Fluctuations and correlations in two-dimensional Bak–Sneppen model

2002 J. Phys. A: Math. Gen. 35 5997

(http://iopscience.iop.org/0305-4470/35/29/305)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 5997–6007 PII: S0305-4470(02)33282-7

Fluctuations and correlations in two-dimensional Bak–Sneppen model

C B Yang

Institute of Particle Physics, Hua-Zhong Normal University, Wuhan 430079, People's Republic of China

Received 29 January 2002, in final form 20 May 2002
Published 12 July 2002
Online at stacks.iop.org/JPhysA/35/5997

Abstract

The self-organized criticality in the two-dimensional Bak–Sneppen model is studied from the fluctuations of the mean fitness and the jump of the minimum sites. An ensemble having the same number of updates from the initial state is used in the investigation. From the Gaussian distribution of the mean fitness in the ensemble, the lattice size dependence of the ensemble-averaged fitness is found. From the dependence, a solution to the gap equation is found and the critical gap and an exponent are calculated. A spatial–temporal correlation function is investigated from the jump of the minimum site.

PACS numbers: 05.40.−a, 64.60.−i

1. Introduction

Self-organized criticality (SOC) is one of the most important topics for investigating the formation of a great variety of patterns in complex systems. Its potential applications range from self-similar, fractal behaviour in nature [1–4], $1/f$ noise in quasar [5], river flow[1] and brain activity [6], to many natural and social phenomena, including earthquakes, economic activity and biological evolution. It was suggested a decade ago that all these phenomena are signatures of spatial–temporal complexity and can be related via scaling relations to the fractal properties of some discontinuous evolutions (also called 'avalanches') of the systems. One of the simplest SOC models is the Bak–Sneppen (BS) model [7]. In the d-dimensional BS model, L^d sites are assigned random numbers ξ_j (called fitness of the sites) drawn uniformly in (0, 1) in the initial state. In each update followed, a site with minimum fitness is determined and the site together with its nearest $2d$ neighbours is given a new random number also drawn uniformly in (0, 1). Periodic boundary conditions are adopted in the model. The maximum of the minimum fitness before the sth update is called the gap $G(s)$ of the system. Needless

[1] See, for example, the empirical observations of H E Hurst described in [1, 3].

to say, $G(s)$ is a staircase function. The process of updates with a constant gap $G(s)$ is called an avalanche. In the large L limit there is a gap equation

$$\frac{\mathrm{d}G}{\mathrm{d}s} = \frac{1-G}{L^d \langle S\rangle_G} \tag{1}$$

where $\langle S\rangle_G$ is the mean lifetime of avalanches with a common gap G. The gap equation can tell us how the system is driven to its critical state. This important equation is, however, impossible to solve analytically because we do not know in advance the expression of $\langle S\rangle_G$. One additional difficulty associated with solving the gap equation is that the equation is exactly true only in the limit $L \to \infty$ because $G(s)$ is not continuous for finite L.

A basic feature of the BS model is its randomness of the fitness on the sites. Because of the randomness of the fitness on all the sites, all observations about the system are also random. Therefore, analysing the fluctuations of quantities of the system is very important and can give us some interesting information about the system, as we have shown in [8]. In [8], we showed for the one-dimensional BS model that the scaled time $t = s/L$ is a better variable for the description of the system approaching the critical state and that important knowledge on the evolution of the system can be obtained by analysing the fluctuations of the mean fitness of the system in an ensemble of update processes at fixed t. The distribution of the mean fitness in an ensemble of updates at fixed t starting from the same initial state is a Gaussian, as can be foreseen from the central limit theorem. The peak position of the distribution has a simple L-dependence from which the gap G can be obtained as a function of the scaled time t, rather than the update number s, for an infinite lattice. $G(t)$ is the solution to equation (1) for $L \to \infty$. Then we can get $\langle S\rangle_G$ and some characteristics of the BS model, such as the critical value f_C and some critical exponents. In [9], the flying of the minimum site is investigated by introducing a directional short distance Δ between the minimum sites at two successive updates. For fixed t, the distribution of Δ in an ensemble of updates is peaked at 0. With the increase of t, the distribution gets more and more obviously peaked. The appearance of a peak in the distribution of Δ indicates the existence of correlations between positions of the minimum sites at successive updates. From the discussion we see that the study of fluctuations is a new way of investigating the properties of the BS model.

In this paper we extend the discussions in [8, 9] to the two-dimensional case. We will give the solution to the gap equation as a function of proper time t and from the solution calculate characteristic quantities of the model. We will also study the correlations between positions of the minimum sites at two successive updates. The organization of the paper is as follows. In section 2, we investigate the fluctuations of the mean fitness in a two-dimensional BS model. In section 3, we focus on the correlations between the locations of minimum sites at two successive updates and calculate the t-dependence of a spatial–temporal correlation function. Section 4 ponders over the method of fast simulation of the model and possible applications of the method. Section 5 is a brief summary.

2. Fluctuations in the two-dimensional BS model

In [8], we suggested an investigation of the BS model through its event-by-event fluctuations of the mean fitness over the lattice, $f_i = \frac{1}{L^d}\sum_{j=1}^{L^d} \xi_j$, at fixed scaled time (we called it the proper time) $t = s/L^d$. In the d-dimensional BS model there are $2d+1$ sites involved in an update. Thus, $2d+1$ times the proper time is the mean number of updates each site undergoes in the s-updates of the system. The method used in [8] is quite different from the methods used by other authors in the study of the same model. Usually, one studies the BS model in computer simulations by counting the number of appearances of some quantities in *a long*

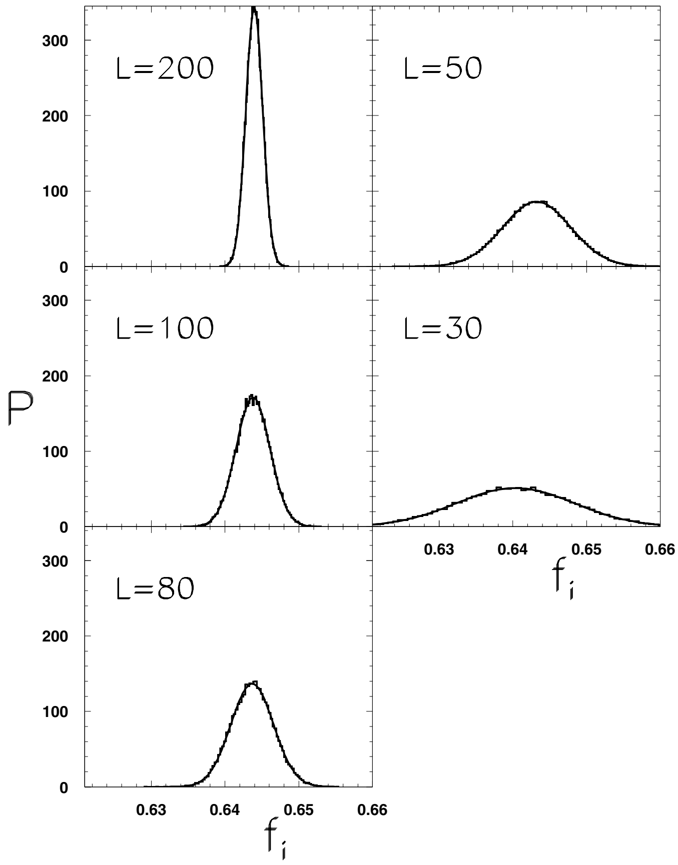

Figure 1. Distribution of the mean fitness on two-dimensional lattices with $L = 30, 50, 80, 100$ and 200 at $t = 2$.

period of update from the initial state. In other words, an average of quantities is performed over a whole update process. After the average, the time degree plays no role in the final results. In our new way, we focus on the fluctuations of f_i *in an ensemble of updates* at the same proper time t. From the course of statistical physics we learnt that these two averages may give different results unless the system is ergodic. In the BS model, the system is surely not ergodic, at least not before it reaches the critical state. Therefore, the ensemble average proposed in [8] is a novel way for the study of complex systems.

Now we can study the fluctuations of the mean fitness f_i in an ensemble of updates at the same proper time t starting from the same initial state for the two-dimensional BS model. As in [8], a Gaussian distribution of f_i can be expected from the central limit theorem

$$f_i = f_P + \frac{r_i}{\sqrt{L^2}} \tag{2}$$

where f_P gives the peak position and r_i is a Gaussian random number with zero mean. Computer simulations of the BS model on two-dimensional lattices of different sizes support this expectation. In figure 1, the distribution of f_i for $L = 30, 50, 80, 100$ and 200 is shown for $t = 2$. Gaussian fits to the curves are also drawn there. Perfect agreement can be seen there.

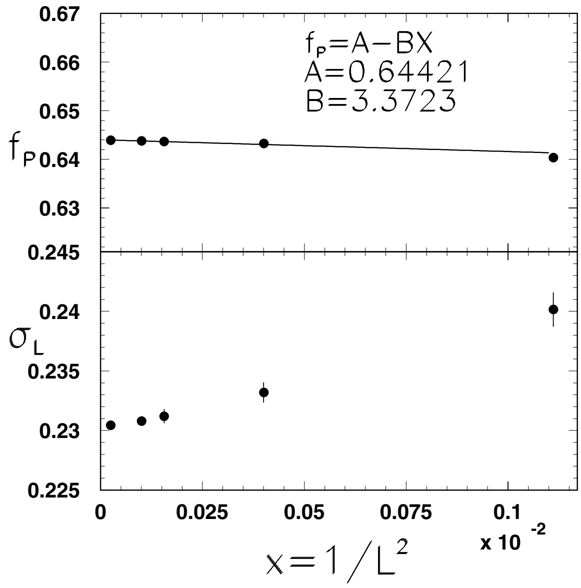

Figure 2. Lattice size dependence of the peak position and width of the Gaussian noise r_i in figure 1.

From figure 1, one can find a right shift of the peak position with the increase of lattice size L. The width σ of the distribution also shows L-dependence. If the width of the distribution of r_i were independent of L, the L-dependence of σ would be trivially inversely proportional to L in the two-dimensional BS model. The fitted width of the distribution of r_i, which is denoted as $\sigma_L = L\sigma$, is shown in figure 2 together with the L-dependence of the fitted f_P. The L-dependence of f_P and σ_L indicates that the distribution of ξ_j on the sites depends on the lattice size L. In particular, we can see from figure 2 that

$$f_P = a - \frac{b}{L^2} \tag{3}$$

is a very good approximation. f_P for other t can be calculated in the simulations through an ensemble average of f_i at the same t. We denote the result of this average as $f_L(t)$. For different L, the behaviour of $f_L(t)$ is shown in figure 3. When t is large, small L-dependence of $f_L(t)$ can be observed. In the limit $L \to \infty$, $f_L = (1+G)/2$. Therefore, the solution $G(t)$ to the gap equation, equation (1), can be obtained from the simulation results with two different lattice sizes L_1 and L_2 from equation (3) which is valid for any t with a and b dependent on t.

From equation (1) one can see that $-\ln(1-G)$ is more interesting than G itself because $-\ln(1-G)$ as a function of t not only records how $G(t)$ changes with time t but also contains directly the information about the mean lifetime $\langle S\rangle_G$ of avalanches with gap G. The reciprocal of the first derivative of $-\ln(1-G)$ with respect to t is exactly $\langle S\rangle_G$ according to equation (1). Therefore we plot $-\ln(1-G)$ as a function of t in figure 4. In the figure, another curve is drawn according to

$$-\ln(1-G) = A\left(1 - \left(\frac{t_0}{t+t_0}\right)^{\delta} \frac{1+0.35t+t^2}{1+0.55t+0.8t^2}\right) \tag{4}$$

with $A = 0.399$, $t_0 = 0.67$ and $\delta = 1.429$.

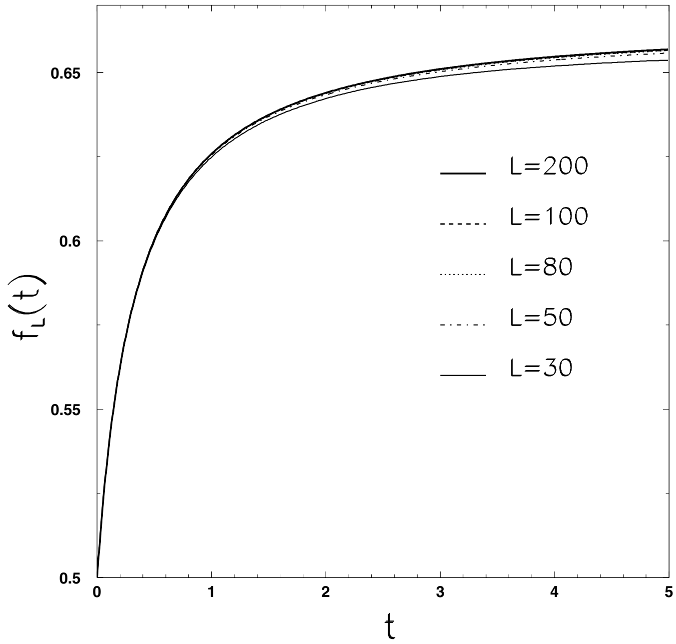

Figure 3. The ensemble-averaged mean fitness f_L as a function of t for different lattice sizes.

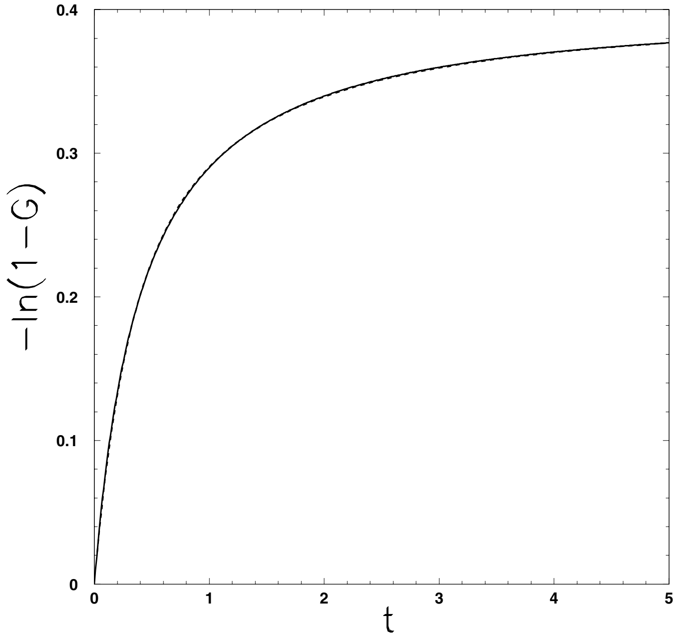

Figure 4. $-\ln(1 - G)$ as a function of time t for an infinite lattice in the two-dimensional BS model.

The agreement of equation (4) with the result of $-\ln(1-G)$ from the simulations is almost perfect. This almost perfect agreement, however, does not necessarily imply that equation (4)

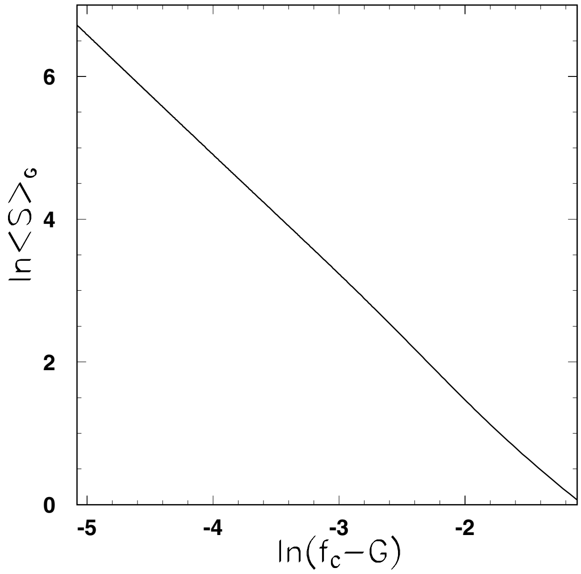

Figure 5. Mean lifetime of avalanches with gap G as a function of $\ln(f_C - G)$.

is the only possible parametrization of $-\ln(1 - G)$ from the simulations. In practice, many different expressions can be used quite satisfactorily in fitting a curve, and normally, the nonlinear fitting is a tedious and challenging task, especially when the number of parameters is large. The expression above is obtained from eyes-guided fitting based on the following considerations: (1) $-\ln(1-G)$ must be zero at $t = 0$; (2) it should approach $A = -\ln(1-f_C)$ in the limit of $t \to \infty$, and (3) the speed of the approach is not too fast and can be taken as a power of t. Then, A, t_0 and δ can be obtained from the fitting data points with large t. The rational expression is inserted to fit the points for medium t. Considering the almost perfect agreement of equation (4) with the results from simulations, we can use equation (4) safely to calculate some characteristic quantities for the two-dimensional BS model. First of all, we can easily get the critical value f_C which is the gap G at $t \to \infty$. Thus we have $f_C = 1 - \exp(-A) = 0.329$, which agrees perfectly with the result from simulations done in [10]. Then we can see how fast the system is driven to the critical state. From equation (4) we get $f_C - G \propto t^{-\delta}$ for $t \gg 1$. Thus, δ is an exponent describing the speed of the gap G approaching the critical value f_C. Finally, we can calculate the mean lifetime of avalanches with gap G, $\langle S\rangle_G$, as a function of t. Since we have determined $G(t)$, we can show $\langle S\rangle_G$ as a function of G also. $\langle S\rangle_G$ is shown in the log–log plot (figure 5) as a function of $f_C - G$. The curve in the log–log plot is not a perfect straight line. Therefore, the naive ansatz used [10] is not valid to some extent. The slope in the small $\ln(f_C - G)$ region, which is $(\delta + 1)/\delta = 1.7$, is in good agreement with the exponent given in [10] but a little larger than that in the large $\ln(f_C - G)$ region.

Now we see that our new method of studying the BS model can give the same information about the critical state of the system as other methods. More information about the transition is also given in this new way.

3. Correlations in the jumping of the minimum site

SOC is a process of a complex system driving itself from an uncorrelated initial state to a highly correlated critical state. Therefore, understanding the emergence and evolution of the

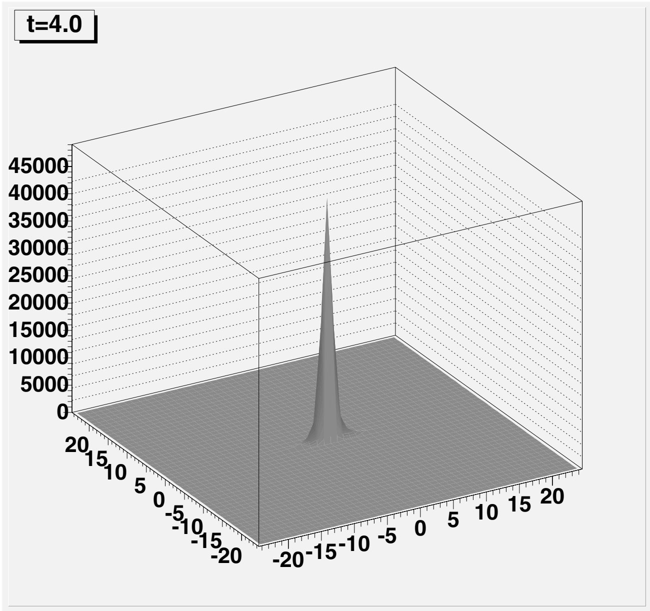

Figure 6. Distribution of jumps $(\delta x, \delta y)$ of the minimum site in the two-dimensional BS model with $L = 100$ at $t = 4$.

correlations in the system is crucial in studying the complex system. For the BS model very few attempts have been made to investigate the spatial–temporal correlations in the evolution to the critical state. Tracking a long update process, one can get in computer simulations the probability for a minimum site to become the minimum site again for the first time after S update and that for two minimum sites at successive updates to have a distance R. These two distributions can give us some information about the correlations in space and time. But they are not direct measures of the spatial–temporal correlations in the evolution of the system in the BS model. In [9], we analysed the jumping of the minimum site in the one-dimensional BS model and found that the distribution of the jump becomes more and more peaked at 0 as the update proceeds. The peaked distribution of the jump is an indication of the correlation between minimum sites at two successive updates. We defined a correlation function to associate the width of the distribution with the correlation. In this paper, we extend the discussion in [9] to the two-dimensional case and study the spatial–temporal correlation in the BS model.

In the two-dimensional BS model with L^2 sites, a minimum site can be specified by two Cartesian coordinates (x, y). The jump of the minimum sites is characterized by a vector $\vec{\Delta} = (\delta x, \delta y)$ as the extension of the directional shorter distance for the one-dimensional case. For simplicity, we let the jump in both x and y directions be in the range $[-L/2, L/2)$. At a very early stage of evolution, each site has the same probability of being the minimum site. Thus, the values of δx and δy in a jump can take any possible value with equal probability. When the gap G is increased, only the sites involved in an avalanche of gap G have random numbers less than G and may be the minimum site. Therefore, the jump of the minimum site is bounded among the active sites. As a consequence, the distribution of the jumps becomes more and more peaked with increasing gap G. A typical distribution of the jump of the minimum site in a two-dimensional BS model is shown in figure 6 for $t = 4$ and $L = 100$. The distribution is normalized to the number of events, $N = 100\,000$, in the ensemble of updates. There is an extremely sharp peak. For other L and large enough t, similar distributions can be verified.

As in [9] we calculate the mean value of the jump squared, $\langle\vec{\Delta}^2\rangle$. If there is no correlation between the locations of the minimum sites at two successive updates, δx and δy can take any value in $[-L/2, L/2)$ with the same probability $p(\vec{\Delta}, t) = 1/L^2$. In this case $\langle\vec{\Delta}^2\rangle$ will be

$$\sum_{\vec{\Delta}} \vec{\Delta}^2 p(\vec{\Delta}, t) = \frac{2}{L} \sum_{-L/2}^{L/2-1} i^2 = \frac{L^2+2}{6}.$$

From figure 6, one can see very clearly that the distribution of $\vec{\Delta}$ is not uniform. To calculate $\langle\vec{\Delta}^2\rangle$ in our simulation, we set up an ensemble with $N = 100\,000$ updates at the same t from the initial state. Then

$$\langle\vec{\Delta}^2\rangle = \frac{1}{N} \sum_{i=1}^{N} \vec{\Delta}_i{}^2 \tag{5}$$

in which $\vec{\Delta}_i$ is the jump at time t for the ith simulation event in the ensemble. The value of $\langle\vec{\Delta}^2\rangle$ in the simulation is less than that for the case without correlation. The larger the discrepancy, the stronger the correlation. We can define a correlation function in the same way as in [9]

$$C(t) = 1 - \frac{6\langle\vec{\Delta}^2\rangle}{L^2+2}. \tag{6}$$

This correlation function is spatial–temporal because it involves *a spatial jump of minimum site between two successive temporal updates*. The calculated correlation function is shown in figure 7 as a function of t in the range (0, 5) from simulations with a different lattice size L. At large t, the finite-size effect plays a role in the correlation function. For very large lattice, the correlation function can be parametrized as

$$C(t) = 1 - \left(\frac{t_0}{t_0+t}\right)^{\delta} \frac{1+1.1t+0.32t^2}{1+0.35t+0.65t^2} \tag{7}$$

with $t_0 = 0.215$ and $\delta = 1.49$. δ indicates how fast the correlation function reaches its saturated value at large t. When $t \ll t_0$ the correlation is very weak but the correlation function increases linearly with time t. When $t \gg t_0$ the correlation is strong and the correlation function increases very slowly. Therefore, t_0 is regarded as a turning point from weak to strong correlation. Across the point, the behaviour of the correlation function is changed. We can use the value of t_0 to mark when the system is near the critical state.

4. Fast simulation of the Bak–Sneppen model

Because of the intrinsic complexity of SOC, no simple mathematical formula can be employed to describe the behaviour. Therefore, computer simulation is at present the main tool for the investigation of properties of complex systems. In many cases, the use of discrete space and time is necessary at the current computation level. With the increase in the size of the discrete spatial–temporal lattice, the simulation will become slower and slower. This hampers the study of complex systems. Thus, a very important practical topic is how to improve the simulation speed.

Once the maximum ‘proper time’ $t_{\max}$ is fixed in the simulations, the number of update steps $N_{\max} = L^d t_{\max}$ is proportional to the lattice size L^d. It is clear that the main computation in each step is to find the minimum site. If we compare the fitness on all the sites directly, $L^d - 1 \simeq L^d$ comparisons are needed. Therefore, the time consuming process of simulation

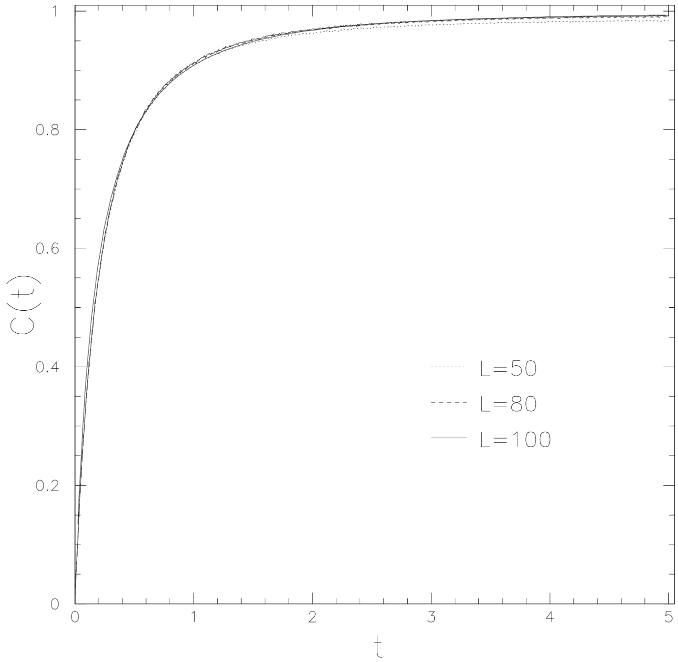

Figure 7. Behaviour of the correlation function $C(t)$ for t in (0, 5) from simulations of the two-dimensional BS model with different lattice size L.

of the BS model is proportional to L^{2d}. In this section, we try to introduce a new method to quicken the simulation of the BS model. We will show below that the time consuming process will become proportional to L^d for fixed $t_{\max}$.

The only way to quicken the simulation is to reduce the computation in finding the minimum site when $t_{\max}$ is fixed. In our problem, an array of a list can be introduced for this purpose. In each element of the list array, the positions of sites with fitness in a certain interval can be stored. In simulating the BS model one can divide the total interval of fitness (0, 1) into N_L parts. Then a list array of size N_L should be used, and the ith ($i = 0, 1, \ldots, N_L - 1$) list contains the positions of sites with fitness in the interval $[i/N_L, (i+1)/N_L)$. One can see that the minimum fitness can be found in the jth list when all i lists with $i < j$ are, by chance, empty. If there are N_j elements in the list, $N_j - 1$ comparisons are enough to find the minimum. In our simulation we choose $N_L = L^d/50$. Thus, on average, only 50 comparisons are needed to find the minimum site, no matter how big a lattice is used. The list array is filled during the initialization of the random fitness on the sites. When a site undergoes an update of random fitness, a new rank i for the new fitness is calculated. If the new rank is the same as the one the fitness is in before the update, no change in the list is necessary. Otherwise, a node is deleted from the old list and added to the new list. With updating, fewer and fewer fitnesses lie below a certain gap G, so that the procedure of finding the update centre may become faster and faster. This is an advantage of using list array. Now the step needed in the simulation of the BS model is proportional only to the update steps for fixed $t_{\max}$. By the way, the method using list array can be easily coded with the C++ programming language (see [11, 12]). We compared simulations of the BS model with and without list array for a two-dimensional lattice of $L = 100$ in each direction and found that the new method is about

100 times faster. If binary search trees are used for the non-empty lists with lowest ranks, the update process can be even faster.

The method discussed above is applicable to some other lattice models (such as a self-organized growth model described in [13] for the quenched Herring–Mullins equation). In fact, there exists a whole class of models in which the rules consist of selecting the site with the extremal (global maximum or minimum) value of some variable and then changing the values of the variable on the site and its nearest neighbours according to some stochastic rule. These models, referred to as extremal models, were extensively studied (for a review, see [10]). They were employed to describe a variety of physical phenomena such as fluid invasion in disordered porous media [14], low-temperature creep [15], earthquake dynamics [16], etc.

It is more interesting to note that a similar method can be used in the sorting processes. For the basic task of sorting N elements, the best algorithms require on the order of several times $N\log_2 N$ operations. For $N < 50$, roughly, Shell's method is better, but it goes as $N^{3/2}$ in the worst case. For large N, Quicksort is, on average, the fastest known sorting algorithm. But in the worst case it can be a N^2 method! Another fast sorting method is the Heapsort. See [17] for a more detailed discussion.

Selection is sorting's austere sister. The most common use of selection is in the statistical characterization of a set of data. The operation count scales as N. By using a list array the problem can be converted into a N^0 one. This method is more efficient when more than one median value needs to be selected.

5. Summary

Using an ensemble of update processes starting from the initial state, we study the fluctuations of the mean fitness in the ensemble of updates at the same proper time t and the correlation between the locations of two minimum sites at successive updates. We found the lattice size dependence of the ensemble-averaged fitness and got a solution to the gap equation. Some characteristic quantities are calculated from the solution. From the jumping of the minimum site a spatial–temporal correlation function is defined and calculated. A method for a fast simulation of the BS model together with its possible applications is briefly discussed.

Acknowledgment

This work is supported in part by NNSF in China under grant no 10075035.

References

[1] Mandelbrot B B 1983 *The Fractal Geometry of Nature* (New York: Freeman)
[2] Coleman P H, Pietronero L and Sanders R H 1988 *Astron. Astrophys.* **200** L32
Pietronero L 1987 *Physica* A **144** 257
[3] Rinaldo A, Rodriguez-Iturbe I, Rigon R, Ijjasz-Vasquez E and Bras R L 1993 *Phys. Rev. Lett.* **70** 822
[4] Barton C C and Lapointe P R (ed) 1994 *Fractals in the Earth Sciences* (New York: Plenum)
[5] Press W H 1978 *Comment. Astrophys.* **7** 103
[6] Gruneis F, Nakao M and Yamamato M 1990 *Biol. Cybern.* **62** 407
[7] Bak P and Sneppen K 1993 *Phys. Rev. Lett.* **71** 4083
[8] Yang C B and Cai X 2001 *Eur. Phys. J.* B **21** 109
[9] Yang C B and Cai X 2001 *Eur. Phys. J.* B **22** 375
[10] Paczuski M, Maslov S and Bak P 1996 *Phys. Rev.* E **53** 414
[11] Stroustrup B 1995 *The C++ Programming Language* 2nd edn (Reading, MA: Addison-Wesley)
[12] Barton J J and Nackman L R 1995 *Scientific and Engineering C++, An Introduction with Advanced Techniques and Examples* (Reading, MA: Addison-Wesley)

[13] Park K and Kim I 2000 *Phys. Rev.* E **61** 4606
[14] Wilkinson D and Willemsen J F 1983 *J. Phys. A: Math. Gen.* **16** 3365
[15] Zaitsev S I 1992 *Physica* A **189** 411
[16] Ito K 1995 *Phys. Rev.* E **52** 3232
[17] Press W H, Teukolsky S A, Vettering W T and Flannery B P 1993 *Numerical Recipes in C* 2nd edn (Cambridge: Cambridge University Press) ch 8