates the network and acquires nearly all the links. If the mutation rate is high enough to be over the error threshold, no single node dominates.

Figures 7 and 8 show simulations where nodes are assigned intrinsic fitness values uniformly randomly in the range (0,1), for different values of p. We see that when p is high (0.9), i.e., mutation rate is low, the degree distribution stretches out along the bottom, and one or a few nodes acquire nearly all the links. When $p=0.4$, though, we do not get this behavior, because the mutation rate is over the error threshold.

Since we generally do not see a single node dominating in real-world networks, we are led to one of two conclusions: either mutation rates in real-world networks are rather high, or the intrinsic fitnesses of the nodes are all equal. The former seems somewhat untenable. The latter suggests that most networks undergo neutral evolution [18].

Fitness landscapes can also be dynamic. Golder and Huberman give examples of short-term dynamics in collaborative tagging systems (in particular Del.icio.us) [22]. Figures 9 and 10, which are taken from their paper, show two instances of the rate at which two different web sites acquired bookmarks. The first one shows a peak right after it appears, before the rate of bookmarking drops to a baseline level. The second instance shows a web site existing for a while before it suddenly shows a peak in the rate of bookmarking. Both are examples of dynamic, i.e., changing, fitness. Wilke *et al.* have shown that, in the case of molecular evolution, a rapidly changing fitness landscape is equivalent to the time-averaged fitness landscape [23]. Thus, while the short-term dynamics show peaks in link (or bookmark) acquisition, the long-term dynamics could still be neutral or nearly neutral.

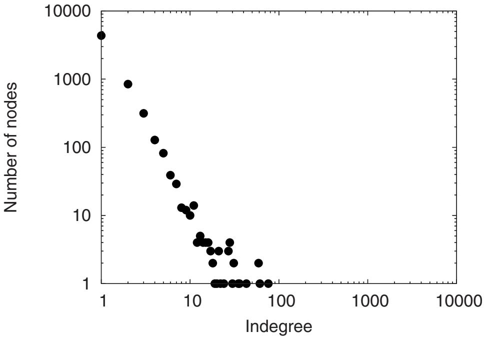

FIG. 7. $p=0.4$, $N=100\,000$, number of edges$=10\,000$, node fitnesses are uniformly randomly distributed between 0 and 1.

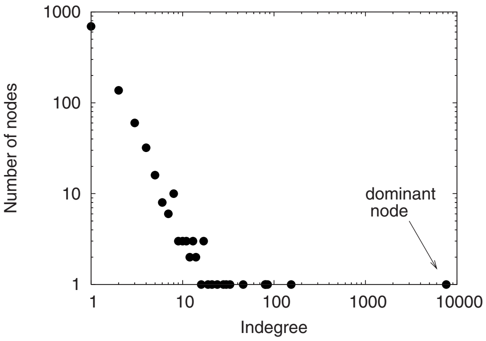

FIG. 8. $p=0.9$, $N=100\,000$, number of edges$=10\,000$, node fitnesses are uniformly randomly distributed between 0 and 1.

VII. DYNAMICAL PROPERTIES OF REAL-WORLD NETWORKS

Leskovec *et al.* point out that, though models like preferential attachment are good at generating networks that match static "snapshots" of real-world networks, they do not appropriately model how real-world networks change over time [24]. They point out two main properties that are observed for several real-world networks over time: *densification power laws* and *shrinking diameters*. The term densification power law refers to the fact that the number of edges grows superlinearly with respect to the number of nodes in the

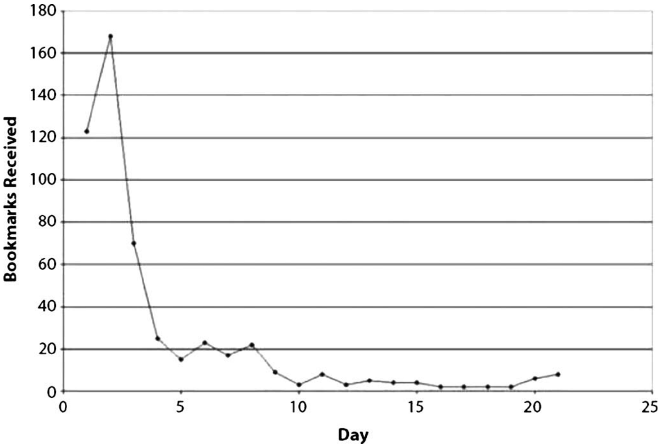

FIG. 9. This is Fig. 6a from [22]. It shows the number of bookmarks received against time (day number). This particular site acquires a lot of bookmarks almost immediately after it appears, but thereafter receives few bookmarks. Reproduced with permission from S. A. Golder and B. A. Huberman, "Usage patterns of collaborative tagging systems," J. Info. Sci., Copyright (© CILIP, 200), by permission of Sage Publications, Ltd.

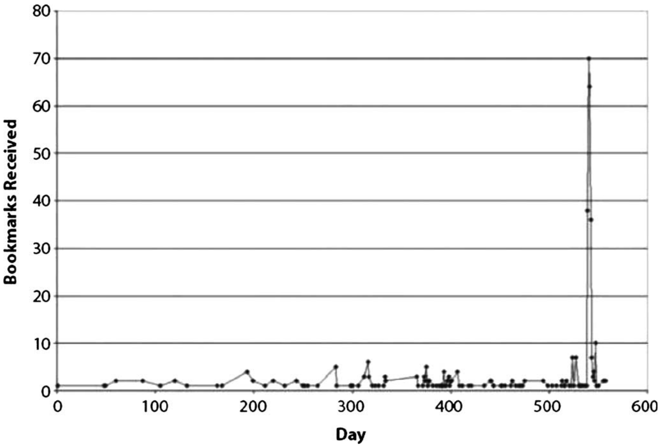

FIG. 10. This is Fig. 6b from [22]. It shows the number of bookmarks received against time (day number). This particular site suddenly acquires a lot of bookmarks in a short period of time, although it has existed for a long time. Reproduced with permission from S. A. Golder and B. A. Huberman, "Usage patterns of collaborative tagging systems," J. Info. Sci., Copyright (© CILIP, 200), by permission of Sage Publications, Ltd.

network. In particular, it grows as a power law. This means that these networks are getting more densely connected over time. The second surprising property of the dynamics of growing real-world networks is that the diameter (or 90th percentile distance, which is called the effective diameter) decreases over time. In most existing models of scale-free network generation, it has been shown that the diameter increases very slowly over time [25]. Leskovec *et al.* stress the importance of modeling these dynamical aspects of network growth, and they present an alternate algorithm that displays both the above properties.

Noisy preferential attachment can also show these properties if we slowly decrease the mutation rate over time. Figures 11 and 12 show the effective diameter of the network and the rate of change of the number of nodes with respect to the number of edges for a simulation in which the mutation rate was changed from 0.3 to 0.01 over the course of the simulation run.

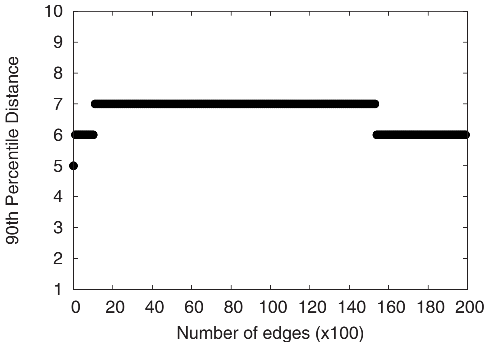

FIG. 11. Effective diameter of the network when the mutation rate decreases over time from 0.3 to 0.01. It increases quickly at first and then decreases slowly over time.

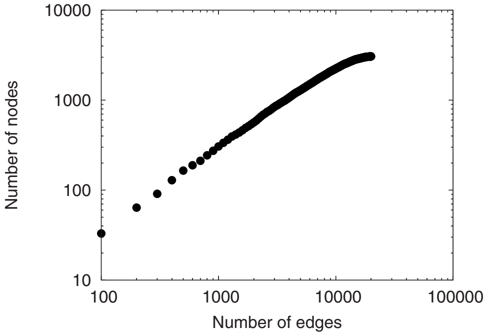

FIG. 12. Number of nodes grows as a power law with respect to the number of edges (or time, since one edge is added at each time step). The slope of the line is approximately 0.86.

VIII. CONCLUSIONS

We have shown that, when modeled appropriately, the preferential attachment model of network generation can be seen as a special case of the process of molecular evolution because they share a common underlying probabilistic model. We have presented a more general model of network generation, based on this underlying probabilistic model. Further, this model of network generation, which we call noisy preferential attachment, unifies the Erdős-Rényi random graph model with the preferential attachment model.

The preferential attachment algorithm assumes that the fitness of a node depends only on the number of links it has. This is not true of most real networks. On the worldwide web, for instance, the likelihood of linking to an existing web page depends also on the content of that web page. Some websites also experience sudden spurts of popularity, after which they may cease to acquire new links. Thus the probability of acquiring new links depends on more than the existing degree. This kind of behavior can be modeled by the noisy preferential attachment algorithm by including intrinsic fitness values for nodes.

The noisy preferential attachment algorithm can also be used to model some dynamical aspects of network growth such as densification power laws and shrinking diameters by gradually decreasing mutation rate over time. If true, this brings up the intriguing question of why mutation rate would decrease over time in real-world networks. On the worldwide web, for example, this may have to do with better-quality information being available through the emergence of improved search engines, etc. However, the fact that many different kinds of network exhibit densification and shrinking diameters suggests that there may be some deeper explanation to be found.

From a design point of view, intentional modulation of the mutation rate can provide a useful means of trading off between exploration and exploitation of network structure. We have been exploring this in the context of convergence in a population of artificial language learners [26].

The larger contribution of this work, however, is to bring together the fields of study of networks and evolutionary dynamics, and we believe that many further connections can be made.

ACKNOWLEDGMENTS

We appreciate the helpful comments of Roberto Aldunate and Jun Wang. The work was supported under NSF Grant No. IIS-0340996.

APPENDIX A

Here we solve the difference equation

$$x_{i,t+1} = \left(1 - \frac{i(p-q)}{n_0+t} - q\right)x_{i,t} + \left(\frac{(i-1)(p-q)}{n_0+t} + q\right)x_{i-1,t}. \tag{A1}$$

$x_{0,t}$ is a special case,

$$Nx_{0,t+1} = \begin{cases} Nx_{0,t} - 1 & \text{with probability } Nx_{0,t}q, \\ Nx_{0,t} & \text{otherwise}. \end{cases}$$

Expanding and simplifying as above, we get

$$x_{0,t+1} = (1-q)x_{0,t}.$$

The solution to this difference equation is simply

$$x_{0,t} = (1-q)^t x_{0,0}, \tag{A2}$$

where $x_{0,0}=(N-1)/N$ is the initial value of the number of empty urns. Note that here, and henceforth, we are assuming that initially all the urns are empty except for one, which has one ball in it. Therefore $x_{1,0}=1$ and $x_{i,0}=0\ \forall i>1$. This also means that $n_0=1$. These conditions together specify the entire initial state of the system.

Equation (A1) is difficult to solve directly, so we shall take the approach of finding the solution to $x_{1,t}$ and $x_{2,t}$ and then simply guessing the solution to $x_{i,t}$.

Substituting $i=1$ in Eq. (7) gives us

$$x_{1,t+1} = \left(1 - \frac{(p-q)}{n_0+t} - q\right)x_{1,t} + qx_{0,t}.$$

Substituting the solution for $x_{0,t}$ from Eq. (A2) gives us

$$x_{1,t+1} = \left(1 - \frac{(p-q)}{n_0+t} - q\right)x_{1,t} + q(1-q)^t x_{0,0}. \tag{A3}$$

The complete solution for $x_{1,t}$ is (see Appendix B),

$$x_{1,t} = (1-q)^t\left(A(t+1) + \frac{B}{t^r}\right), \tag{A4}$$

where $A=qx_{0,0}/(1+p-2q)$ and $B=2(p-q)/(1+p-2q)N\Gamma(1-r)$ are constants. Let us now use this result to derive the solution for $x_{2,t}$. Substituting $i=2$ in Eq. (A1), we get

$$x_{2,t+1} = \left(1 - \frac{2(p-q)}{n_0+t} - q\right)x_{2,t} + \left(\frac{p-q}{n_0+t} + q\right)x_{1,t}.$$

Substituting the solution for $x_{1,t}$ from Eq. (A4) and replacing n_0 by 1 for convenience gives us

$$x_{2,t+1} = \left(1 - \frac{2(p-q)}{1+t} - q\right)x_{2,t} + (1-q)^t\left(A(t+1) + \frac{B}{t^r}\right) \times \left(\frac{p-q}{1+t} + q\right). \tag{A5}$$

The solution to this (after some work) turns out to be (see Appendix B),

$$x_{2,t} = (1-q)^t\left(A(t+1)\frac{r}{2r+1} + \frac{B}{t^r} + \frac{D}{t^{2r}}\right) + \frac{q(1-q)^t}{1+p-2q}\left(A(t+1)\frac{2rt+t+2r}{2(2r+1)} + \frac{B}{t^r}(t+2)\right). \tag{A6}$$

In the above expression, compared to the first term, the remaining terms are negligible. To see this, consider that B/t^r can be at most B (as $r\to 0$), and at least B/t (as $r\to 1$). B itself is less than $1/N$. Therefore the contribution of the second term is upper bounded by $1/N$. A similar observation will hold for D/t^{2r}. This is far less than the contribution due to the first term, since A (which is also close to $1/N$) is multiplied by $(t+1)$. The remaining terms are approximately of the form t^2/N^2 (and higher i will contain higher powers). We can ignore these as long as $t \ll N$. Thus, we can write the solution for $x_{2,t}$ approximately as

$$x_{2,t} = \frac{Ar}{2r+1}(t+1)(1-q)^t = \frac{q}{1+p-2q}\frac{N-1}{N}(t+1)(1-q)^t = \frac{1-p}{N}\frac{r}{(r+1)(2r+1)}(t+1)(1-q)^{t-1}.$$

We can continue on with $x_{3,t}$:

$$x_{3,t+1} = \left(1 - \frac{3(p-q)}{1+t} - q\right)x_{3,t} + \left(\frac{2(p-q)}{1+t} + q\right)x_{2,t}.$$

If we follow through with this as for $x_{2,t}$, we will see the 2 from the constant in the second term $[2p/(t+1)]$ appear as a factor in the first term of the solution for $x_{3,t}$. In the general expression for the solution, this appears as $\Gamma(i)$. Therefore, we can guess the approximate expression for $x_{i,t}$ to be

$$x_{i,t} = \frac{1-p}{N}\frac{r^{i-1}\Gamma(i)}{\prod_{k=1}^{i}(kr+1)}(t+1)(1-q)^{t-1}, \tag{A7}$$

which is the same as Eq. (8).

APPENDIX B

Equation (A3) is

$$x_{1,t+1} = \left(1 - \frac{(p-q)}{n_0+t} - q\right)x_{1,t} + q(1-q)^t x_{0,0}.$$

This equation is of the form $y(t+1)=p(t)y(t)+r(t)$. The general form of the solution is

$$y(t) = u(t)\left(\sum \frac{r(t)}{Eu(t)} + C\right), \tag{B1}$$

where $u(t)$ is the solution of the homogeneous part of the above equation, i.e., $u(t+1)=p(t)u(t)$, and E is the time-shift operator, i.e., $Eu(t)=u(t+1)$. Now, the homogeneous part of Eq. (A3) is

$$\begin{aligned} u(t+1) &= \left(1 - q - \frac{p-q}{n_0+t}\right)u(t) \\ &= \left(\frac{(1-q)t + (1-q)n_0 - (p-q)}{n_0+t}\right)u(t) \\ &= (1-q)\left(\frac{t + n_0 - \dfrac{p-q}{1-q}}{t+n_0}\right)u(t). \end{aligned}$$

The solution to this difference equation is

$$u(t) = C(1-q)^t \frac{\Gamma(t+n_0-r)}{\Gamma(t+n_0)}, \tag{B2}$$

where $r=(p-q)/(1-q)$, C is a constant, and $\Gamma(\cdot)$ is the Gamma function, which is a "generalization" of the factorial to the complex plane. It is defined recursively as $\Gamma(n+1)=n\Gamma(n)$. The derivation of Eq. (B2) is given in Appendix C. From Eqs. (A3), (B1), and (B2), we get

$$\begin{aligned} x_{1,t} &= C(1-q)^t \frac{\Gamma(t+n_0-r)}{\Gamma(t+n_0)} \\ &\quad \times\left(\sum \frac{qx_{0,0}(1-q)^t\Gamma(t+1+n_0)}{C(1-q)^{t+1}\Gamma(t+1+n_0-r)} + D_1\right) \\ &= \frac{C(1-q)^t}{(t+n_0-1)^{\underline{r}}}\left(\frac{qx_{0,0}}{C(1-q)}\sum (t+n_0)^{\underline{r}} + D_1\right) \\ &\qquad (t^{\underline{r}} \text{ is read as ``} t \text{ to the } r \text{ falling''}) \\ &= \frac{q(1-q)^{t-1}x_{0,0}}{(t+n_0-1)^{\underline{r}}}\frac{(t+n_0)^{\underline{r+1}}}{r+1} + \frac{D(1-q)^t}{(t+n_0-1)^{\underline{r}}} \\ &\qquad (\text{where } D = CD_1 \text{ is another constant}) \\ &= \frac{q(1-q)^t x_{0,0}}{1+p-2q}\frac{\Gamma(t+n_0-r)}{\Gamma(t+n_0)}\frac{\Gamma(t+n_0+1)}{\Gamma(t+n_0-r)} + \frac{D(1-q)^t}{(t+n_0-1)^{\underline{r}}} \\ &= \frac{q(1-q)^t x_{0,0}(t+n_0)}{1+p-2q} + \frac{D(1-q)^t}{(t+n_0-1)^{\underline{r}}}. \end{aligned}$$

Let us evaluate the constant by applying the initial conditions $t=0$, $x_{0,0}=(N-1)/N$, $x_{1,0}=1/N$, and $n_0=1$. We get

$$\frac{1}{N} = \frac{q(N-1)/N}{1+p-2q} + D\Gamma(1-r)$$

$$1 = \frac{q(N-1)}{1+p-2q} + ND\Gamma(1-r).$$

Therefore,

$$D = \frac{2(p-q)}{(1+p-2q)N\Gamma(1-r)}. \tag{B3}$$

This gives us the complete solution for $x_{1,t}$ as

$$x_{1,t} = (1-q)^t\left(A(t+1) + \frac{B}{t^{\underline{r}}}\right),$$

where $A=qx_{0,0}/(1+p-2q)$ and $B=D=2(p-q)/(1+p-2q)N\Gamma(1-r)$ are constants. This is the same as Eq. (A4).

1. Solution to Eq. (A5)

Equation (A5) is

$$x_{2,t+1} = \left(1 - \frac{2(p-q)}{1+t} - q\right)x_{2,t} + (1-q)^t\left(A(t+1) + \frac{B}{t^{\underline{r}}}\right) \times\left(\frac{p-q}{1+t} + q\right).$$

Again, this equation is of the form of Eq. (B1). The solution to the homogeneous part in this case is

$$u(t) = C(1-q)^t \frac{\Gamma\left(t + 1 - \dfrac{2(p-q)}{1-q}\right)}{\Gamma(t+1)}. \tag{B4}$$

This is solved in exactly the same way as Eq. (B2) (see Appendix C). Now, from equations (B1), (A5), and (B4), we get,

$$\begin{aligned} x_{2,t} &= \frac{C(1-q)^t}{t^{\underline{2r}}}\left(\sum \frac{(1-q)^t\left(A(t+1) + \dfrac{B}{t^{\underline{r}}}\right)\left(\dfrac{p-q}{t+1} + q\right)}{C(1-q)^{t+1}\dfrac{1}{(t+1)^{\underline{2r}}}} + D_1\right) \\ &= \frac{C(1-q)^t}{t^{\underline{2r}}}\Bigg[\frac{1}{C(1-q)}\Bigg(A(p-q)\sum (t+1)^{\underline{2r}} + Aq\sum (t \\ &\quad + 1)(t+1)^{\underline{2r}} + B(p-q)\sum \frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}(t+1)} + Bq\sum \frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}}\Bigg) \\ &\quad + D_1\Bigg]. \end{aligned}$$

Solving the summations (see Appendix C), we get

$$\begin{aligned} x_{2,t} = \frac{C(1-q)^t}{t^{\underline{2r}}}\Bigg\{&\frac{1}{C(1-q)}\Bigg[\frac{A(p-q)(t+1)^{\underline{2r+1}}}{2r+1} \\ &+ Aq\left(\frac{t(t+1)^{\underline{2r+1}}}{2r+1} - \frac{(t+1)^{\underline{2r+2}}}{(2r+1)(2r+2)}\right) + B(p-q)\frac{t^{\underline{2r}}}{rt^{\underline{r}}} \\ &+ Bq\frac{(t+2)t^{\underline{2r}}}{(1+r)t^{\underline{r}}}\Bigg] + D\Bigg\}. \end{aligned}$$

Simplifying,

$$\begin{aligned}x_{2,t} &= (1-q)^t\left(\frac{Ar(t+1)}{2r+1}+\frac{Aq(t+1)(2rt+t+2r)}{(1-q)(2r+1)(2r+2)}\right.\\&\quad\left.+\frac{B}{t^{\underline{r}}}+\frac{Bq(t+2)}{(1-q)(1+r)t^{\underline{r}}}\right)+\frac{D(1-q)^t}{t^{\underline{2r}}}\\&=(1-q)^t\left(A(t+1)\frac{r}{2r+1}+\frac{B}{t^{\underline{r}}}+\frac{D}{t^{\underline{2r}}}\right)\\&\quad+\frac{q(1-q)^t}{1+p-2q}\left(A(t+1)\frac{2rt+t+2r}{2(2r+1)}+\frac{B}{t^{\underline{r}}}(t+2)\right).\end{aligned}$$

This is the same as Eq. (B5).

APPENDIX C

1. Derivation of Eq. (B2)

Equation (B2) is the solution to the following difference equation:

$$u(t+1)=(1-q)\left(\frac{t+n_0-\dfrac{p-q}{1-q}}{t+n_0}\right)u(t).$$

Note that all the factors in this equation are positive. Taking logarithms, we get

$$\log u(t+1)=\log\left[(1-q)\left(\frac{t+n_0-r}{t+n_0}\right)\right]+\log u(t),$$

$$\Delta\log u(t)=\log\left[(1-q)\left(\frac{t+n_0-r}{t+n_0}\right)\right],$$

$$\log u(t)=\sum[\log(1-q)+\log(t+n_0-r)-\log(t+n_0)]+D.$$

Remembering that $\Sigma a=ta$ and $\Sigma\log(t+a)=\log\Gamma(t+a)$, we get

$$\log u(t)=t\log(1-q)+\log\Gamma(t+n_0-r)-\log\Gamma(t+n_0)+D.$$

Therefore

$$u(t)=C(1-q)^t\frac{\Gamma(t+n_0-r)}{\Gamma(t+n_0)}.$$

This is the same as Eq. (B2).

2. Derivation of Eq. (B5)

Equation (B5) is the solution to the following difference equation:

$$\begin{aligned}x_{2,t}=\frac{C(1-q)^t}{t^{\underline{2r}}}&\left[\frac{1}{C(1-q)}\left(A(p-q)\sum(t+1)^{\underline{2r}}+Aq\right.\right.\\&\times\sum(t+1)(t+1)^{\underline{2r}}+B(p-q)\\&\left.\left.\times\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}(t+1)}+Bq\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}}\right)+D_1\right].\end{aligned}$$

We shall solve each of the summations individually. At several points, we will use the summation by parts formula,

$$\sum[Ey(t)\Delta z(t)]=y(t)z(t)-\sum[z(t)\Delta y(t)]. \qquad \text{(C1)}$$

The first summation term can be obtained directly:

$$\sum(t+1)^{\underline{2r}}=\frac{(t+1)^{\underline{2r+1}}}{2r+1}+C_1. \qquad \text{(C2)}$$

The second summation term can be obtained using the summation by parts formula. Let $Ey(t)=t+1$. Then $y(t)=t$, and $\Delta y(t)=1$. Let $\Delta z(t)=(t+1)^{\underline{2r}}$. Then $z(t)=(t+1)^{\underline{2r+1}}/(2r+1)$. We get

$$\sum(t+1)(t+1)^{\underline{2r}}=\frac{(t+1)(t+1)^{\underline{2r+1}}}{2r+1}-\sum\frac{(t+1)^{\underline{2r+1}}}{2r+1},$$

$$\sum(t+1)(t+1)^{\underline{2r}}=\frac{(t+1)(t+1)^{\underline{2r+1}}}{2r+1}-\frac{(t+1)^{\underline{2r+2}}}{(2r+1)(2r+2)}+C_2. \qquad \text{(C3)}$$

Before proceeding, we pause to calculate $\Sigma(1/t^{\underline{r}})$. Note that

$$\Delta\frac{1}{t^{\underline{r}}}=\frac{1}{(t+1)^{\underline{r}}}-\frac{1}{t^{\underline{r}}}=\frac{t+1-r}{(t+1)t^{\underline{r}}}-\frac{1}{t^{\underline{r}}}=\frac{-r}{(t+1)t^{\underline{r}}},$$

$$\frac{t+1}{-r}\Delta\frac{1}{t^{\underline{r}}}=\frac{1}{t^{\underline{r}}}.$$

Taking the summation, we get

$$\sum\frac{1}{t^{\underline{r}}}=\frac{1}{-r}\sum\left(Et\Delta\frac{1}{t^{\underline{r}}}\right).$$

Using the summation by parts formula, we get

$$\sum\frac{1}{t^{\underline{r}}}=\frac{1}{-r}\left(\frac{t}{t^{\underline{r}}}-\sum\frac{1}{t^{\underline{r}}}\right),$$

$$\left(1-\frac{1}{r}\right)\sum\frac{1}{t^{\underline{r}}}=\frac{-t}{rt^{\underline{r}}},$$

$$\sum\frac{1}{t^{\underline{r}}}=\frac{t}{(1-r)t^{\underline{r}}}. \qquad \text{(C4)}$$

We now proceed to the third summation term in the difference equation for $x_{2,t}$,

$$\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}(t+1)}=\sum\frac{t^{\underline{2r-1}}}{t^{\underline{r}}}.$$

We shall again use the summation by parts formula. Let $Ey(t)=t^{\underline{2r-1}}$. Therefore $y(t)=(t-1)^{\underline{2r-1}}$, and $\Delta y(t)=(2r-1)\times(t-1)^{\underline{2r-2}}$. Let $\Delta z(t)=1/t^{\underline{r}}$. Therefore $z(t)=t/(1-r)t^{\underline{r}}$ [from Eq. (C4)]. We get

$$\begin{aligned}\sum\frac{t^{\underline{2r-1}}}{t^{\underline{r}}}&=\frac{t(t-1)^{\underline{2r-1}}}{(1-r)t^{\underline{r}}}-\sum\frac{2r-1}{1-r}\frac{t(t-1)^{\underline{2r-2}}}{t^{\underline{r}}}\\&=\frac{t(t-1)^{\underline{2r-1}}}{(1-r)t^{\underline{r}}}-\frac{2r-1}{1-r}\sum\frac{t^{\underline{2r-1}}}{t^{\underline{r}}},\end{aligned}$$

$$\left(1+\frac{2r-1}{1-r}\right)\sum\frac{t^{\underline{2r-1}}}{t^{\underline{r}}}=\frac{t}{1-r}\frac{(t-1)^{\underline{2r-1}}}{t^{\underline{r}}}\sum\frac{t^{\underline{2r-1}}}{t^{\underline{r}}}=\frac{t^{\underline{2r}}}{rt^{\underline{r}}}.$$

Therefore,

$$\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}(t+1)}=\frac{t^{\underline{2r}}}{rt^{\underline{r}}}. \tag{C5}$$

The fourth summation term in the difference equation for $x_{2,t}$ is similar to the third one,

$$\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}}=\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}(t+1)}(t+1).$$

Let $Ey(t)=(t+1)$. Then $y(t)=t$ and $\Delta y(t)=1$. Let $\Delta z(t)=\Sigma[(t+1)^{\underline{2r}}/t^{\underline{r}}(t+1)]$. Then $z(t)=t^{\underline{2r}}/rt^{\underline{r}}$ [from Eq. (C5)]. Therefore, using the summation by parts rule, we get

$$\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}}=t\frac{t^{\underline{2r}}}{rt^{\underline{r}}}-\frac{1}{r}\sum\frac{t^{\underline{2r}}}{t^{\underline{r}}}. \tag{C6}$$

Now,

$$\sum\frac{t^{\underline{2r}}}{t^{\underline{r}}}=\sum\frac{(t+1-2r)t^{\underline{2r-1}}}{t^{\underline{r}}}=\frac{(t-2r)t^{\underline{2r}}}{rt^{\underline{r}}}-\frac{1}{r}\sum\frac{t^{\underline{2r}}}{t^{\underline{r}}}$$

$$=\frac{t-2r}{1+r}\frac{t^{\underline{2r}}}{t^{\underline{r}}}.$$

Substituting back in Eq. (C6), we get

$$\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}}=t\frac{t^{\underline{2r}}}{rt^{\underline{r}}}-\frac{1}{r}\left(\frac{t-2r}{1+r}\frac{t^{\underline{2r}}}{t^{\underline{r}}}\right)=\frac{t^{\underline{2r}}}{rt^{\underline{r}}}\left(t-\frac{t-2r}{1+r}\right).$$

Therefore, we have

$$\sum\frac{(t+1)^{\underline{2r}}}{t^{\underline{r}}}=\frac{(t+2)t^{\underline{2r}}}{(1+r)t^{\underline{r}}}. \tag{C7}$$

Combining Eqs. (C2), (C3), (C5), and (C7), we get the solution for $x_{2,t}$, i.e., Eq. (B5).

[1] S. Milgram, Psychol. Today **2**, 60 (1967).
[2] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[3] C. Asavathiratham, S. Roy, B. Lesieutre, and G. Verghese, IEEE Control Syst. **21**, 52 (2001).
[4] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[5] R. Ferrer i Cancho and R. V. Solé, Proc. R. Soc. London, Ser. B **268**, 2261 (2001).
[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[7] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
[8] H. A. Simon, Biometrika **42**, 425 (1955).
[9] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[10] K. M. Page and M. A. Nowak, J. Theor. Biol. **219**, 93 (2002).
[11] M. A. Nowak, Z. Phys. Chem. **216**, 5 (2002).
[12] M. Eigen and P. Schuster, Naturwiss. **64**, 541 (1977).
[13] M. Eigen, J. McCaskill, and P. Schuster, J. Phys. Chem. **92**, 6881 (1988).
[14] N. L. Komarova, J. Theor. Biol. **230**, 227 (2004).
[15] F. Chung, S. Handjani, and D. Jungreis, Ann. Comb. **7**, 141 (2003).
[16] M. Benaïm, S. Schreiber, and P. Tarrès, Ann. Appl. Probab. **14**, 1455 (2004).
[17] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[18] M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge University Press, Cambridge, U.K., 1983).
[19] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz, Phys. Rev. Lett. **89**, 258702 (2002).
[20] G. Bianconi and A.-L. Barabási, Europhys. Lett. **54**, 436 (2001).
[21] M. Eigen, J. McCaskill, and P. Schuster, Adv. Chem. Phys. **75**, 149 (1989).
[22] S. A. Golder and B. A. Huberman, J. Inf. Sci. **32**, 198 (2006).
[23] C. O. Wilke, C. Ronnewinkel, and T. Martinetz, Phys. Rep. **349**, 395 (2001).
[24] J. Leskovec, J. Kleinberg, and C. Faloutsos, in Proceedings of KDD'05, Chicago, 2005 (unpublished).
[25] B. Bollobás and O. Riordan, Combinatorica **24**, 5 (2004).
[26] S. Swarup and L. Gasser, in From Animals to Animats 9: Proceedings of the Ninth International Conference on the Simulation of Adaptive Behavior, Rome, 2006 (unpublished).
[27] G. U. Yule, Philos. Trans. R. Soc. London, Ser. B **213**, 21 (1925).