Evolving network with different edges

Jie Sun,[1,2] Yizhi Ge,[1,3] and Sheng Li[1,*]
[1]*Department of Physics, Shanghai Jiao Tong University, Shanghai, China*
[2]*Department of Mathematics and Computer Science, Clarkson University, Potsdam, New York 13699, USA*
[3]*Department of Physics, Georgetown University, Washington, DC 20057, USA*
(Received 13 November 2005; revised manuscript received 24 August 2007; published 12 October 2007)

We propose a scale-free network similar to Barabasi-Albert networks but with two different types of edges. This model is based on the idea that in many cases there are more than one kind of link in a network and when a new node enters the network both old nodes and different kinds of links compete to obtain it. The degree distribution of both the total degree and the degree of each type of edge is analyzed and found to be scale-free. Simulations are shown to confirm these results.

DOI: 10.1103/PhysRevE.76.046108 PACS number(s): 84.35.+i, 05.40.−a, 05.50.+q, 87.18.Sn

I. INTRODUCTION

Recently complex networks have attracted interest in many fields, includes biology, sociology, physics, etc. Contributions include the analysis of random networks with Erdos-Renyi graphs [1] and small world networks such as the Watts-Strogatz network [1–7]. However, real world networks such as the Internet, the movie actor network, and science collaboration networks (see [5,8]) all show a power-law degree distribution property, which cannot be explained by the two models mentioned. A scale free model [9,10] that displays this property was proposed by Barabasi and Albert (BA) in 1999. On the other hand, this model suffers a drawback in that the exponent of the power law is always fixed, while in real networks it varies. Further study has shown that preferential attachment plays a key role in the scaling properties of the evolving network [9–13]. In 2000, Dorogovtsev, Mendes, and Samukhin used a different initial attractiveness to vary the power-law exponent [13].

Social network models describe complex human interactions. To model the complicated relationships between elements of social networks, researchers introduced multiple type of vertices into network models [14]. Weighted edges representing varying kinds of social relationships were also introduced [15]. Both models show scaling properties. It is challenging to find more fundamental network structures.

Similarly, our model can be understood from another perspective. Considering the Internet as a network, edges may be partitioned based on what kind of connection they represent [e.g., hypertext transfer protocol (HTTP), file transfer protocol (FTP), etc.]. Generally, different types of links need to be added to growing networks, where competition exist between different links as well as the nodes. It is attractive to discover the topologies of the whole network and different subnetworks containing only one type of two edge, and to consider the relation between different edges. In the model we construct, the entire network is similar to the partial preference BA model. The significant difference is that we add some tunable parameters into the preferential attachment to alter the exponent of the power-law degree distribution and the shift constant of the degree distribution function.

Our work begins with the discussion of a scale-free network similar to the BA model with a change of the preferential attachment. This is a special example of the growing networks put forward by Dorogovstev, Mendes, and Samukhin in 2000 [13] in which they add a constant in the preferential attachment term. In our main work, the tunable parameters in the model play different roles in different edge preferential attachments, and are regarded as a weight on the attractiveness of different links in the evolving network. The research produces some interesting results: the whole network and the two subnetworks all evolve just the same as the ordinary scale-free network, while the two kinds of links connected to a certain node always differ. The degree distributions of the two subnetworks are unequal.

II. MODEL AND THEORETICAL APPROACH

Before introducing our findings, we first review the work by Dorogovstev, Mendes, and Samukhin (DMS) in 2000 [13]. In their paper, the preferential attachment in the BA model is no longer fixed; instead, a changeable initial attractiveness, defined at each site, together with the incoming links of the site, determines the probability whether a new link will point to this site. In the long-time limit, the exponent of the degree distribution varies from 2 to ∞, depending on the initial attractiveness. In this paper, we will make a simple and clear revision, and attain concise results.

The preferential attachment in the BA model is changed such that the possibility of a site's degree increasing is now proportional to $\Pi(k_i)=m\frac{k_i+f}{\Sigma(k_i+f)}$, where k_i is the degree of site i while f is a given constant. This model may be considered as a special case of the model in [13] if one takes their $A^{(0)}$ as $m+f$. DMS found the degree distribution to have the exponent $-(2+a)$, where $a=\frac{A^{(0)}}{m}$. A linear shift $ma=A^{(0)}$ is found in the distribution function

$$P(q) \approx (1+a)\frac{\Gamma[(m+1)a+1]}{\Gamma(ma)}(q+ma)^{-(2+a)}. \qquad (1)$$

For our model

*lisheng@sjtu.edu.cn

$$\frac{\partial k_i}{\partial t} = m\frac{k_i + f}{\sum (k_i + f)}, \tag{2}$$

where m is the number of links added from each new node. We solve this system of differential equations for each k_i by simple separation of variables. Then by asymptotic approximation of that solution, as t approaches infinity, we can conclude the degree distribution,

$$P(k) = \left(2 + \frac{f}{m}\right)(m+f)^{(2+f/m)}(f+k)^{-(3+f/m)}. \tag{3}$$

Compare the degree functions (1) and (3), taking the power term into account, and we find that the two results are the same in essence. In our derivation the tunable parameter is f, while in [13] it is the initial attractiveness $A^{(0)}$. By transforming $A^{(0)}$ to $m+f$ the two functions show the same property. This transformation is reasonable; $A^{(0)}$ should be the initial degree a site has when it enters the network, which is exactly $m+f$. In [13], the exponent $\gamma=2+a$ varies from 2 to ∞, while in our model, $\gamma=3+\frac{f}{m}$, $-m< f<\infty$ gives the same range. The advantage in using continuum theory is that no special function appears. Another difference between our model and the model in [13] is that we consider the network as undirected while in [13] networks are considered as directed.

We now propose here the details of our network model. Let our graph have two kinds of edges, as follows: at each time interval, a new node is added to the network connecting to m existing nodes. We divide the m edges into two types, X and Y. We propose that node i in the original network connects to the new node with preferential attachments:

$$\frac{\partial x_i}{\partial t} = m\frac{x_i + y_i + f + g y_i}{2\left(\sum x_j + y_j + f\right)}, \tag{4}$$

$$\frac{\partial y_i}{\partial t} = m\frac{x_i + y_i + f - g y_i}{2\left(\sum x_j + y_j + f\right)}, \tag{5}$$

where x_i represents the number of X edges node i has connected (at time t) and y_i represents the number of Y edges, with $x_i+y_i=k_i$; we do not assume this to be exactly the reflection of real world networks, but just a step forward in the direction of finding out complex relations between networks that share the same nodes but different types of edges.

Here in our model both nodes and different kinds of edges compete. Although any newly added node would have a fixed number of m new adding links, it is the parameters in our equations above that decide the portion of one kind of links versus the other. This idea represents the fact that in real life networks, even though the strong nodes tend to have more and more links, if we take a look at some other aspect and some other kind of links that strong guy might not be that strong since any one would only have a finite and limited effect. More emphasis on one kind of link would decrease the focus on the other and thus reduce the number of links of the other kind.

The parameter g here plays the key role in determining the relation between X and Y. When g is very small, X and Y behave the same. When g becomes relatively large, new edges will always be added as X edges.

First, we discuss the Y subnetwork. We add m edges at every time step, so

$$\sum (x_i + y_i + f) = (2m+f)t. \tag{6}$$

Substituting this into Eq. (5) yields

$$\frac{\partial y_i}{\partial t} = m\frac{(m+f)\left(\frac{t}{t_i}\right)^{m/(2m+f)} - g y_i}{2(2m+f)t}. \tag{7}$$

Then

$$y_i = \frac{m+f}{2+g}\left(\frac{t}{t_i}\right)^{m/(2m+f)} + \text{const}\; t^{-[mg/2(2m+f)]}. \tag{8}$$

When t approaches infinity, this becomes

$$y_i \approx \frac{m+f}{2+g}\left(\frac{t}{t_i}\right)^{m/(2m+f)}. \tag{9}$$

Then we obtain

$$P_y(y_i < y) = P\left[t_i > y^{-(2+f/m)}\left(\frac{m+f}{2+g}\right)^{(2+f/m)} t\right] = 1 - \frac{1}{m_0 + t} y^{-(2+f/m)}\left(\frac{m+f}{2+g}\right)^{(2+f/m)} t. \tag{10}$$

For large t, $P(y_i)$ is approximated by

$$P_y(y) = \frac{\partial P(y_i < y)}{\partial y} \approx \left(2 + \frac{f}{m}\right)\left(\frac{m+f}{2+g}\right)^{(2+f/m)} y^{-(3+f/m)}. \tag{11}$$

The degree distribution of Y is a power law with the same exponent as that of total degree distribution but without the shift f. Then, for the X subnetwork we obtain

$$x_i = k_i - y_i = \frac{1+g}{2+g}(m+f)\left(\frac{t}{t_i}\right)^{(m/2m+f)} - f - \text{const}\; t^{-[mg/2(2m+f)]}. \tag{12}$$

In the long-time limit, we neglect the $t^{-[mg/2(2m+f)]}$ term, and obtain

$$x_i = \frac{1+g}{2+g}(m+f)\left(\frac{t}{t_i}\right)^{[m/(2m+f)]} - f. \tag{13}$$

The degree distribution of the X subnetwork is

$$P_x(x) = \left(2 + \frac{f}{m}\right)\left(\frac{1+g}{2+g}(m+f)\right)^{(2+f/m)} (x+f)^{-(3+f/m)}. \tag{14}$$

It is a power-law distribution *with* shift f. From the calculations above, we obtain the ratios of total degree to x degree and y degree

$$P_{x+y}(k-f):P_x(k-f):P_y(k) = 1:\left(\frac{1+g}{2+g}\right)^{2+f/m}:\left(\frac{1}{2+g}\right)^{2+f/m}. \tag{15}$$

The subscript $x+y$ in $P_{x+y}(k)$ means the total degree.

III. NUMERICAL SIMULATIONS

Figures 1(a)–1(c) show the results of simulations of networks with $N=10^6$ nodes, $m=5$, and $f=-2,2,5$. The simulated exponents agree with the theoretical results. The predicted ratio of x degree to y degree is also confirmed by numerical simulations (see Fig. 2). One should notice that in the figures, we have made shifts for x degree distributions and total degree distributions from $P_x(k)$ and $P_{x+y}(k)$ to $P_x(k-f)$ and $P_{x+y}(k-f)$, respectively. Therefore, their log-log plots behave as straight lines parallel to those of y degree.

However, we found that when we take g to be small, the simulations differ from the theoretical predictions. With small g, when f is positive, the x degree is larger than expected while the y degree matches the predicted value [see Fig. 3(a)]; and when f is negative, the y degree is larger than expected while the x degree is as expected [see Fig. 3(b)]. We neglected the last term of Eq. (8), but when g is small, decay of this term is slow. This explains the discrepancy at small g.

We also looked at the fluctuation of x degree x_k to theoretical value as a function of total degree k. It illustrates how the competition generates heterogeneity in edge composition of each vertex. We calculated the relative standard deviation compared to theory: the fluctuations decrease quickly with the increase of total degree (see Fig. 4).

IV. STUDY OF PARAMETERS

The parameter m is treated as a fixed number in a certain network, for we can identify at least vaguely how many links are added in each time interval. The value of f varies from $-m$ to ∞, but if f is too large both the total degree distribution and the X degree distribution approach exponential,

$$(k+f)^{-(3+f/m)} \approx f^{-(3+f/m)}e^{-(3+f/m)(k/f)} = f^{-(3+f/m)}e^{-k(1/m+3/f)}. \tag{16}$$

If $f \gg m$, the characteristic degree of total network and X subnetwork is m. In addition, the $t^{-[mg/2(2m+f)]}$ term will not decay fast enough to be neglected if g, f, and m are chosen such that $mg/2(2m+f)$ is too small.

The most intriguing parameter is g. As mentioned, g should not be chosen to make the $t^{-[mg/2(2m+f)]}$ term too large. Simulations have shown that small g makes the degree distribution depart from predictions. g should not be too large, as well; a very large g leads the network to have few Y links, as we have found in simulations. Based on these above reasons, we need g to be big enough to make the $t^{-[mg/2(2m+f)]}$ term small enough to be neglected; but g should not be so large as to make $\frac{\partial x_i}{\partial t}=m\frac{x_i+y_i+f+gy_i}{2(\Sigma x_j+y_j+f)}$ larger than $\frac{\partial y_i}{\partial t}=m\frac{x_i+y_i+f-gy_i}{2(\Sigma x_j+y_j+f)}$, which leads to a network having few Y edges.

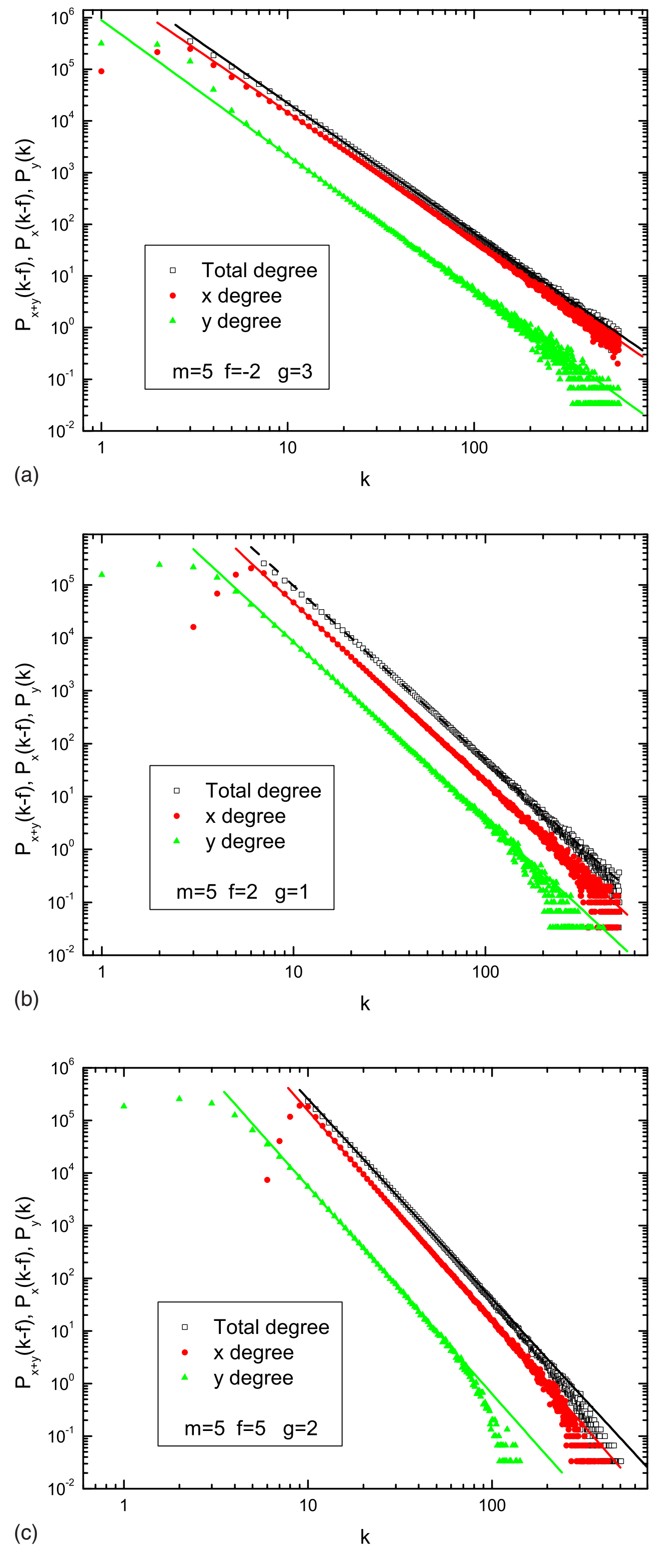

FIG. 1. (Color online) Simulations of the distribution of total degree, X degree, and Y degree, for networks of 10^6 nodes and $m=5$. The lines are linear fits of the main part of the data. (a) $f=-2$, $g=3$. The slopes of the lines are $\gamma_k=-2.5$, $\gamma_x=-2.5$, and $\gamma_y=-2.6$. The prediction is $\gamma=-2.6$. (b) $f=2$, $g=1$. The slopes of the lines are $\gamma_k=-3.3$, $\gamma_x=-3.4$, and $\gamma_y=-3.4$. The prediction is $\gamma=-3.4$. (c) $f=5$, $g=2$. The slopes of the lines are $\gamma_k=-3.8$, $\gamma_x=-4.0$, and $\gamma_y=-3.9$. The prediction is $\gamma=-4$.

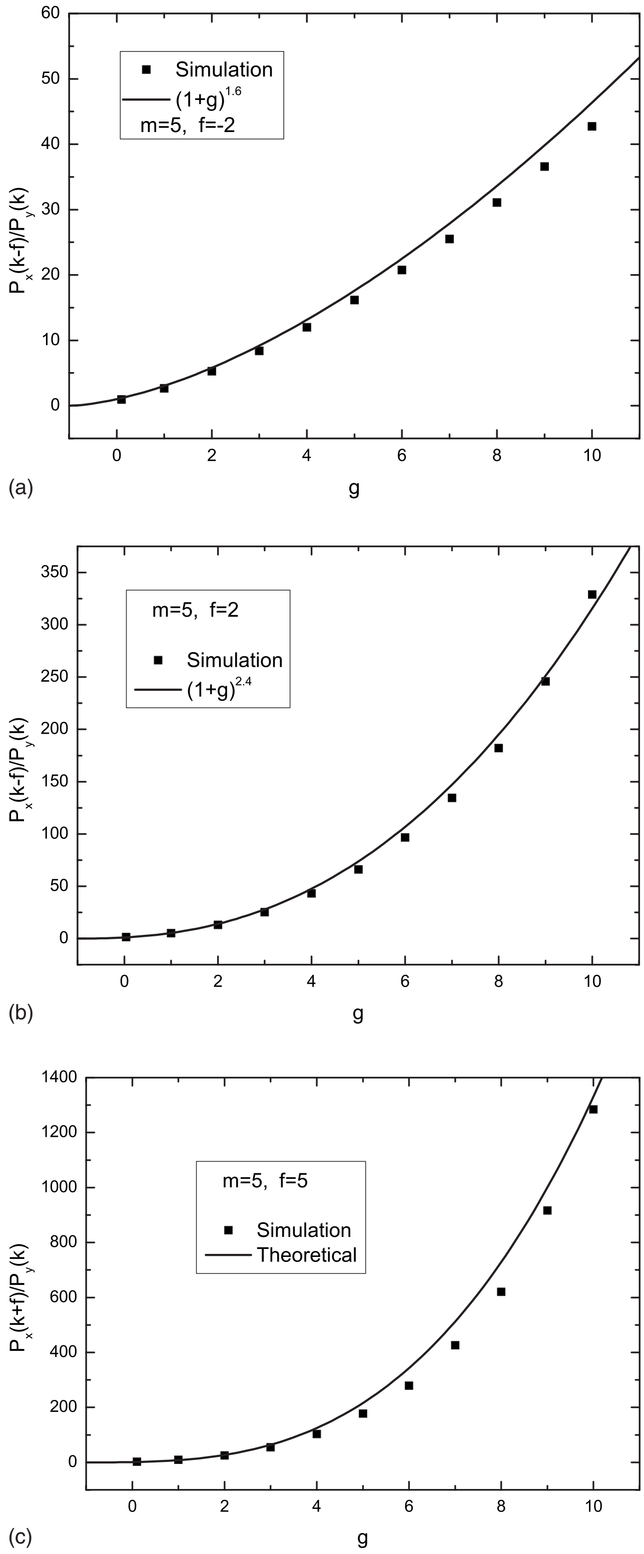

FIG. 2. The ratios of the distribution of X degree to the distribution of Y degree with respect to different g. The curves are theoretical predictions. $N=10^6$ nodes and $m=5$. (a) $f=-2$; (b) $f=2$; (c) $f=5$.

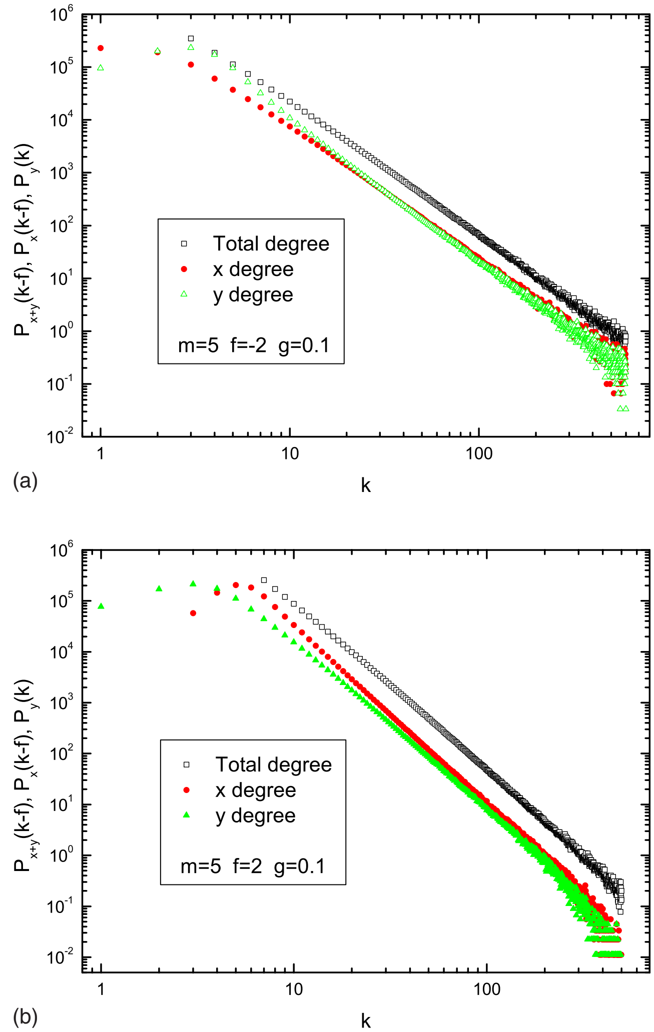

FIG. 3. (Color online) The degree distribution of total degree, x degree, and y degree with small $g=0.1$, for networks of 10^6 nodes and $m=5$. (a) $f=-2$, (b) $f=2$.

For a given node with degree k at given time, the probability $p_x(k)$ that how many X edges it has follows Eq. (4). Figure 4 shows the fluctuation decreases quickly with k. It seems that the distribution of fluctuation should follow the binomial distribution. However, here the degree of a node is varying with time. It increases from m to k and p_x varied synchronously. This leads to the fact that the fluctuation of the number of X edges is much smaller than that of the binomial distribution.

V. DISCUSSIONS

The mathematical expressions of degree distribution of the subnetworks provide deep insight into the dynamics of evolving systems. We build a competitive environment where not only nodes but also different types of edges compete. This model can reflect many properties of social network. A newly added node has a fixed number of m initial edges. However, it is the other nodes that decided how many X edges and Y edges there would be. Obviously, other nodes compete for these m edges based on what they already have. For instance, let X edge denotes financial relationship between individual persons, and Y represents other connection.

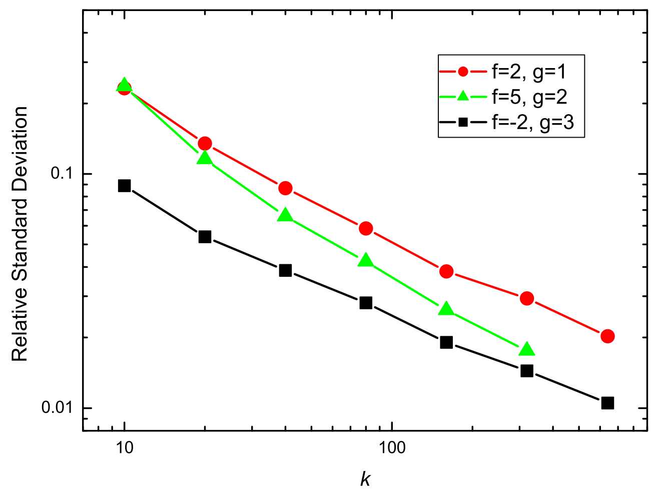

FIG. 4. (Color online) The relative standard deviation of X degree respective to total degree, for networks of 10^6 nodes and $m=5$.

Rich people tend to have more financial relationship with other people, while interestingly, more financial links signifies the richness. But when we focus on other links between people, rich people are not necessarily so lucky. Due to the limitation of personal capability, time, and devotion, one cannot have infinite connections with others. Therefore, it is at the expense of less Y edges to obtain more X edges, and vice versa. The above study shows the relationship between X and Y edges.

An important characteristic of this model is that the degree distribution of the X subnetwork shows a linear shift while the Y subnetwork does not. Simulations confirm this, although the exact value of the shift may has some errors, due to the application of mean-field approximation. This difference between the X and Y subnetwork comes from the different preferential attachment.

These results may be generalized to directed networks. Every time step, we introduce m directed edges. Then

$$\sum (k_i + f) = (m + f)t. \tag{17}$$

Similarly to the undirected result, we obtain

$$P_{x+y}(k) = \left(1 + \frac{f}{m}\right) f^{1+f/m} (k+f)^{-(2+f/m)}. \tag{18}$$

The degree distributions of the X network and Y network are

$$P_x(k) = \left(1 + \frac{f}{m}\right)\left(\frac{f(2+g) - 1}{2+g}\right)^{(1+f/m)} (k+f)^{-(2+f/m)}, \tag{19}$$

$$P_y(k) = \left(1 + \frac{f}{m}\right)\left(\frac{1}{2+g}\right)^{(1+f/m)} k^{-(2+f/m)}. \tag{20}$$

The ratios of total degree to X degree and Y degree are

$$P_{x+y}(k-f):P_x(k-f):P_y(k) = 1:\left(\frac{f(2+g)-1}{f(2+g)}\right)^{1+f/m}:\left(\frac{1}{f(2+g)}\right)^{1+f/m}. \tag{21}$$

Compared with the results of undirected network (15), the proportion of Y edges is larger than in the undirected network.

ACKNOWLEDGMENTS

The authors would like to thank Erik M. Bollt for discussion and comments and James P. Bagrow for feedback. This paper was supported by the National Science Foundation of China under Grant No. 10105007, and PRP Foundation of Shanghai Jiao Tong University Grant No. T0720701.

[1] P. Erdos and A. Renyi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
[2] B. Derrida and Y. Pomeau, Europhys. Lett. **1**, 45 (1986).
[3] B. Derrida and D. Stauffer, Europhys. Lett. **2**, 739 (1987).
[4] R. Monasson and R. Zecchina, Phys. Rev. Lett. **75**, 2432 (1995).
[5] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[6] D. J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton University Press, Princeton, NJ, 1999).
[7] M. E. J. Newman and D. J. Watts, Phys. Lett. A **263**, 341 (1999).
[8] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[9] A-L. Barabasi and R. Albert, Science **286**, 509 (1999).
[10] A-L. Barabasi, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).
[11] S. N. Dorogovtsev and J. F. F. Mendes, Phys. Rev. E **62**, 1842 (2000).
[12] S. N. Dorogovtsev and J. F. F. Mendes, Europhys. Lett. **52**, 33 (2000).
[13] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[14] D.-H. Kim, B. Kahng, and D. Kim, Eur. Phys. J. B **38**, 305 (2004).
[15] M. Boguna, R. Pastor-Satorras, A. Diaz-Guilera, and A. Arenas, Phys. Rev. E **70**, 056122 (2004).