Eur. Phys. J. B **22**, 375–379 (2001)

THE EUROPEAN
PHYSICAL JOURNAL B
EDP Sciences

Jump of the minimal site and a new correlation function in the Bak-Sneppen model

C.B. Yang[a] and X. Cai

Institute of Particle Physics, Hua-Zhong Normal University, Wuhan 430079, PR China

Received 14 April 2001 and Received in final form 28 June 2001

Abstract. A new type of spatio-temporal correlation function for the process approaching the self-organized criticality is investigated within the Bak-Sneppen model for biological evolution. In terms of the "directional shorter distance" between the two sites with minimum fitness at two successive updates, the correlation function is defined and studied numerically for the nearest- and random-neighbor versions of the model. Qualitatively different behaviors of the jump of the minimal site in the two models are presented, and the behaviors of the correlation functions are shown also different.

PACS. 05.40.-a Fluctuation phenomena, random processes, noise, and Brownian motion – 87.10.+e General theory and mathematical aspects – 64.60.Ak Renormalization-group, fractal, and percolation studies of phase transitions

1 Introduction

The phenomenon of "self-organized criticality" (SOC) has become a topic of considerable interest [1–8] because of the potentially wide applications to complex systems ranging from the behavior of sandpile and the description of the growth of surfaces to generic description of biological evolution. One of the key features of SOC is that the dynamics of complex systems in nature does not follow a smooth, gradual path. Instead it often occurs in terms of punctuation, or "avalanches" in other word. Numerous numerical studies have found power law distributions for the size and lifetime of avalanches and claimed SOC to occur in many specific models, and the transition to the SOC state was studied in [9–11]. Such complexity also shows up in simple mathematical models for biological evolution far from equilibrium.

It seems that the phenomenon SOC cannot be adequately characterized by the power-law distributions of avalanche size and lifetime, as concluded in [12]. In [12] were shown "some striking observable differences between two 'self-organized critical' models which have a remarkable structural similarity". The two models, as called nearest- and random-neighbor versions of the Bak-Sneppen (BS) model, were introduced in [13–15] and used to mimic biological evolution. The one-dimensional BS model involves L sites on a circle. Each site represents a species in the "food-chain". BS model provides a coarse-grained description of the evolution of the ecosystem of interacting species driven by mutation and natural selection. In computer simulations of the model, the L sites can be numbered $1, 2, \ldots, L$ in clockwise order on a circle, so the sites numbered 1 and L are nearest neighbors. Of course, the starting site for the numbering can be arbitrary. Random numbers are assigned to each site and are taken as a measure of the "survivability" of the species, therefore they are called fitness. Initially, the random number on each site is drawn uniformly from the interval (0, 1). In each update, the least survivable species with minimum fitness is to be found. In the local (or nearest-neighbor) version of the BS model, the evolution rule demands that the minimal site and its two nearest neighbors undergo fitness mutations and obtain new random numbers which are also drawn uniformly from (0, 1). In the second version, $K-1$ other randomly chosen sites besides the minimal one are involved in the fitness update (so this version is called random-neighbor model). As shown in [15–17], the randomness of the neighbors in each fitness update makes the second version analytically solvable. Investigation in [12] showed that some behaviors of the nearest- and random-neighbor models are qualitatively identical. They both have a nontrivial distribution of heights of minimum fitness, and each has power-law avalanche size and lifetime distributions. But the spatial and temporal correlations between the minimum fitness show different behaviors in the two models and thus can be used to distinguish them.

In [18] a spatio-temporal correlation between the locations $X(s)$ of minimum fitness at two consecutive updates is newly defined and investigated. The spatio-temporal correlation is an important topic in the study of SOC since the word "self-organization" used in contemporary studies for complex systems refers to a dynamical process in which a complex system starts from a state without correlation and ends up to a complex state with a high degree

[a] e-mail: yangcb@oregon.uoregon.edu

of correlation in both space and time dimensions. So, the process for a system to approach its SOC state should be characterized by the increasing strength of some properly defined spatio-temporal correlation functions. To the best of our knowledge, only a few attempts have been made along this direction. This paper is a new attempt along this line. In this paper, a new spatio-temporal correlation function is defined and investigated.

This paper is organized as following. In Section 2, we will discuss the jump of minimal site and the correlation between two minimal sites at two successive updates. Then in Section 3 we define a new spatio-temporal correlation function based on the jump of the minimal sites discussed in Section 2. Section 4 is the main conclusion.

2 Distribution of the jump of the minimal site and correlation

We begin with the local version of the BS model. Though an infinite series of exact equations can be derived for the BS model [19], an exact calculation is not possible for quantities such as the critical value and various exponents. All investigations show that the spatial and/or temporal correlations play important roles for the system approaching the critical point. To see how the spatio-temporal correlations in the process to its SOC state in the BS model are generated in this model, one can look at the evolution. Denote s the number of updates from the initial state. In the first a few updates from the initial state the location $X(s)$ of the minimum fitness is random and can jump throughout the whole lattice due to the independence and randomness of the fitness on all sites, and $X(s)$ can take a random value from 1 to L with equal probability, if the simulation is repeated many times. A useful quantity to describe the evolution is the gap $G(s)$ which is defined as the maximum of the minimum fitness before sth update step. When a gap G is for the first time reached, all sites have fitness not smaller than the gap. In the updates followed, the sites involved form a compact set in the one-dimensional BS model, and only those sites may have fitness less than G, before a higher new gap is reached. So, it is clear that the random location $X(s)$ of minimal site at time s should be near the former location $X(s-1)$ as long as the gap G remains unchanged. Therefore, the distribution of $X(s)$ will be peaked at $X(s-1)$ when s is large. With the update going on, $X(s)$ has stronger and stronger tendency to be in the neighborhood of $X(s-1)$ and the peak will be more and more obvious. Such a peak indicates the existence of correlation between $X(s)$ and $X(s-1)$. So that correlation between $X(s)$ and $X(s-1)$ appears naturally.

To describe the neighboring relation between $X(s)$ and $X(s-1)$, a directional shorter distance $\Delta(s)$ is introduced in [18]. Generally $\Delta(s)$ should be defined as a vector in higher dimensional BS model. For this one-dimensional study, the direction of vector $\Delta(s)$ can be represented by a sign. Thus, $\Delta(s)$ can have positive or negative values. The absolute value of $\Delta(s)$ can be thought as the shorter distance between $X(s)$ and $X(s-1)$ on the circle. In other words, $|\Delta|-1$ equals to the number of sites between $X(s)$ and $X(s-1)$ on the shorter curve on the circle. Therefore Δ is independent of the numbering of the sites. If the shorter curve from $X(s-1)$ to $X(s)$ is clockwise, $\Delta(s)$ is positive. Otherwise, it is negative. For definiteness, we assume $-L/2 \leq \Delta(s) < L/2$. Then we have

$$|\Delta(s)| = \begin{cases} |X(s) - X(s-1)| & \text{if } |\Delta(s)| < L/2 \\ L - |X(s) - X(s-1)| & \text{else.} \end{cases} \tag{1}$$

Obviously, as a measure of the jump of the minimal site in the updates, $\Delta(s)$ is also random for a simulation process from the initial state due to the randomness of the locations of the minimal sites in the update process. As mentioned in [10], the minimal site jumps throughout the system in a correlated and anomalous fashion which has some similarity to the usual Lévy flight picture. In [10] the jump is defined as $r = |X(s) - X(s-1)|$ (the notation used there is different), and the distribution of jumps is obtained from a simulation of 5×10^7 update steps in the *stationary state* for a system with size $L = 3000$. In other words, the jump distribution is calculated in [10] over a long period of update time. On the contrary, the directional shorter distance discussed in current paper is defined for a specific update step s from the initial state, so its distribution can be obtained only from an ensemble of simulations from the initial state. In addition, the absolute value of $\Delta(s)$ is not always equal to $|X(s)-X(s-1)|$ for a given numbering scheme for the sites. Thus $\Delta(s)$ in this paper is different from the jump r used in [10].

The meaning of the distribution of Δ is also different from that of x in the spatial correlation function $P_x(x)$ used in references [20,12]. First, x is no less than 0 in references [20,12], but Δ in this paper can be positive and negative. Second, no time information is recorded in $P_x(x)$, *i.e.*, $P_x(x)$ is calculated for the whole evolution process. But the distribution of Δ is obtained for specific scaled time t from an ensemble of evolution from the initial state by counting Δ at that update step. Therefore, the distribution of Δ contains both spatial and temporal information about the evolution process.

To have a look at the distribution of $\Delta(s)$, 0.5 million simulations from the initial state are performed for lattices with different size L. As suggested in [21], the number s of update steps from the initial state is a natural but not the best quantity for the description for a system to approach its SOC state, since the speed of the process depends on the size L of the lattice in the problem. The larger the system size L, the slower the process. A better quantity is the scaled time $t = s/L$ for the one-dimensional BS model. The scaled time t has a simple physical meaning. In the one-dimensional BS model, this quantity is equal to one-third of the average number of updates undergone for each site. For illustration, the distribution of $\Delta(s)$ is shown at $t = 2.0$, 4.0, 6.0 and 8.0 for $L = 200$, 400, 600 and 800 in Figure 1. It can be seen that for each t Δ is more likely to take small values. The larger t, the larger the probability

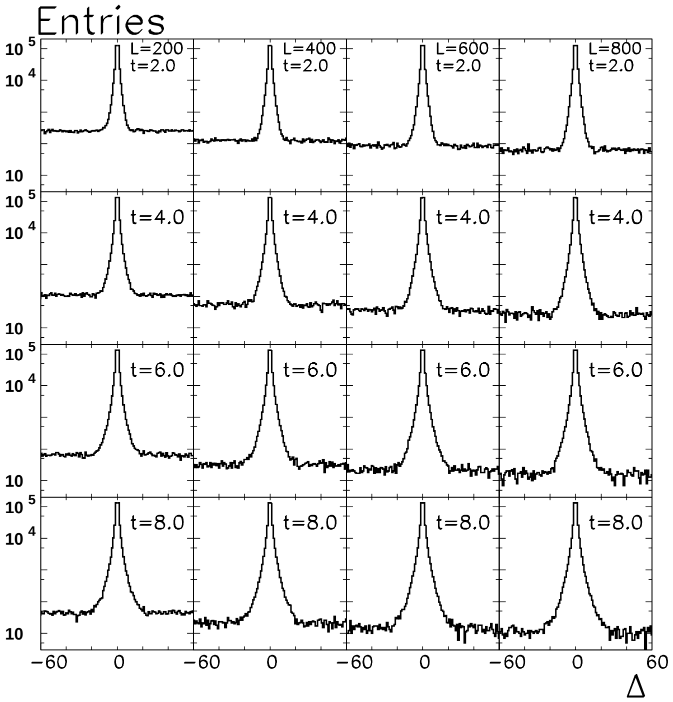

Fig. 1. Distribution of the directional shorter distance Δ in the local version of the BS model for systems with $L = 200$, 400, 600 and 800 at proper time $t = 2.0$, 4.0, 6.0 and 8.0. The distribution is normalized to the number of simulation events.

for Δ to take small values, thus the higher the peaks in Figure 1 with lower tails for larger $|\Delta|$. This is an indication of the local property of the evolution of the nearest version of the BS model. In fact, Δ can take large values when and only when an avalanche is over and a larger new gap is reached for the first time. At those update steps, the numbers assigned to the sites distribute independently and uniformly above the gap, therefore each site has the same probability to be the next minimal site, so that Δ can, with equal probability, take any possible integer value from $-L/2$ to $L/2 - 1$ when L is even, as used in present discussion.

3 Correlation between two successive minimal sites

The distribution of directional shorter distance between two successive minimal sites in the evolution can be used to measure the correlation between $X(s)$ and $X(s-1)$. As mentioned above, Δ can take any possible value with equal probability if there is no correlation between the locations of minimum fitness at two consecutive updates. The fact that there is much larger probability for Δ to take small values reveals the existence of correlation between $X(s)$ and $X(s-1)$. Since the peak in Δ distributions increases with t, indicating that $X(s)$ is more and more likely to be in the neighborhood of $X(s-1)$, there is more and more stronger correlation between $X(s)$ and $X(s-1)$. It would be useful to link the correlation with some characteristic quantity of the distribution of Δ. One of the most important characteristic quantities for the distribution is the moment

$$\langle \Delta^2(s) \rangle = \sum_{-L/2}^{L/2-1} \Delta^2 P(s, \Delta) = \frac{1}{N} \sum_{i=1}^{N} \Delta_i^2(s), \qquad (2)$$

in which the larger integer N is the number of simulations performed from the initial state, and $P(s, \Delta)$ is the probability for the directional shorter distance to take value Δ at sth update. When there is no correlation between $X(s)$ and $X(s-1)$, $\Delta(s)$ can take any value in $(-L/2, L/2-1)$ with equal probability, and therefore $\langle \Delta^2(s) \rangle$ is a constant for all s. The constant is equal to

$$\frac{1}{L} \sum_{i=-L/2}^{L/2-1} i^2 = \frac{L^2+2}{12} \cdot$$

The correlation between $X(s)$ and $X(s-1)$ shown above makes $\langle \Delta^2(s) \rangle$ smaller. The stronger the correlation, the smaller the $\langle \Delta^2(s) \rangle$. Since s and t have one-to-one correspondence for fixed L, one can use the proper time t as the variable for the evolution process. Thus one can define a new correlation function as

$$C(t) = 1 - \frac{12\langle \Delta^2(t) \rangle}{L^2+2} \cdot \qquad (3)$$

This correlation function has the demanded property that it increases with the correlation between minimal sites at two successive update steps. $C(t) = 0$ if no correlation between $X(s)$ and $X(s-1)$. The stronger the correlation, the smaller the $\langle \Delta^2(t) \rangle$, therefore the larger the correlation function $C(t)$. Since two minimal sites at two consecutive update steps are involved in the definition of this correlation function, it is a measure of some kind of spatio-temporal correlation in the BS model. To calculate $\langle \Delta^2(t) \rangle$ and $C(t)$, 0.5 million simulations from the initial state are done, and the calculated correlation function $C(t)$ as a function of proper time t is shown in Figure 2 for systems with size $L = 200$, 400, 600 and 800. The number of fitness updates in each simulation is such that corresponds to a scaled time $t = 10$. One can see that $C(t)$ is an increasing function of t. A remarkable property of $C(t)$ is that it has extremely weak dependence on the system size L, and the curves for different sizes almost coincide, which means that the correlation strength between $X(s)$ and $X(s-1)$ depends only on the proper time t (or in other words, on the average number of update steps each site undergone). This result is not surprising, considering the following two facts: (1) Such correlation is generated from the updates of the random fitness on each site; (2) t is the measure of mean number of updates of the fitness on each site. Of course, one cannot expect $C(t) \to 1$ when $t \to \infty$ for fixed finite L. From the update rules of the local BS model, it is not too difficult to see that Δ can take values $-1, 0, 1$ with equal probability, as shown by the wide heaps in Figure 1. Thus $\langle \Delta^2(t) \rangle$ can only approach $2/3$

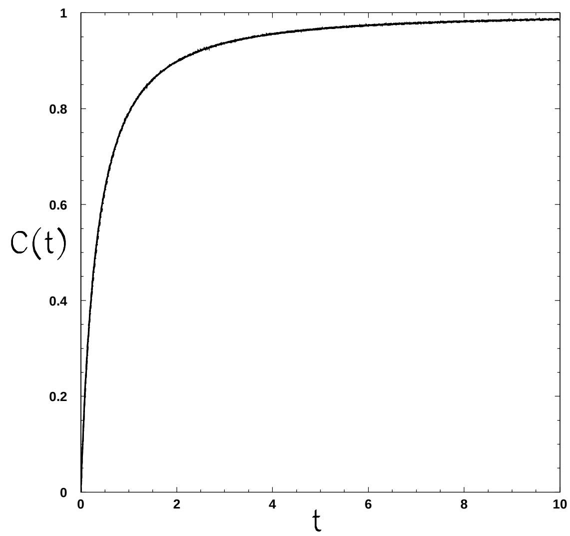

Fig. 2. The correlation function $C(t)$ in the local version of the BS model for the same systems as in Figure 1. The four curves cannot be distinguished by eyes. A curve from equation (4) is also drawn as a thick solid curve.

instead of 0. When L is large enough, the limiting value of $C(t)$ can be approximately taken to be 1.0.

We would like to point out that the correlation function $C(t)$ shown in Figure 2 can be parameterized quite well by following expression

$$C(t) = 1 - \left(\frac{t_0}{t_0 + t}\right)^{\delta} \left(\frac{1 + 0.24t + 0.005t^2}{1 + 2.0t + 0.1t^2}\right), \qquad (4)$$

with $t_0 = 0.85, \delta = 0.83$. The value of t_0 can tell us when the behavior of the correlation function changes. As can be seen from Figure 2 the correlation function takes small values but increases quickly when t is small. From last expression one can get $C(t) \sim 2.74t$ for $t \ll t_0$. For large t the correlation function takes large value but increases very slowly in the form $1 - C(t) \propto t^{-\delta}$ as can be shown from last expression for $t \gg t_0$. Thus the behavior of the correlation function changes at $t \simeq t_0$. We can say that the system approaches its critical state since $t \simeq t_0$.

One may be interested in the relationship between δ in this paper and other exponents known for the BS model. In [10] various exponents were derived from a scaling argument. It is clear that such a scaling argument is valid only for systems very close to their critical points. As we have mentioned above, the quantities calculated in this paper are obtained for specific t from *average over an ensemble of evolution* from the initial state, while those in previous studies were calculated from *average over a long update process* in the *stationary state*. Results from the two different averages may be very different. In addition, the quantities under investigation are also different. For the study of a system in equilibrium, the same result can be expected for the same observable from time and ensemble averages provided that the system is assumed ergodic. In the study of SOC, however, the systems are far from equilibrium, and the time evolution of such systems may depend on the history. For the BS model, the evolution detail of a system depends on the configuration of the fitness in the initial state. Because of such a dependence, different results for the same quantity may be obtained if the quantity is averaged along two different evolution tracks. If a system under investigation exhibits a bifurcation behavior, the two observations can be very different. After an ensemble average at a suitably defined time, however, the initial configuration dependence is eliminated, and more reliable results can be expected. In short, the time and ensemble averages can give different results, even when the same quantity is concerned, which usually measure different aspects of the same dynamics. The exponent δ may be related to those seen in the damage spreading [22]. In the damage spreading method one essentially monitors the time evolution of two or more copies of the same system with different initial configurations subjected to a specific dynamics and to the same thermal noise. We used exactly the same technique in this paper. It has been shown that the variation of the damage and related quantities with time, temperature, initial conditions and any other relevant parameters leads to information about the criticality of the system. But it is also found out that the results from the damage spreading process can be quite different for distinct dynamics [23]. In earlier studies of SOC such effect has not be taken into account, and the main focus is on various distributions which are obtained from observation over only one long evolution of the system. In the ensemble average of $\Delta^2(s)$ in this paper the divergency (or damage) among different copies of the system in the ensemble contributes. Thus some information on the damage spreading may be included in the exponent δ. Then we can say that the exponent δ in this paper may be independent of those for the *stationary state*. And very probably, many other independent exponents of such a type for states before the stationary one may exist.

Similar calculations can be done for the random neighbor version of the BS model. For illustration, only results with a system with $L = 200$ and $K = 3$ are presented here. The distributions of Δ are shown in Figure 3 for $t = 2.0, 4.0, 6.0, 8.0$. The distributions for this version are very different from those in Figure 1 for the local version of the BS model. 1 million simulations from the initial state are performed. Except $\Delta = 0$, the distributions are flat for all t. With the increase of t, the distribution of Δ has a slightly higher peak at 0 and lower constant value for other Δ. The particularity of $\Delta = 0$ in this version of the BS model can be easily understood from following two facts: (1) The three sites involved in an update step have larger probability to be the next minimal site; (2) If the last minimal site is, by accident, the next minimal site again, $\Delta = 0$, otherwise Δ can take any value because of the randomness of the two neighbors and of the fitness on all the sites. So, the distribution for $\Delta \neq 0$ must be flat, as shown in the figure. The calculated $C(t)$ is shown as a function of t in Figure 4. For comparison, the corresponding correlation function in the local version is re-shown in the same figure. One can see that the correlation in the random version is much weaker than in the local one. In

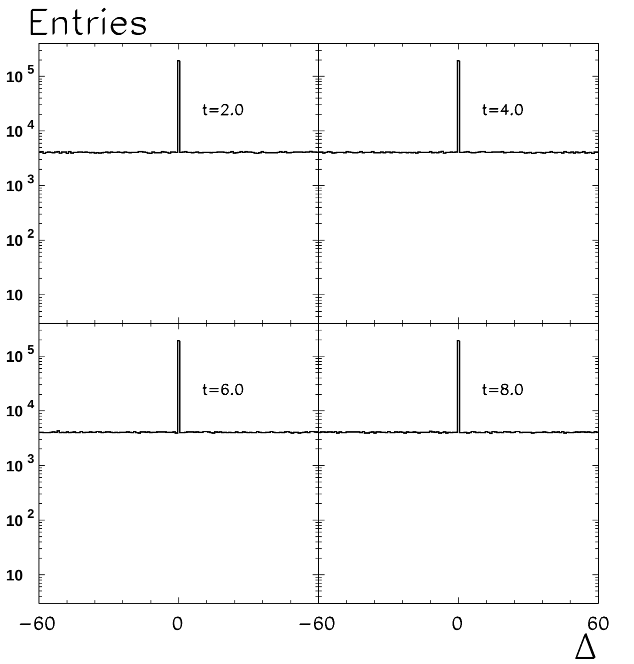

Fig. 3. Distribution of the directional shorter distance Δ in the random neighbor version of BS model for systems with $L = 200$ at proper time $t = 2.0$, 4.0, 6.0 and 8.0. The distribution is normalized to the number of simulation events.

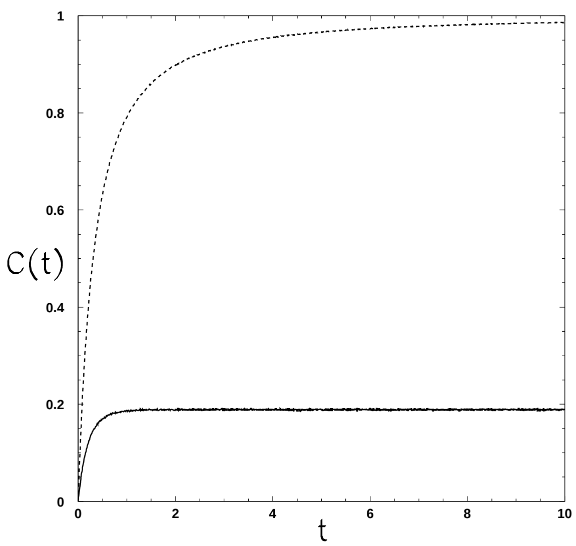

Fig. 4. The correlation function $C(t)$ in the random neighbor version of the BS model for the system with size $L = 200$. For comparison, the corresponding curve for the local version is re-shown by dashed curve.

the small t region, $C(t)$ for the random version increases with t and reaches its saturating value quickly. For $t > 1$ no more increase of $C(t)$ can be observed. So the degree of correlation between $X(s)$ and $X(s-1)$ for later updates remains the same, which is much weaker than in the local version of the model. Thus, the different behaviors of the correlation can be used to distinguish the two versions of the BS model. Since strong correlation in the stationary state is one of the key features of SOC, one can claim that the random neighbor version of the BS model is not a SOC model, as concluded in [12].

4 Conclusion

In conclusion, a new spatio-temporal correlation function in the BS model between the minimal sites at consecutive update steps is investigated for the local and nonlocal version of the BS model. Similar correlation function can be used in other minimal models to judge the SOC nature of the models.

This work was supported in part by the NNSF in China and NSF in Hubei, China. One of the authors (C.B.Y) would like to thank Alexander von Humboldt Foundation of Germany for the financial support granted to him.

References

1. P. Bak, C. Tang, K. Wiesenfeld, Phys. Rev. A **38**, 364 (1987); Phys. Rev. Lett. **59**, 381 (1987).
2. K. Chen, P. Bak, S.P. Obukhov, Phys. Rev. A **43**, 625 (1991).
3. P. Bak, K. Chen, M. Creuts, Nature **342**, 780 (1989).
4. K. Sneppen, Phys. Rev. Lett. **69**, 3539 (1992); K. Sneppen, M.H. Jensen, *ibid.* **70**, 3833 (1993); **71**, 101 (1993).
5. P. Bak, K. Chen, Sci. Am. **264**, 46 (1991).
6. K. Chen, P. Bak, Phys. Lett. A **140**, 46 (1989).
7. P. Bak, K. Chen, C. Tang, Phys. Lett. A **147**, 297 (1990).
8. A. Sornette, D. Sornette, Europhys. Lett. **9**, 197 (1989).
9. M. Paczuski, S. Maslov, P. Bak, Europhys. Lett. **27**, 97 (1994).
10. M. Paczuski, S. Maslov, P. Bak, Phys. Rev. E **53**, 414 (1996).
11. A. Corral, M. Paczuski, Phys. Rev. Lett. **83**, 572 (1999).
12. J. de Boer, A.D. Jackson, Tilo Wetig, Phys. Rev. E **51**, 1059 (1995).
13. P. Bak, K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).
14. H. Flyvbjerg, P. Bak, K. Sneppen, Phys. Rev. Lett. **71**, 4087 (1993).
15. J. de Boer, B. Derrida, H. Flyvbjerg, A.D. Jackson, T. Wettig, Phys. Rev. Lett. **73**, 906 (1994).
16. Yu.M. Pis'mak, J. Phys. A **28**, 3109 (1995).
17. Yu.M. Pis'mak, Phys. Rev. E **56**, R1326 (1997).
18. C.B. Yang, X. Cai, Z.M. Zhou, Phys. Rev. E **61**, 7243 (2000).
19. S. Maslov, Phys. Rev. Lett. **77**, 1182 (1996).
20. K. Ito, Phys. Rev. E **52**, 3232 (1995).
21. C.B. Yang, X. Cai, Eur. Phys. J. B **21**, 109 (2001).
22. L. da Silva, F.A. Tamarit, A.C.N. Magalhães, J. Phys. A **30**, 2329 (1997) and references therein.
23. F. Wang, M. Suzuki, Physica A **223**, 34 (1996).