Power-law distributions from additive preferential redistributions

Suhan Ree*
Department of Industrial Information, Kongju National University, Yesan-Up, Yesan-Gun, Chungnam, 340-702, South Korea
(Received 8 September 2005; published 15 February 2006)

We introduce a nongrowth model that generates the power-law distribution with the Zipf exponent. There are N elements, each of which is characterized by a quantity, and at each time step these quantities are redistributed through binary random interactions with a simple additive preferential rule, while the sum of quantities is conserved. The situation described by this model is similar to those of closed N-particle systems when conservative two-body collisions are only allowed. We obtain stationary distributions of these quantities both analytically and numerically while varying parameters of the model, and find that the model exhibits the scaling behavior for some parameter ranges. Unlike well-known growth models, this alternative mechanism generates the power-law distribution when the growth is not expected and the dynamics of the system is based on interactions between elements. This model can be applied to some examples such as personal wealths, city sizes, and the generation of scale-free networks when only rewiring is allowed.

DOI: 10.1103/PhysRevE.73.026115 PACS number(s): 89.75.Da, 89.65.Gh, 89.75.Hc

I. INTRODUCTION

Power-law distributions have been observed in diverse fields for more than a century [1]. Some well-known examples exhibiting "scaling" behavior are city sizes [2–4], word frequencies [5], sizes of business firms [6], personal incomes [7,8], personal wealths [9–11], sizes of web sites [12], numbers of links of web pages [13], connections of routers in the Internet [14], species in genera [15], interactions of proteins [16], citations of scientific papers [17], and so on, covering many research fields such as biology, economics, sociology, engineering, and physics. Many generative models have been introduced so far to explain this ubiquitous phenomenon [18], and most of them use simple mechanisms that give rise to the power-law distributions. One group of models uses stochastic multiplicative processes [3,7,19], and another group uses preferential growing mechanisms [5,15,20]. These models have their root in the Gibrat's law of proportional growth [6], and are based on two assumptions: the growth of the system and the noninteraction between elements. There are also nongrowth models in which the main mechanism is the interaction between randomly chosen elements, resulting in the multiplicative changes of values [9,10,21,22]. Systems showing the scaling behavior consist of N elements (N may vary with time), while each element i ($1 \leqslant i \leqslant N$) is represented by the quantity k_i, and the probability of an element having the value k, $P(k)$, has the form $k^{-\gamma}$ for a given range of k.

Here we introduce a nongrowth model exhibiting the power-law distribution with the Zipf exponent ($\gamma=2$), in which quantities of elements are redistributed through binary random interactions with a simple additive preferential rule. The model assumes that N and the sum of all k_i's are conserved, and that, when two elements i and j are chosen randomly at a given time, k_i and k_j will be changed additively while preserving k_i+k_j. This model can be a mechanism that explains scaling behavior of many socioeconomical systems, especially when the growth is not expected and interactions between elements are vital to their dynamics. Moreover, this model can be extended to generate scale-free networks through rewiring only, because the rewiring process by changing an end point of a link changes the degrees of two nodes additively while preserving the sum of the degrees of all nodes.

In this paper, the model and its stationary distributions are investigated both numerically and analytically. In Sec. II, the model is described in detail. In Sec. III, the master equation is obtained. Stationary distributions are found numerically first, and then analytically solved. And the condition for the power-law distributions in the parameter space is also found using both numerical and analytic methods. In Sec. IV, three possible applications of this model are discussed. Finally, in Sec. V, we summarize our results.

II. MODEL

Let us introduce our stochastic model in detail. The model assumes that k_i's are non-negative integers, and we define α as the average quantity per element,

$$\alpha \equiv \sum_{k=0}^{\infty} kP(k) = \langle k \rangle. \tag{1}$$

At each time step T, two elements i and j are randomly chosen, and the element i gives one unit of the quantity to the element j with the *exchange probability* R; hence their quantities are changed additively, $k_i \to k_i - 1$ and $k_j \to k_j + 1$, while $k_i + k_j$ is conserved (as a result, α becomes a conserved quantity). In other words, i is the giver and j is the taker, while the probability of nonexchange is $1-R$. One simple way to give an advantage to an element with bigger k is letting R dependent on k_i and k_j as below,

*Electronic address: suhan@kongju.ac.kr

$$R = \begin{cases} 1 & (0 < k_i \leqslant k_j), \\ \beta & (k_i > k_j), \\ 0 & (k_i = 0), \end{cases} \tag{2}$$

where β is a constant in the range of $0 \leqslant \beta \leqslant 1$. In this model, we can represent the system with three independent parameters: N, α, and β. When a distribution is given initially at $T=0$, $P(k)$ will evolve as T increases, and eventually reach a stationary distribution. To express the cumulative distribution, we also define $P(\geqslant k) \equiv \Sigma_{k'=k}^{\infty} P(k')$.

The parameter β plays an important role in this model. Two special cases of $\beta=0$ and $\beta=1$ have been previously discussed in the context of conserved exchange processes [9,10,23]. The focus of this paper, however, is the general case of $0<\beta<1$. When $\beta<1$, the time-reversal symmetry of the dynamics is broken, and at the same time the elements with bigger k ("the rich") get an advantage over those with smaller k ("the poor"). Then this model becomes one of rich-get-richer mechanisms, which will generate broad stationary distributions. In the next section, we will show that the stationary distribution from this model exhibits the power law when β is less than a certain critical value, and that this critical value will depend on the value of α.

III. STATIONARY DISTRIBUTIONS

First we look at the dynamics of an element with the quantity k to focus on the evolution of an element. To gain one unit, an element should be chosen as the taker with the probability $1/N$, and the probability of an element gaining one unit, $T_+(k)$, depends on the choice of the giver. Similarly the probability of an element losing one unit, $T_-(k)$, when chosen as a giver, can be found

$$T_+(k) = [1 - P(0) - (1-\beta)P(\geqslant k+1)],$$

$$T_-(k) = (1-\delta_{k0})[\beta + (1-\beta)P(\geqslant k)]. \tag{3}$$

Then, for an element, the expected change of k after a time step, Δk, is $[T_+(k) - T_-(k)]/N$. Since Δk is not proportional to k, the Gibrat's law is not satisfied. In a sense, each element is performing the random walk if we regard k as the position, while the transition probability found in Eq. (3) is asymmetrical, position-dependent, and time varying.

If we use the continuous approximation as $N \to \infty$, the master equation can be obtained,

$$\Delta P(k) = [P(k-1)T_+(k-1) - P(k)T_-(k)] - [P(k)T_+(k) - P(k+1)T_-(k+1)], \tag{4}$$

where $\Delta P(k)$ is the expected change of $P(k)$ after one time step. Then from the condition for the stationary distribution, $\Delta P(k)=0$ $(\forall k, k \geqslant 0)$, we find that stationary distributions should satisfy these nonlinear equations

$$P(k+1) = \frac{T_+(k)}{T_-(k+1)} P(k) = \frac{1 - P(0) - (1-\beta)P(\geqslant k+1)}{\beta + (1-\beta)P(\geqslant k+1)} P(k), \tag{5}$$

for $k \geqslant 0$, because, in Eq. (4), there are two parts, two terms each, and each part should be zero when $\Delta P(k)=0$. Even though we can theoretically find $P(k)$ as a function of α and β using Eqs. (1) and (5), these nonlinear equations are not easily solved analytically except for some special cases.

A. Case of $\beta=0$

This is a trivial winner-take-all situation. When $\beta=0$, the rich will always win for every binary interaction. Even without solving Eq. (5), the stationary state and its distribution are trivially found. In the stationary state, one element has all quantities, $k=\alpha N$, and the other elements have no quantity, $k=0$; therefore the stationary distribution is

$$P(k) = \frac{N-1}{N}\delta_{k0} + \frac{1}{N}\delta_{k,\alpha N}. \tag{6}$$

As $N \to \infty$, $P(k)$ becomes δ_{k0} approximately.

B. Case of $\beta=1$

When $\beta=1$, the model describes the conserved random exchange process, which was already discussed previously [9,23]. From Eq. (5), we easily find $P(k)$ exactly as $[1-P(0)]^k P(0)$. After substituting $P(k)$ into Eq. (1) to find $P(0)$, we obtain the stationary distribution,

$$P(k) = \frac{1}{1+\alpha}\left(\frac{\alpha}{1+\alpha}\right)^k. \tag{7}$$

As $\alpha \to \infty$, $P(k)$ becomes $(1/\alpha)\exp[-k/\alpha]$, which is the Boltzmann-Gibbs distribution.

C. Case of $0<\beta<1$

In this general case, the rich have an advantage over the poor, but lose to the poor from time to time. This property keeps the stationary distribution balanced somewhere between those from two extreme cases discussed above. Since the analytic method cannot be solely used in this case, the model is numerically investigated first.

After performing extensive numerical simulations while varying N, α, and β, we found that the stationary distributions exhibit the power law when α and β satisfy a certain condition; that is, it is scaling when (α,β) is inside a region in (α,β)-space, represented by a condition such as $f(\alpha,\beta)<\epsilon$, where $f(\alpha,\beta)$ and ϵ will later be found in this section. In Fig. 1, we show a scaling case of $N=10^5$, $\alpha=1$, and $\beta=0.1$. As time increases, the initial distribution, δ_{k1}, evolves to a power-law stationary distribution, which is shown using the cumulative distribution, $P(\geqslant k)$. (In all simulations here, α is a positive integer, and the initial distributions of the δ-function form, $\delta_{k\alpha}$, will be used.)

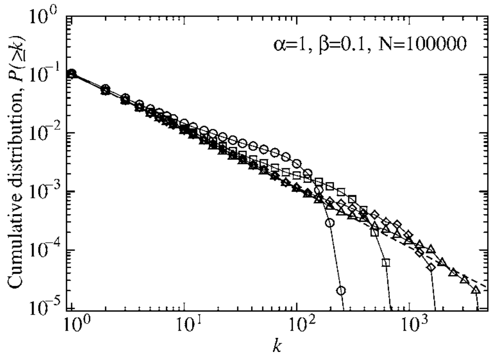

FIG. 1. Evolution of the cumulative distribution $P(\geqslant k)$ when $N=10^5$, $\alpha=1$, and $\beta=0.1$ in a log-log plot. From an initial distribution $P(k)=\delta_{k1}$, we observe how $P(\geqslant k)$ evolves as the number of timesteps T varies from 0 to 10^9 (○), 10^{10} (□), 10^{11} (◇), 10^{12} (△). We can observe that $P(\geqslant 1)\simeq\beta(=0.1)$, which leads us to $P(0)=1-P(\geqslant 1)\simeq 1-\beta$. The dashed line represents the theoretical stationary distribution $P(\geqslant k)=1/(9k+1)$ at $\beta=0.1$.

One common property that stands out in all scaling cases like the one in Fig. 1 is that $P(0)\simeq 1-\beta$ whenever the distribution is scaling. This property can be analytically proven by solving Eq. (5) when $P(0)$ is given as $1-\beta$. Since $P(\geqslant k+1)=1-\Sigma_{k'=0}^{k}P(k')$, $P(k+1)$ can be represented as a function of $P(0),\ldots, P(k)$, and β. When $k=0$, $P(1)$ can be found as a function of $P(0)$ and β, and when $k=1$, $P(2)$ can also be found as a function of $P(0)$ and β using $P(1)$ obtained already. If we repeat this process, $\{P(k)|k\geqslant 1\}$ will all be found as a function of $P(0)$ and β. When we substitute $P(0)=1-\beta$, found numerically in scaling cases, we can obtain $P(k)$ and $P(\geqslant k)$ in closed forms as below

$$P(k)=\frac{\beta}{1-\beta}\frac{1}{[k+1/(1-\beta)][k+\beta/(1-\beta)]},$$

$$P(\geqslant k)=\frac{\beta}{1-\beta}\frac{1}{k+\beta/(1-\beta)}. \tag{8}$$

This is the Zipf's law, $P(k)\propto k^{-2}$ and $P(\geqslant k)\propto k^{-1}$, valid when $k(1-\beta)$ is big enough. The shape of $P(k)$ in Eq. (8) does not depend on α, but as we will show later α will play a significant role in deciding whether the system is scaling or not.

As shown above, the relation $P(0)=1-\beta$ is the condition for the scaling stationary distributions. In other words, a scaling condition for α and β can be found if we find a condition with which the condition $P(0)=1-\beta$ holds. To observe when the relation $P(0)=1-\beta$ holds, we find $P(0)$ for various α and β values using numerical simulations. In Fig. 2, we show $P(0)$ versus α, and $P(0)$ versus β when $N=10^4$. When β is given, $P(0)$ equals to $1-\beta$ when α is greater than a certain critical value, $\alpha_c(\beta)$, and when α is given, $P(0)$ equals to $1-\beta$ when β is less than a certain critical value, $\beta_c(\alpha)$ [$\beta=\beta_c(\alpha)$ is the inverse functions of $\alpha=\alpha_c(\beta)$]. Therefore we find that a critical relation exists for the system to exhibit the scaling behavior, and the boundary between the scaling region and nonscaling region is represented by $\alpha=\alpha_c(\beta)$.

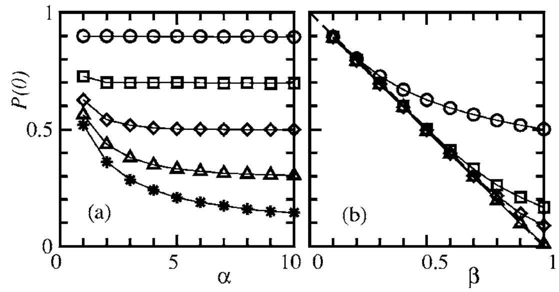

FIG. 2. When $N=10^4$ and $T=10^9-10^{10}$, $P(0)$ values are found numerically for various α and β values (averaged over ten runs). (a) $P(0)$ versus α when $\beta=0.1$ (○), 0.3 (□), 0.5 (◇), 0.7 (△), 0.9 (*). $P(0)$ is close to $1-\beta$ when α is greater than a certain value for a given β. (b) $P(0)$ when β for $\alpha=1$ (○), 5 (□), 10 (◇), 100 (△). $P(0)$ is close to $1-\beta$ when β is less than a certain value for a given α. Moreover, we observe that $P(0)=1/(1+\alpha)$ from Eq. (7) are satisfied when $\beta=1$. The dashed line represents $P(0)=1-\beta$.

How do we estimate this critical boundary in (α,β)-space analytically? One possible argument uses the highest k value, k_M. Since N is finite in the model, the power law will be valid only for a finite range of k, and the position of cutoff, k_M, depends on N, α, and β. Especially when Eq. (8) is satisfied (scaling cases), we can estimate k_M by solving the equation below,

$$\alpha=\sum_{k=0}^{k_M}kP(k) \simeq \frac{\beta}{1-\beta}\int_0^{k_M}dk\left[\frac{k}{[k+1/(1-\beta)][k+\beta/(1-\beta)]}\right], \tag{9}$$

which is a modification of Eq. (1) by letting $P(k)=0$ when $k>k_M$ [reasonable because $P(k)$ obtained from the continuous approximation is not valid when N is finite and k is high]. By solving Eq. (9), the estimated value of k_M for scaling cases is

$$k_M\simeq\frac{\beta}{1-\beta}\beta^{1/(1-\beta)}\exp\left[\frac{\alpha}{\beta/(1-\beta)}\right]. \tag{10}$$

Then we can find the ratio of the number of elements that are supposed to be in the region $k>k_M$ to the total number of elements N, which can be obtained from $P(\geqslant k)$ at $k=k_M$,

$$P(\geqslant k_M)\simeq\frac{\beta}{1-\beta}\frac{1}{k_M}\simeq\beta^{1/(1-\beta)}\exp\left[\frac{-\alpha}{\beta/(1-\beta)}\right]\equiv f(\alpha,\beta). \tag{11}$$

If the ratio $f(\alpha,\beta)$ is small enough, these elements that were supposed to be in $k>k_M$ can be regarded as additional elements with $k=0$, changing $P(0)$ into $P(0)+f(\alpha,\beta)$, and they will not disrupt the stationary power-law distribution. However, when $f(\alpha,\beta)$ is not small, the whole distribution can be disrupted [see how $P(0)$ affects all elements in Eq. (3)], and the distribution will settle into another type of stationary distributions which decay much faster than scaling ones do. Therefore, the scaling condition can be written as $f(\alpha,\beta)<\epsilon$ where $0<\epsilon\ll 1$. In Fig. 3, we plot this condition

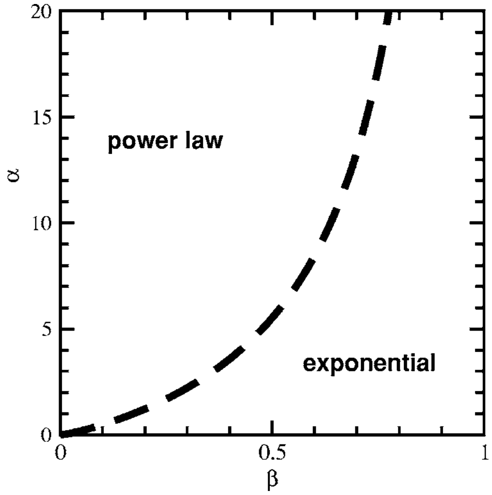

FIG. 3. The scaling condition in (α,β) space. The dashed line represents $f(\alpha,\beta)=\epsilon$ when $\epsilon=10^{-3}$. As α increases, the range of β for the power law approaches $0<\beta<1$.

when $\epsilon=10^{-3}$, estimated from the simulation results. The critical boundary $\alpha=\alpha_c(\beta)$ that separates the scaling region from the nonscaling region was obtained from $f(\alpha,\beta)=\epsilon$,

$$\alpha_c(\beta) = \frac{\beta}{1-\beta}\ln[\epsilon^{-1}\beta^{1/(1-\beta)}]. \tag{12}$$

The scaling region shown in (α,β) space corresponds well with results in Fig. 2, and is surprisingly big. If α is big enough, the system exhibits the power law for almost any value of β, which means that just a slight advantage given to the rich is enough to make the system follow the Zipf's law. In Fig. 4, we observe several cases with various parameter values using numerical simulations. In Fig. 4(a), we fix α at 5 and vary β, to observe that $\beta_c(5)\sim 0.5$. In Fig. 4(b), we vary α and β to observe that $\beta_c(\alpha)$ increases as α increases. When (α,β) is in the power-law region $[\beta<\beta_c(\alpha)]$, the shape of the stationary distribution is determined by β only, and α only changes the position of the cutoff, k_M. On the other hand, the shape of the stationary distribution will be determined by both α and β for nonscaling cases $[\beta>\beta_c(\alpha)]$.

IV. POSSIBLE APPLICATIONS

So far we have proposed a simple model using general terms such as elements and quantities. Here we discuss three examples where this mechanism can be applied.

A. Personal wealth

The first example is the wealth distribution with people and their assets, which is known to exhibit the power law especially for the richest people. In a society, the population does not always grow, and their total amount of assets can be assumed to be conserved. People also interact in many ways, changing their assets, and the rich have an advantage over the poor. In our model, α becomes the average amount of assets per person, and β is the parameter representing the

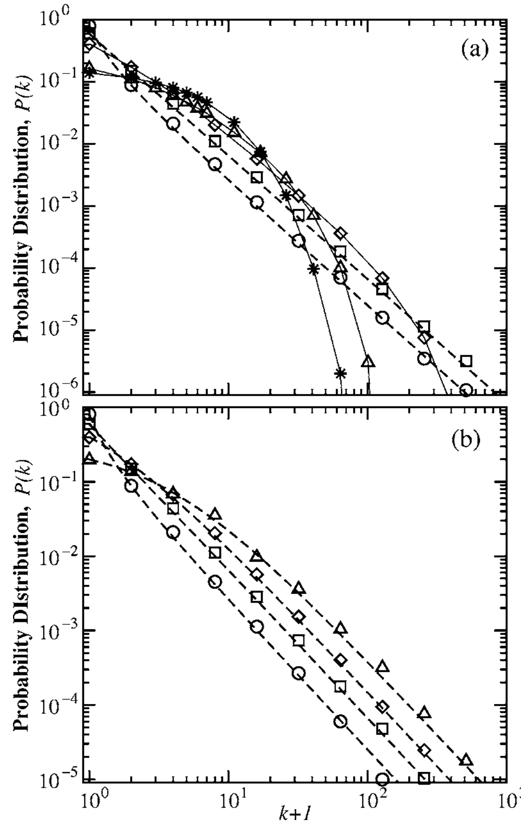

FIG. 4. Stationary distributions for various (α,β) values. Dashed lines represent theoretical stationary probability distributions for given β values, and data points are logarithmically binned for scaling cases. (a) When $N=10^6$ and $T=7\times 10^{11}$, α is fixed at 5, and $\beta=0.2$ (○), 0.4 (□), 0.6 (◇), 0.8 (△), 1.0 (*). The power-law distributions are observed clearly when $\beta=0.2$ and 0.4. (b) Examples of various (α,β) values exhibiting power-law distributions after $T=10^{12}$: $N=10^6$, $\alpha=1$, $\beta=0.2$ (○); $N=10^5$, $\alpha=10$, $\beta=0.4$ (□); $N=10^4$, $\alpha=100$, $\beta=0.6$ (◇); $N=10^3$, $\alpha=1000$, $\beta=0.8$ (△).

advantage for the rich. Because α is usually big enough, the power law with the Zipf exponent, $\gamma=2$, will emerge for almost any value of β, while the empirical data show $\gamma\simeq 2.091$ [1]. There are other nongrowth models for the power-law wealth distributions, which use the binary interactions of the traders [9–11].

B. City sizes

The second example is the distribution of city sizes with cities and their sizes. This is the original Zipf's law, and the Zipf exponent has become famous for this phenomenon (originally Zipf used the rank statistics and the exponent is 1, which is equivalent to $\gamma=2$ in our case). Our model can be applied to this case when the number of cities is fixed, and the overall population does not grow. Here an interaction is the migration of a person (or a family) from one city to another. People tend to move from a small city to a larger city; hence, β is the parameter representing this tendency. Then, the Zipf's law will emerge from our model. It will be unrealistic if $P(0)$ is not close to 0 because there are no empty cities usually. But when α is big enough and β is close to 1, the distribution will still be scaling and $P(0)$ will be close to 0. Even with a drawback of not taking account of the growth of cities from within unlike other models [3], this mechanism has some merits to be regarded as another valid explanation of the Zipf's law: (1) the model produces the Zipf exponent naturally with a simple mechanism; (2) the migration of people between cities is well represented by the model; (3) the attractiveness of the bigger cities is also well represented. There also can be a different approach. For example, Zanette and Manrubia [4] used the stochastic linear model, which assumes neither the growth nor the binary interactions.

C. Scale-free networks

The last example is the network with nodes and their degrees. A network is an entity that consists of nodes and links, while the degree of a node is the number of links connected to a given node. In many systems represented by networks, degrees of nodes have been found to follow power-law distributions (hence called scale-free networks). Based on the mechanism of linear preferential attachment proposed by Ref. [20], many extended models have been followed [24–27]. In these models, the assumption of the growth of nodes and links is crucial, and interactions between nodes are either ignored or used as an extra feature [26]. This approach is valid for many scale-free networks, but not suitable for nongrowing networks, in which node interactions are vital to their dynamics. Our model can generate this kind of scale-free networks by interpreting the interaction between nodes as the rewiring process. When nodes i and j are chosen, the rewiring process changes the link from (i',i) to (i',j) where i' is a *pivot* node chosen from nodes that are linked to i (loops and multiple links are allowed). Therefore, from our model, networks with power-law degree distributions can be generated through only rewiring, and the results will be presented in a forthcoming paper [28]. The network concept is actually related to many scaling phenomena, since they can be represented by networks directly [13,14,16,17,20] or indirectly [29].

V. CONCLUSIONS

We have proposed a preferential-redistribution mechanism that generates power-law distributions with the Zipf exponent for certain parameter ranges, and this scaling region in our parameter space has been found analytically using some numerical results. Since this scaling region is big enough and the mechanism is very simple, our model can be a good candidate to be used as a base mechanism for models describing some scaling phenomena, and three possible applications have been discussed here. Like other models, our model has limited applicability, but we believe that it can be extended to suit specific needs as a part of more realistic models, or generalized to have more flexible features.

ACKNOWLEDGMENTS

This work was supported by Grant No. R05-2002-000799-0 from the Basic Research Program of the Korea Science and Engineering Foundation.

[1] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).
[2] G. Zipf, *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Cambridge, 1949).
[3] X. Gabaix, Quart. J. Econom. **114**, 739 (1999).
[4] D. H. Zanette and S. C. Manrubia, Phys. Rev. Lett. **79**, 523 (1997).
[5] H. A. Simon, Biometrika **42**, 425 (1955).
[6] R. Gibrat, *Les Inégualités Économiques* (Libraire du Recueil Sirey, Paris, 1931).
[7] D. Champernowne, Econom. J. **63**, 318 (1953).
[8] V. Pareto, *Cours d'Economie Politique* (Droz, Geneva, 1896).
[9] S. Ispolatov, P. L. Krapivsky, and S. Redner, Eur. Phys. J. B **2**, 267 (1998).
[10] A. Chatterjee, B. K. Chakrabarti, and R. B. Stinchcombe, Phys. Rev. E **72**, 026126 (2005).
[11] Y.-B. Xie, B.-H. Wang, B. Hu, and T. Zhou, Phys. Rev. E **71**, 046135 (2005).
[12] B. A. Huberman and L. A. Adamic, Nature (London) **401**, 131 (1999).
[13] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[14] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
[15] G. U. Yule, Philos. Trans. R. Soc. London, Ser. B **213**, 21 (1924).
[16] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, Nature (London) **411**, 41 (2001).
[17] S. Redner, Eur. Phys. J. B **4**, 131 (1998).
[18] M. Mitzenmacher, Internet Mathematics **1**, 226 (2004).
[19] M. Levy and S. Solomon, Int. J. Mod. Phys. C **7**, 595 (1996).
[20] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[21] D. ben-Avraham, E. Ben-Naim, K. Lindenberg, and A. Rosas, Phys. Rev. E **68**, 050103(R) (2003).
[22] A. Fujihara, T. Ohtsuki, and H. Yamamoto, Phys. Rev. E **70**, 031106 (2004).
[23] A. Drăgulescu and V. M. Yakovenko, Eur. Phys. J. B **17**, 723 (2000).
[24] P. L. Krapivsky, G. J. Rodgers, and S. Redner, Phys. Rev. Lett. **86**, 5401 (2001).
[25] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[26] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
[27] B. Bollobás and O. Riordan, Internet Mathematics **1**, 1 (2003).
[28] S. Ree, e-print cond-mat/0601590.
[29] C. Andersson, A. Hellervik, K. Lindgren, A. Hagson, and J. Tornberg, Phys. Rev. E **68**, 036124 (2003).