

Probabilistic approach to the Bak-Sneppen model

G. Caldarelli,¹ M. Felici,^{1,2} A. Gabrielli,¹ and L. Pietronero¹

¹*INFN, Unità di Roma 1 “La Sapienza,” Piazzale Aldo Moro 2, 00185 Roma, Italy*

²*Laboratoire PMC, Ecole Polytechnique 91128 Palaiseau, France*

(Received 29 October 2001; published 18 March 2002)

We study here the Bak-Sneppen model, a prototype model for the study of self-organized criticality. In this model several species interact and undergo extinction with a power-law distribution of activity bursts. Species are defined through their “fitness” whose distribution in the system is uniform above a certain threshold. Run time statistics is introduced for the analysis of the dynamics in order to explain the peculiar properties of the model. This approach based on conditional probability theory, takes into account the correlations due to memory effects. In this way, we may compute analytically the value of the fitness threshold with the desired precision. This represents a substantial improvement with respect to the traditional mean field approach.

DOI: 10.1103/PhysRevE.65.046101

PACS number(s): 02.50.-r, 64.60.Ak, 04.20.Jb

I. INTRODUCTION

In the recent years, many models mimicking the scale-free behavior exhibited by natural phenomena such as river basins [1,2], fracture dynamics [3,4], earthquakes [5], have been extensively studied. The main features of these models are the lack of spatial and temporal characteristic scales, the evolution through intermittent bursts of activity and the absence of the fine tuning of some parameter to reach the critical state. To describe all this, the concept of self-organized criticality (SOC) [6] has been introduced, and many works have been devoted to clarify its real nature [7].

We focus here on the evolution model introduced by Bak and Sneppen (BS) in 1993 [8]. This model is the prototype of a wide class of SOC models, characterized by a deterministic dynamics in a medium with quenched disorder. The quenched noise is independent of time and represents the disordered environment where the system evolves. BS model is defined by a discrete set of L species arranged on an one-dimensional network. Each species i is defined by a “fitness” x_i given by a real number in the interval $(0,1)$. Time evolution is discretized and at any time step the species with the lowest fitness is removed from the set/citenote. Also, at the same time the two species neighbors are removed. This should model a food web where extinction of one species affects also the survival of predation and prey species. Three new species with randomly extracted new fitness then enter the system. Hereafter, we shall call this process an “updating” of the species fitnesses. After a transient period, the system reaches spontaneously a stationary state characterized by two main features: (i) the fitnesses are uniformly distributed between a threshold value x_c and 1; (ii) the dynamics evolves as a sequence of *critical avalanches*, whose duration s is power-law distributed [9]: $P(s) \sim s^{-\tau}$, where $\tau = 1.07$ [10].

To study in an analytical way this kind of processes, a method based on conditional probability has been recently introduced. This method, called run time statistics (RTS), provides a powerful tool to study how the system stores information on the disorder during its evolution. This method has been also applied to the class of models derived by the invasion percolation [11,12] to compute the asymptotic be-

havior of the histogram equation. We apply here RTS to the BS model in order to find an analytical approach to solve the model.

As regards the analytical results available for the BS, the only successful approach has been the mean field approximation. Scale relations and an equation describing the hierarchy of the avalanches [13], allow us to reduce the number of independent exponents to one. A kind of ϵ expansion was also introduced to compute the critical exponents by performing an expansion around their mean field value [14]. This alternative approach allows us to compute in a perturbative way the value of the critical threshold linking it to the avalanche exponential τ . In the following sections we are going to introduce the main features of the model, the basic computation of the run time statistics, and the result of this approach with respect to the BS. A preliminary paper with some of the results has already been published in Ref. [15], here instead we are going to fully develop the derivation of such results.

II. CRITICAL DYNAMICS

The dynamics of the BS model can be viewed as a branching process of branching ratio L , if L is the number of species in the system. At every time step, one of the L quenched numbers assigned to the sites is selected to be updated. The same is done for the two nearest neighbors. It is then possible to represent this process as a treelike picture where every node represents a state from which the system can reach L possible states at the subsequent time step. Therefore, at time t there are L^t possible states that are reached through L^t different paths. Because of the deterministic nature of the model, time evolution is determined by the realization of the disorder $\{x\}$. This means that if we know the initial set of values x_i , $i = 1, \dots, L$, and the three numbers extracted at each time step then we also know *a priori* the evolution of the system. Otherwise, to give a description of the behavior of the generic system, we should consider a statistical average over the possible realizations of the quenched disorder. RTS provides an iterative algorithm to assign to each path (sequence of events) its statistical weight according to the laws of conditional probability.

We consider the system in the critical state. In this stationary critical state of BS model almost all the fitnesses lie above a threshold x_c . The distribution function of the fitnesses in the critical state, called the *histogram* $\Phi(x)$, is found to be uniform above x_c , and almost zero below x_c . The critical state is characterized by a power-law distribution of the duration s of critical avalanches: $P(s) \sim s^{-\tau}$. Critical avalanches are defined as a sequence of events $x_{min}(s) < x_c$, where $x_{min}(s)$ is *signal*, i.e., the value of the minimal fitness at time s [9]. Then a critical avalanche begins each time the signal reaches a value larger than x_c . Because of the shape of the histogram, that is of order $1/L$ for $x < x_c$ and constant for $x > x_c$. In the critical state the signal reaches most values close to x_c and the dynamics is a sequence of critical avalanches [9]. Then the site giving rise to a new avalanche, called the *initiator*, has a fitness close to x_c . This means that when the initiator is selected, all the other sites have fitnesses larger than x_c .

Since the critical avalanches are independent of each other [9], we can consider the dynamics within one generic critical avalanche. This is then a fair description of the dynamics of the system in the critical state. Let us define the set of active sites A_t as the sites covered by the avalanche, i.e., the sites whose fitness has been updated at least once since the beginning of the avalanche. If we consider the dynamics within an avalanche, the possible events at time t regard only the sites in A_t because the selection of any other site would imply the end of the avalanche. Indeed, sites not belonging to A_t have a fitness larger than x_c (by definition of A_t , their fitness remained unchanged since the first step of the avalanche). Therefore, the evolution of the avalanche can be seen as a branching process where the branching ratio is not fixed.

Without loss of generality, we take the time origin at the beginning of a critical avalanche, and the origin of the coordinates in the initiator site. At $t=0$ the initiator is the site with the smallest fitness (i.e., extremal rule) and its quenched variable (in the stationary state) has a value close to x_c . Since we are assuming that the stationary state in the system is the critical one, all the quenched numbers in the system are distributed following the stationary distribution $\Phi(x)$, apart from corrections of order $1/L$. Updating at time $t=0$ affects the initiator and the two nearest neighbors. $A_{t=1}$ is then composed by the three sites $\{-1, 0, 1\}$. The three variables corresponding to these sites are distributed following the uniform probability density $f_0(x)$. If the avalanche proceeds, there are three possible events leading to three different configurations for the system (see Fig. 1). Then a path of length 2 can be realized in the following three ways: (i) growth of the initiator at time $t=0$ and then growth of its left neighbor at time $t=1$; (ii) growth of the initiator at time $t=0$, and successive growth of the initiator at time $t=2$; (iii) growth of the initiator at time $t=0$ and then growth of the right neighbor at time $t=1$. As the length t (i.e., the number of steps) increases, the number N_t of possible paths of that length increases fast. For example, avalanches lasting two time steps can occur in the previous three ways, but there are eleven ways to form an avalanche of three time steps (see

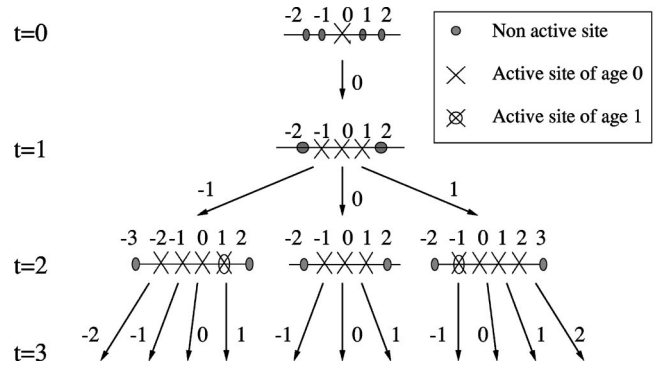


FIG. 1. Diagrammatic plot of the first three steps in an avalanche tree. The initiator $i=0$ is selected at $t=0$. At each step one has to consider all the possible offsprings. Nonactive sites are represented by a filled circle; updated active sites are represented by with a simple cross. Crossed empty circles represents instead active sites not updated at the previous time steps.

Fig. 1). In Table I we report the number N_t of possible paths whose length is t . It can be shown that N_t grows roughly as $t!$. Given this picture of the stationary state, it is evident that a description of the model evolution can be achieved through a description of the growth paths. Run time statistics (whose approach is described in the following section) helps in sorting out the most probable paths, thereby extracting the available information on the process.

III. THE RTS APPROACH

Through the RTS we are able to compute the statistical weight of all the possible paths corresponding to a critical avalanche of a fixed duration [16,17]. The RTS has been developed in order to extract the maximal information available from the knowledge of the dynamical history followed by the process. The information is stored in *effective* probability density functions (PDF's) for the variables $\{x\}$. These effective density functions at a certain time t , are used to compute the conditional probabilities of all the possible events at time $t+1$ given the state at time t . The probability of a sequence of events (that is a fixed path), is then factorized in the product of these one-step probabilities. The time-dependent PDF's are obtained by applying the laws of conditional probability. At the beginning the only information available on the disorder is the probability density $f_0(x)$ from which the quenched variables are extracted. The conditional probability laws in the following steps modify the shape of the probability distribution once the previous history is known.

In the case of BS, at every time step the smallest number is removed from the system. It is then intuitive that if the lifetime of a species, i.e., the number of tests the species

TABLE I. Number of possible paths as a function of time.

t	1	2	3	4	5	6	7	8	9
N_t	1	3	11	47	227	1215	7107	44959	305091

survived in the search for the minimum, is large, its fitness is (probably) also large. Henceforth, the longer the lifetime the higher the probability to deal with a large value of the fitness x , and the smaller the probability to be removed. This concept can be easily formalized by using the theory of conditional probability.

Let us introduce the time-dependent PDF's $f_{i,t}(x)$ giving the probability density of the quenched variable of site i at time t . If $m_{i,t}(x)$ is the probability density function of fitness at time t (assuming it corresponds to site i), we have that $m_{i,t}(x)dx$ is the probability that the fitness of site i has a value between x and $x+dx$ (given that at time t of a fixed path, i is the minimum with fitness smaller than x_c ; this last condition ensures that the system is still under the same critical avalanche). Then $m_{i,t}(x)$ is given by

$$\begin{aligned} m_{i,t}(x) &= P(x < x_i < x + dx | x_i = \min\{x\}_{A_{t-1}}, x_i < x_c) \\ &= \frac{P(x < x_i < x + dx \cap x_i = \min\{x\}_{A_{t-1}}, x_i < x_c)}{P(x_i = \min\{x\}_{A_{t-1}}, x_i < x_c)} \\ &= \frac{1}{\mu_{i,t-1}} f_{i,t-1}(x) \prod_{k \in A_{t-1} - \{i\}} \int_x^1 dx_k f_{k,t-1}(x_k) \\ & \quad x \leq x_c, x > x_c, \end{aligned} \quad (1)$$

where we have defined the one-step probability $\mu_{i,t}$ as

$$\begin{aligned} \mu_{i,t} &= P(x_i = \min\{x\}_{A_t}, x_i < x_c) \\ &= \int_0^{x_c} dx f_{i,t-1}(x) \prod_{j \in A_{t-1} - \{i\}} \int_x^1 dx_j f_{j,t-1}(x_j). \end{aligned} \quad (2)$$

$\mu_{i,t}$ represents the probability that site i has the minimal fitness at time t (smaller than x_c), given the path followed up to time $t-1$ [all the information about the past steps is included in the effective PDF's $f_{j,t}(x)$]. In both Eqs. (1) and (2) we consider only the fitnesses $\{x\}_{A_t}$ of the active sites A_t because the others actually do not participate in the dynamics, being larger than x_c . It is important to notice that

$$\sum_{i \in A_t} \mu_{i,t} < 1. \quad (3)$$

This is because of the condition that the minimum fitness is less than x_c . The complementary probability ($1 - \sum_{i \in A_t} \mu_{i,t}$) is then the probability that the minimum is larger than x_c , that is, that the avalanche stops at time t .

In a similar way we can obtain the probability densities $f_{k,t}(x)$, with $k \in A_{t-1}$,

$$\begin{aligned} f_{k,t}(x) &= P(x < x_k < x + dx | x_i = \min\{x\}_{A_{t-1}}, x_i < x_c) \\ &= \frac{P(x < x_k < x + dx \cap x_i = \min\{x\}_{A_{t-1}}, x_i < x_c)}{P(x_i = \min\{x\}_{A_{t-1}}, x_i < x_c)} \\ &= \begin{cases} \frac{1}{\mu_{i,t}} f_{k,t}(x) \int_0^x f_{i,t}(x_i) dx_i \prod_j \int_{x_i}^1 dx_j f_{j,t}(x_j) & x \leq x_c \\ \frac{1}{\mu_{i,t}} f_{k,t}(x_c) \int_0^{x_c} f_{i,t}(x_i) dx_i \prod_j \int_{x_i}^1 dx_j f_{j,t}(x_j) & x > x_c \end{cases} \end{aligned} \quad (4)$$

where $j \in A_{t-1} - \{i, k\}$. In this way the effective PDF's are conditioned to the whole history from time 0 (beginning of the avalanche) to time t because of the step-by-step algorithm through which they are obtained. We also notice that if the minimum fitness is less than x_c , this implies that $f_{i,t}(x) = f_{i,t}(x_c)$ for $x > x_c$. As it has been pointed out in Ref. [9], the dynamics involves only the quenched numbers below x_c , this is the reason why the system does not acquire information on the variables in the region $x > x_c$.

These formulas for the effective probability densities hold if we assume that the probability density of the whole set of variables $F_t(\{x\})$ could be factorized in the product of the one-variable probability densities at any time. Actually this is true only at time $t=0$, when the $\{x\}$ are uncorrelated, while at later times the extremal dynamics induces correlations among them. Nevertheless, the approximation

$$F_t(\{x\}) \approx \prod_j f_{j,t}(x) \quad (5)$$

has proven to lead to results in good agreement with data from simulations [18], and to give rise to an error that is negligible for large values of the system size L [19].

Once the one-step probabilities $\mu_{i,t}$ have been computed, the probability of a given path C_t (i.e., a fixed sequence of events from time 0 to time t) is given by

$$W(C_t) = \prod_{t'=1}^t \mu_{i,t'}. \quad (6)$$

This probability is the probability of the path C_t averaged on the disorder [20].

Let us consider what would be the rigorous computation of this quantity. The extremal rule can be formulated by defining a growth probability $\eta_i(\{x\})$ for site i to be selected, given by

$$\eta_i(\{x\}) = \prod_j \theta(x_j - x_i) = \begin{cases} 1, & x_i = \min\{x\} \\ 0, & x_i \neq \min\{x\}. \end{cases} \quad (7)$$

Then the probability of a given path C_t is given by

$$\Omega_{C_t}(\{x\}) = \prod_{t'=1}^T \eta_{i'}(\{x\}_{A_{t'}}), \quad (8)$$

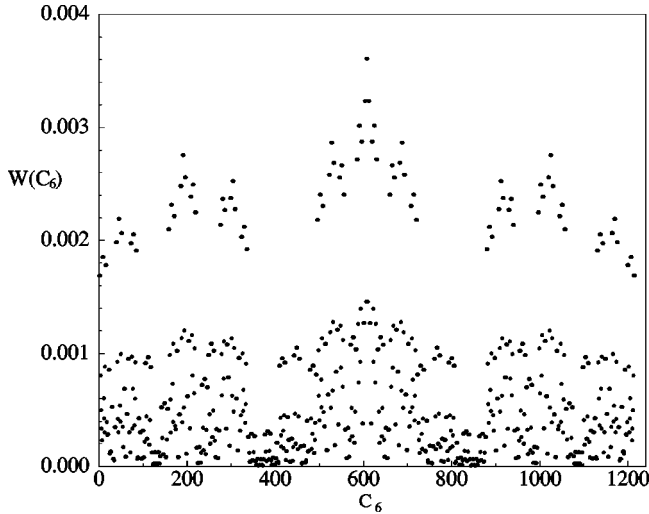


FIG. 2. Statistical weights $W(C_6)$ of the paths C_6 of length 6. Paths are numbered from 1 to N_6 following the order in a treelike diagram like the one shown in Fig. 1.

which can assume only the values 1 or 0. The exact computation of the *mean* probability of the path C_t consists in taking the average over the realizations of the disorder. This average cannot be factorized: this means that it is not possible to obtain the mean probability of a path C_t by simply multiplying the mean probability of the path C_{t-1} , corresponding to the first $t-1$ steps of path C_t , for the mean probability of the last step. On the contrary, the weight $W(C_t)$ is factorized as in Eq. (6) because the one-step probabilities $\mu_{i,t}$ are *conditional* probabilities. The weights $W(C_t)$ of the paths of length 6 are plotted in Fig. 2. In Fig. 3 are plotted the effective probability densities $f_{i,t}(x)$ of a site i for different values of t : the site chosen is not updated from time $t=1$ to time $t=4$, the corresponding PDP is then modified in such a way to approach the histogram shape.

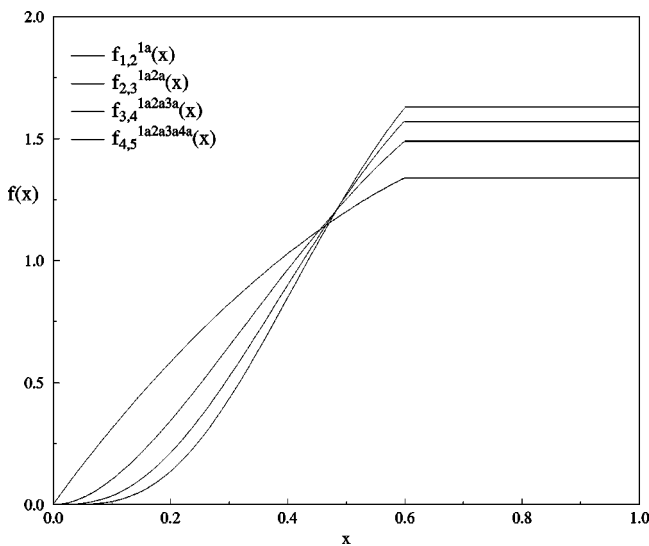


FIG. 3. Effective probability density of the site 1 at different times of the path 1a2a3a4a. The probability densities are labeled with the indexes τ, t giving the “age” of the quenched number, and the time step of the path.

IV. THE HISTOGRAM EQUATION

We now introduce an equation for the histogram $\Phi_t(x)$, that represents the average probability density of the quenched variables at time t . In the limit $t \rightarrow \infty$ we obtain an equation for the stationary histogram. Let us introduce the function $h_t(x)$ defined as

$$h_t(x) = L\Phi_t(x), \quad (9)$$

where L is the size of the system. Then $h_t(x)dx$ is the average number of quenched variables in the system in the interval $[x, x+dx]$ at time t . Since the species updated at each time step is the minimal one with the two nearest neighbors, we can write a balance equation

$$h_{t+1}(x) = h_t(x) - m_{t+1}(x) - [f_1(x) + f_2(x)] + 3f_0(x), \quad (10)$$

where $m_t(x)$ is the probability density of the minimal variable; $f_{1,2}(x)$ are the probability densities that the two nearest neighbors variables would have at time $t+1$ if they were not updated. The PDF's $f_{1,2}(x)$ are given by the first line of Eq. (4) with $k=i-1$ and $k=i+1$, respectively. Because of the self-averaging property of this function, the result obtained by taking the asymptotic limit coincides with the one obtained by averaging over the possible realizations of the disorder:

$$\lim_{t \rightarrow \infty} h_t(x) = L\Phi(x) = L\langle \Phi(x) \rangle_{\{x\}}. \quad (11)$$

We then obtain

$$\langle m(x) \rangle + \langle f_1(x) + f_2(x) \rangle - 3 = 0, \quad (12)$$

where $f_0(x) = 1$.

To compute the average $\langle f_1(x) + f_2(x) \rangle$ we use the weights $W(C_t)$ obtained by applying the RTS algorithm: $\langle f_1(x) + f_2(x) \rangle$ is given by averaging over the paths C_t , weighed with the $W(C_t)$

$$\langle f_1(x) + f_2(x) \rangle = \frac{\sum_{t=1}^{\infty} \sum_{C_t} W(C_t) [f_1^{C_t}(x) + f_2^{C_t}(x)]}{\sum_{t=1}^{\infty} \sum_{C_t} W(C_t)}. \quad (13)$$

To compute this quantity to the order n , we perform an exact enumeration of the paths of length $t \leq n$. For each path C_t , $f_{1,2}^{C_t}(x)$ are computed by iterating the formulas in Eq. (4) to obtain the effective probability densities at time $t+1$ of the nearest-neighbor sites of the site selected to grow at time t of the path C_t . The sum in Eq. (12) contains also terms proportional to $\Phi(x)$. This happens when one of the nearest neighbors of the extremal site does not belong to the set of active sites, its probability density is $\Phi(x)$. Moreover, x_c appears explicitly in Eq. (13) because both the probability densities

$f_{1,2}(x)$ and the weights $W(C_t)$ depend on it (since x_c has been introduced as a parameter by imposing that the system is under an x_c avalanche).

The minimum probability density is averaged in a different way, following a mean field argument. We now consider the generalized run time statistics [19], which gives the correct form for the effective probability densities in the case of a stochastic dynamics. Let us suppose to have a growth probability $\eta(\{x\})$ depending on the quenched disorder different from extremal rule. Also in this case the system stores information during the evolution. The Eq. (4) must now be modified in order to take into account the probability $\eta_i(\{x\})$. The one-step probability is obtained in the following way:

$$\mu_{i,t} = \int_0^1 f_{i,t}(x_i) dx_i \int_0^1 \cdots \int_0^1 \prod_{j \in A_t - \{i\}} dx_j f_{j,t}(x_j) \eta_i(\{x\}_{A_t}), \quad (14)$$

where now A_t is given by all the sites in the system. The probability density of the minimum site is consequently given by

$$m_{i,t+1}(x) = \frac{1}{\mu_{i,t}} \int_0^1 dx_i \delta(x_i - x) f_{i,t}(x_i) \times \int_0^1 \cdots \int_0^1 \prod_k \eta_i(\{x\}_{A_t}) f_{k,t}(x_k) dx_k, \quad (15)$$

where i_t is the quenched number of the site selected at time t .

We now consider the growth probability given by

$$\eta_i(\{x\}_{A_t}) = \frac{e^{-x_i/T}}{\sum_{j \in A_t} e^{-x_j/T}}. \quad (16)$$

In this way in the limit $T \rightarrow 0$ we recover the extremal rule

$$\lim_{T \rightarrow 0} \frac{e^{-x_i/T}}{\sum_{j=1}^L e^{-x_j/T}} = \begin{cases} 1, & x_i = \min\{x\} \\ 0, & x_i \neq \min\{x\}. \end{cases} \quad (17)$$

We then substitute the expression Eq. (16) in Eq. (15) to realize the average over the disorder and then the limit $T \rightarrow 0$.

We average on the paths C_t of length t and then consider the limit $t \rightarrow \infty$. The average on the paths C_t is taken by averaging first on the last step and then on the paths C_{t-1} ,

$$\langle m_{i,t}(x) \rangle_{C_t} = \langle \langle m_{i,t}(x) \rangle_{i_t} \rangle_{C_{t-1}}. \quad (18)$$

The first average gives

$$\langle m_{i,t}(x) \rangle_{i_{t-1}} = \sum_{i=1}^L \mu_{i,t} \left(\frac{1}{\mu_{i,t}} f_{i,t}(x) \int_0^1 \cdots \int_0^1 \prod_{j \neq i} dx_j f_{j,t}(x) \frac{e^{-x/T}}{e^{-x/T} + \sum_{k \neq i} e^{-x_k/T}} \right).$$

Then, averaging over the paths C_{t-1} we obtain

$$\begin{aligned} \langle \langle m_{i,t}(x) \rangle_{i_t} \rangle_{C_{t-1}} &= \sum_{i=1}^L \left\langle f_{i,t}(x) \int_0^1 \cdots \int_0^1 \prod_{j \neq i} dx_j f_{j,t}(x) \frac{e^{-x/T}}{e^{-x/T} + \sum_{k \neq i} e^{-x_k/T}} \right\rangle \\ &\approx \sum_{i=1}^L \langle f_{i,t}(x) \rangle \int_0^1 \cdots \int_0^1 \prod_{j \neq i} dx_j \langle f_{j,t}(x) \rangle \frac{e^{-x/T}}{e^{-x/T} + \sum_{k \neq i} e^{-x_k/T}}. \end{aligned} \quad (19)$$

Taking the limit $t \rightarrow \infty$ we have $\langle f_{j,t}(x) \rangle = \Phi(x)$,

$$\begin{aligned} \langle m(x) \rangle &= L \Phi(x) \int_0^1 \cdots \int_0^1 \left[\prod_{j \neq i} dx_j \Phi(x_j) \right] \\ &\times \frac{e^{-x/T}}{e^{-x/T} + \sum_{k \neq i} e^{-x_k/T}}. \end{aligned} \quad (20)$$

Let us consider the integral in the right-hand side of Eq. (20): it is the average of the function

$$g(\{x_k\}) = \frac{e^{-x/T}}{e^{-x/T} + \sum_{k \neq i} e^{-x_k/T}},$$

which is a function of the $(L-1)$ variables $\{x_k\}_{k \neq i}$. If we put $z_j = e^{-x_j/T}$ and $Z = \sum_k z_k$, according to the central limit theorem the deviation from the mean value $\langle Z \rangle = (L-1)\langle z \rangle$ is negligible. We can then make the approximation $\langle F(Z) \rangle \approx F(\langle Z \rangle)$,

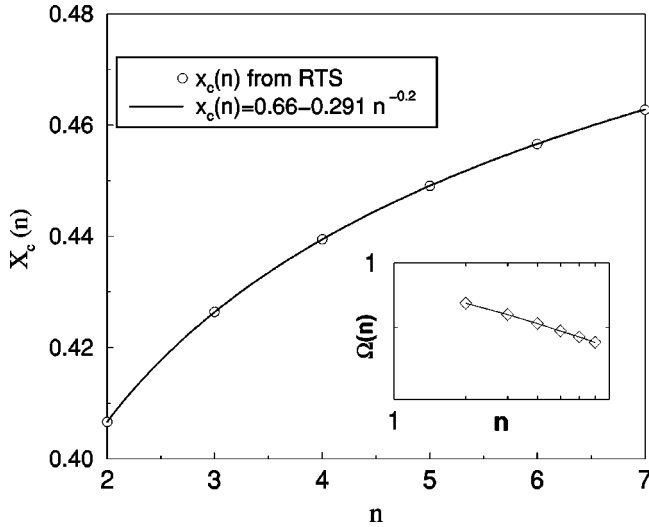


FIG. 4. Empty points represent the values of x_c obtained by the application of the RTS algorithm from $n=2$ to $n=7$. The continuous line represents the fit curve $x_c(n)=0.66-ax^b$ with $a=0.291 \pm 0.003$ and $b=0.20 \pm 0.03$. The inset shows the behavior of $\Omega(n)$ up to $n=7$. Assuming $\Omega(n) \sim n^{-\tau+1}$ as a good approximation also for small n , one finds $\tau \approx 1.05$.

$$\int_0^1 \cdots \int_0^1 \prod_{j \neq i} dx_j \Phi(x_j) \frac{e^{-x/T}}{e^{-x/T} + \sum_{k \neq i} e^{-x_k/T}}$$

$$\approx \frac{e^{-x/T}}{e^{-x/T} + (L-1)\langle e^{-x_k/T} \rangle}.$$

We now introduce the parameter x_c by defining $\langle z \rangle = \langle e^{-z/T} \rangle = e^{-x_c/T}$. We finally obtain

$$\langle m(x) \rangle = L\Phi(x) \frac{1}{1 + (L-1)e^{-(x_c-x)/T}}. \quad (21)$$

We can see that if we assume for $\Phi(x)$ the behavior

$$\Phi(x) = \begin{cases} O\left(\frac{1}{L}\right), & x \leq x_c \\ O(1), & x > x_c, \end{cases} \quad (22)$$

then the average of the minimum density function given above has the expected behavior, in the limit $T \rightarrow 0$,

$$\langle m(x) \rangle = \begin{cases} O(1), & x \leq x_c \\ 0, & x > x_c, \end{cases} \quad (23)$$

Thus the parameter x_c introduced above is the critical threshold in the histogram function.

We can now turn to Eq. (12). After use of Eqs. (21) and (13), the equation takes the form

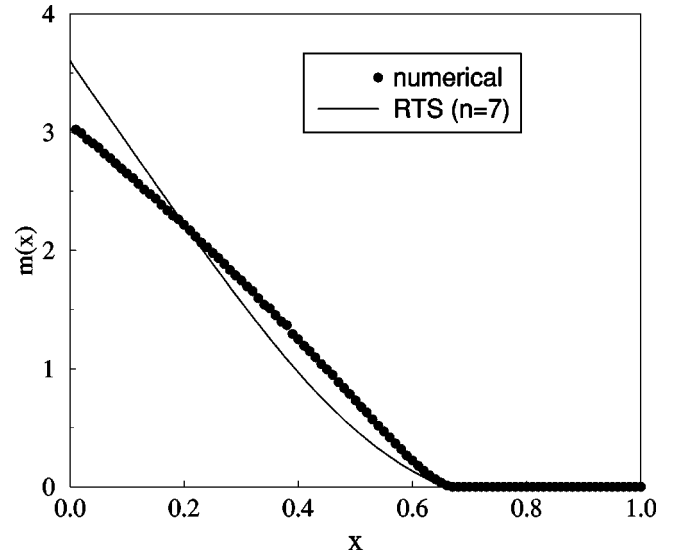


FIG. 5. The continuous line gives the stationary distribution $m(x)$ of the minimal fitnesses evaluated through the application of the RTS algorithm, considering all the possible avalanche paths up to a maximal time $n=7$ after the selection of the initiator, and assuming $x_c(n \rightarrow \infty) = 0.66$. The points give the numerical behavior evaluated in extensive simulations.

$$L\Phi(x) \frac{1}{1 + (L-1)e^{-(x_c-x)/T}} + \frac{A^{(n)}(x_c)\Phi(x) + B^{(n)}(x, x_c)}{D^{(n)}(x_c)} - 3 = 0, \quad (24)$$

where the coefficients $A^{(n)}(x_c)$, $B^{(n)}(x, x_c)$, and $D^{(n)}(x_c)$ are given by Eq. (13) with truncation at order n . There is only one value of x_c for which $\Phi(x)$ is normalized. This value gives the value of the BS threshold.

It is easy to verify that the first order of approximation, $n=1$, corresponds to the mean field approximation. Indeed, the first order of Eq. (13) is

$$\langle f_1(x) + f_2(x) \rangle^{(1)} = 2\Phi(x). \quad (25)$$

This is essentially the random neighbors assumption since in this way all the correlations among the species in the avalanche are neglected. By substituting Eq. (25) in Eq. (24) we obtain

$$\Phi(x) = \frac{3}{2 + \frac{L}{1 + (L-1)e^{-(x_c-x)/T}}} \quad (26)$$

that in the limit $T \rightarrow 0$ becomes

$$\Phi(x) = \begin{cases} \frac{3}{L}, & x \leq x_c \\ \frac{3}{2}, & x > x_c. \end{cases} \quad (27)$$

Then, imposing the normalization condition

$$\int_0^1 \Phi(x) dx = 1 \quad (28)$$

we obtain $\frac{3}{2}(1-x_c)=1$, verified for $x_c=\frac{1}{3}$, which is the value obtained in the mean field case [13]. It is worth noticing that in Eq. (27) we have analytically obtained the behavior $\Phi(x)\sim 1/L$ under x_c .

We solved numerically Eq. (12) from order $n=2$ to order $n=7$. The values of x_c obtained for different values of n are plotted in Fig. 4. By considering part of the correlations among species, x_c becomes larger than the one obtained in the mean field approximation ($x_c=1/3$). The best evaluation is $x_c(n=7)\simeq 0.465$, larger than the mean field result but still quite far from the value obtained from simulations. Nevertheless, it is possible to verify that the behavior of $x_c(n)$ is compatible with an asymptotic value $x_c(n\rightarrow\infty)\simeq 0.66$. We made a fit with the fitting function $x_c(n)=0.66-ax^b$ that is found to be well compatible with the given asymptotic value (see Fig. 4). The fit values are $a=0.291\pm 0.003$ and $b=0.20\pm 0.03$. The small value of b is due to the fact that the avalanche duration distribution $P(s)$ is characterized by a small exponent ($\tau\simeq 1.07$), henceforth all the sizes s are important for statistics.

One can use $x_c(n\rightarrow\infty)$ to evaluate both the avalanche exponent τ and the average minimum distribution $m(x)\equiv\langle m_i(x)\rangle$. The exponent τ can be obtained from the function $\Omega(t)=\sum_{C_t} W(C_t)$ for t ranging from 1 to the maximal

possible n . This function is proportional to the probability that the avalanche lasts at least t time steps. Thus, in the scaling regime, $\Omega(t)\sim t^{-\tau+1}$. Making this hypothesis, and substituting the value $x_c(n\rightarrow\infty)$ in the expressions giving the weights $W(C_t)$, one finds $\tau(n=7)\simeq 1.05$ (see Fig. 4), which is in agreement with the known numerical value.

Finally, we can obtain an approximation of the probability density function of the minimal fitness $m(x)$ in the stationary state. Let us consider the Eqs. (21) and (24). If we take the limit $T\rightarrow 0$ in Eq. (21) we obtain

$$\langle m(x)\rangle = L\Phi(x)\theta(x_c-x). \quad (29)$$

If we now turn to Eq. (24) and we solve it for $\Phi(x)$ using the value $x_c(n\rightarrow\infty)$, we obtain the function $m(x)$ reported in Fig. 5. In the same figure, this result is compared with the numerical distribution of the minimal fitnesses obtained in the numerical simulations. In spite of the strong approximation (the paths considered are only those of length ≤ 7), the agreement is quite good.

In conclusion, this paper presents a perturbative approach to the BS model, based on the probabilistic framework called run time statistics. The detailed derivation of the self-organized threshold x_c , the avalanche exponent τ , and the stationary distribution of minimal fitnesses $m(x)$ is presented here. Through RTS we are able to improve the agreement between the numerical and the theoretical values found for this model.

-
- [1] I. Rodriguez-Iturbe and A. Rinaldo, *Fractal River Basins, Chance and Self-Organization* (Cambridge University Press, Cambridge, UK).
- [2] G. Caldarelli, A. Giacometti, A. Maritan, I. Rodriguez-Iturbe, and A. Rinaldo, *Phys. Rev. E* **55**, R4865 (1997).
- [3] A. Petri, G. Paparo, A. Vespignani, A. Alippi, and M. Costantini, *Phys. Rev. Lett.* **73**, 3423 (1994).
- [4] G. Caldarelli, F.D. Di Tolla, and A. Petri, *Phys. Rev. Lett.* **77**, 2503 (1996).
- [5] Z. Olami, H.J.S. Feder, and K. Christensen, *Phys. Rev. Lett.* **68**, 1244 (1992).
- [6] P. Bak, C. Tang, and K. Wiesenfeld, *Phys. Rev. Lett.* **59**, 381 (1987); *Phys. Rev. A* **38**, 364 (1988).
- [7] H. Jensen, *Self-Organised Criticality* (Cambridge University Press, Cambridge, UK, 1998).
- [8] P. Bak and K. Sneppen, *Phys. Rev. Lett.* **71**, 4083 (1993).
- [9] M. Paczuski, S. Maslov, and P. Bak, *Phys. Rev. E* **53**, 414 (1996).
- [10] P. Grassberger, *Phys. Lett. A* **200**, 277 (1995).
- [11] D. Wilkinson and J.F. Willemsen, *J. Phys. A* **16**, 3365 (1983).
- [12] R. Cafiero, G. Caldarelli, and A. Gabrielli, *Phys. Rev. E* **56**, R1291 (1997); A. Gabrielli, R. Cafiero, and G. Caldarelli, *J. Phys. A* **31**, 7429 (1998).
- [13] J. de Boer, B. Derrida, H. Flyvbjerg, A.D. Jackson, and T. Wettig, *Phys. Rev. Lett.* **73**, 906 (1994).
- [14] M. Marsili, P. De Los Rios, and S. Maslov, *Phys. Rev. Lett.* **80**, 1457 (1998).
- [15] M. Felici, G. Caldarelli, A. Gabrielli, and L. Pietronero, *Phys. Rev. Lett.* **86**, 1896 (2001).
- [16] L. Pietronero and W.R. Schneider, *Physica A* **119**, 249 (1990).
- [17] M. Marsili, *J. Stat. Phys.* **77**, 733 (1994); A. Gabrielli, M. Marsili, R. Cafiero, and L. Pietronero, *ibid.* **84**, 889 (1996).
- [18] R. Cafiero, A. Gabrielli, and M. Marsili, *Phys. Rev. E* **55**, 7745 (1997); A. Gabrielli, R. Cafiero, M. Marsili, and L. Pietronero, *Europhys. Lett.* **38**, 491 (1997).
- [19] A. Gabrielli (unpublished).
- [20] It is, in principle, possible to define a more general stochastic dynamics by assigning to each site of the lattice a probability to be selected for updating that depends on the disorder, that is, on the set of numbers $\{x_i\}$, $i=1, \dots, L$. The extremal rule corresponds to the particular case in which this probability is 1 for the site with minimal fitness in the system, and 0 for all the others. In such a case, the dynamics is deterministic. In general, the main feature of the models with quenched disorder is the presence of memory effects.