Bak-Sneppen model: Local equilibrium and critical value

Daniel Fraiman[*]

Departamento de Matemática y Ciencias, Universidad de San Andrés, 1644 Buenos Aires, Argentina, and CONICET, Argentina

The Bak-Sneppen (BS) model is a very simple model that exhibits all the richness of self-organized criticality theory. At the thermodynamic limit, the BS model converges to a situation where all particles have a fitness that is uniformly distributed between a critical value p_c and 1. The p_c value is unknown, as are the variables that influence and determine this value. Here we study the BS model in the case in which the lowest fitness particle interacts with an arbitrary even number of m nearest neighbors. We show that p_c verifies a simple local equilibrium relation. Based on this relation, we can determine bounds for p_c of the BS model and exact results for some BS-like models. Finally, we show how transformations of the original BS model can be done without altering the model's complex dynamics.

I. INTRODUCTION

In the early 1990s, Per Bak and Kim Sneppen introduced one of the most elegant dynamical models of evolution. The model has attracted the attention of numerous physicists, mathematicians, and biologists. The Bak-Sneppen (BS) evolution model is defined in the following way: There exist N particles, sites, or species in a one-dimensional ring, and each site k is characterized by a quantity X_k, called fitness, which evolves by

$$X_k(t+1) = \begin{cases} X_k(t) & \text{if } d(k,\tilde{k}_t) > a \\ U_{k,t} & \text{if } d(k,\tilde{k}_t) \leqslant a \end{cases}, \qquad (1)$$

where $\tilde{k}_t = \{k : \ X_k(t) \leqslant X_j(t) \ \forall j \in \{1,2,\ldots,N\}\}$ is the particle with the lowest X value at time t. The distance between two particles i and j is $d(i,j) = \min(|i-j|,|i+j-N|)$, just in order to have a ring configuration (periodic boundaries conditions); $U_{k,t}$ are independent and identically distributed random variables with uniform distribution $(0,1)$; and, finally, $a \in \mathbb{N}$ is the number of neighbors on each side that are interacting with any given particle. The initial condition is uniform, i.e., $X_k(0) = U_{k,0}$ for all particles. In Ref. [1] Bak and Sneppen introduced the model for the case $a = 1$ and showed that this extremely simple model, which can be elegantly applied to the evolution of species, exhibits self-organized criticality. Once the system has reached the stationary regime, an unexpected behavior appears in the model. At the thermodynamic limit, all particles appear with a fitness value that is distributed uniformly between a critical value p_c and 1, and there are avalanches of particle extinction. More formally, let $\tilde{X}(t) := X_{\tilde{k}(t)}(t)$ be the lowest fitness value at time t which occurs at site $\tilde{k}(t)$. An avalanche is a succession of events where the lowest fitness value is less than p_c. It starts at time $t+1$ if $\tilde{X}(t) \geqslant p_c$ and $\tilde{X}(t+1) < p_c$ and has a duration τ if $\tilde{X}(t+1) < p_c$, $\tilde{X}(t+2) < p_c, \ldots, \tilde{X}(t+\tau) < p_c$, and $\tilde{X}(t+\tau+1) \geqslant p_c$. This sequence of minimum fitness $\tilde{X}(t+1), \tilde{X}(t+2), \ldots,$

is a dependence sequence which makes calculating the τ distribution $[P(\tau = s)]$ very difficult. Only the first values of the distribution can be easily computed. Nevertheless, it is well known that the mean avalanche duration is infinity, $\langle \tau \rangle = \infty$, due to the power-law tail distribution (fingerprint of criticality) [1–10].

The critical value p_c only depends on the unique parameter of the model, a, described in Eq. (1) which determines the number of interacting neighbor sites, $m := 2a$, that are updated at each temporal step. Although there have been large efforts to calculate the value p_c, at least for $m = 2$, there is no precise result 25 years after the model's introduction. Simulation results show that $p_c(m = 2)$ is approximately 0.667 [3,4]. In this paper, we study the value p_c for different values of m by presenting a novel local equilibrium property. This property relates p_c with the neighbors of the lowest fitness particle, which allows one to obtain nontrivial bounds for p_c. These last results are presented in the Sec. II. In Sec. III we study the BS model under different modifications on the interaction criteria (two nearest neighbors) and on the update procedure (uniform). In all cases we advance the description of the corresponding critical value and the average fitness.

II. BAK-SNEPPEN MODEL

In this section we study the BS model introducing a local equilibrium equation and bounds for p_c.

A. Local equilibrium

The strategy adopted here for studying p_c is to focus on the generator of the avalanche, the lowest fitness site, and also its neighbors. Figure 1(a) shows the fitness of each particle in the stationary regime. A system of 3000 particles with 4 neighbors ($m = 4$) was simulated and a snapshot is shown. The lowest fitness particle (red circle) is responsible for producing avalanches of very large duration that diffuse very slowly. This behavior was named by Bak and Sneppen *punctuated equilibrium*. The avalanches can be interpreted as the "punctuations" that *maintain* equilibrium. Clearly, to maintain equilibrium, the

---

*[*]dfraiman@udesa.edu.ar
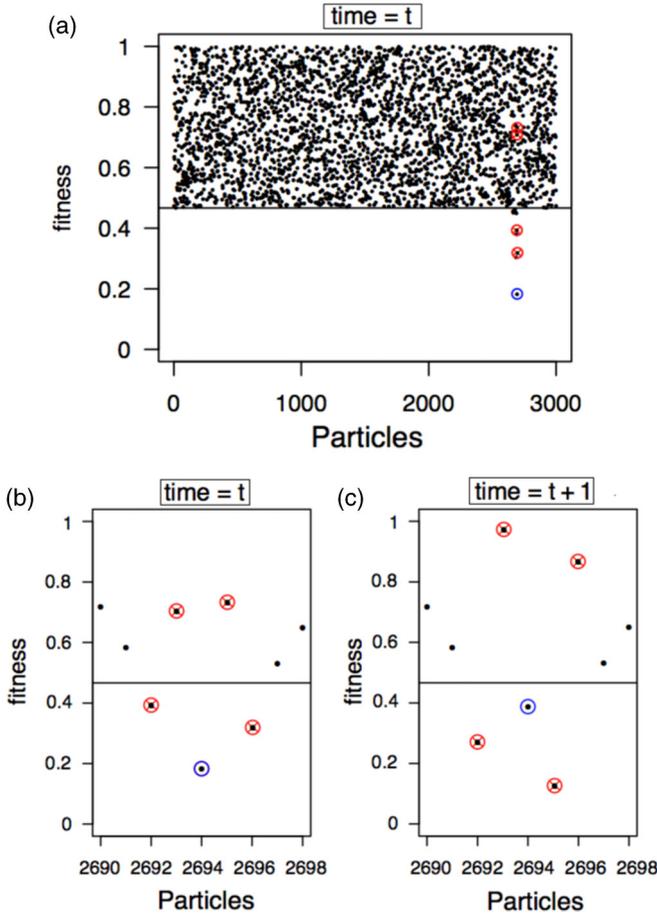
FIG. 1. (a) Fitness of the $N = 3000$ particles evolved by the Bak-Sneppen model with $m = 4$. (b) Zoom-in of the particles near the lowest fitness particle (2694) at time t. The number of neighbors of the lowest particle with fitness below p_c is $S_t = 2$. (c) The same particles shown in panel (b) at time $t + 1$. In this case, the lowest particle is the number 2695 and $S_{t+1} = 1$. The lowest fitness particle at time t is shown with a blue circle, and the m neighbors of the lowest site are shown with red cross circles.

number of particles with a fitness value smaller than p_c during the avalanche (known as "active particles") must be stable. During the avalanche, there must be an equilibrium where the net flux of particles crossing (up or down) p_c is zero. If this is not true, then the avalanche will grow indefinitely (and no critical behavior will be observed) or disappear very quickly, increasing the fictitious value p_c. With this idea in mind, next we write a conserved mass equation with a permeable wall at $X = p_c$.

The lowest fitness particle affects the fitness of its m nearest interacting neighbors [$m/2$ each side, see Figs. 1(b) and 1(c)]. Previous to the fitness update, some of these m neighbors are below p_c and some are above. The lowest fitness particle can also have a fitness value larger or smaller than p_c. The latter is the most likely, as it occurs with probability $q = \frac{\langle \tau \rangle}{\langle \tau \rangle + (1 - (1 - p_c)^{m+1})^{-1}}$, while the former occurs with probability $1 - q$. This probability q is calculated as usual for a two state (active or avalanche state and inactive or "freezed" state) model, q is the fraction of time spent in the avalanche state. That is, $q = \frac{\langle \tau_A \rangle}{\langle \tau_A \rangle + \langle \tau_I \rangle}$, where $\tau_A = \tau$ and τ_I represent the time durations in the avalanche and inactive states, respectively.

Once the system arrive to the inactive state in order to get out of this state at least one of the $m + 1$ updated particles must have a fitness below p_c, and this occurs with probability $\tilde{p} := 1 - (1 - p_c)^{m+1}$ independent of the previous update. Therefore, τ_I verifies to be a geometric random variable with parameter $\tilde{p}$, and that is why $\langle \tau_I \rangle = \tilde{p}^{-1}$. In order for the system to maintain equilibrium, the proportion of particles below p_c must be preserved (i.e., the proportion before the update must be equal to the proportion after the update). Since updates are uniform, this last proportion is just p_c, and the proportion before the update can be easily written using the law of total probability. This preservation gives rise to the following equation:

$$p_c = q \frac{1 + \langle S \rangle}{1 + m} + (1 - q) \frac{0}{1 + m}, \quad (2)$$

where $\langle S \rangle$ is the mean number of interacting neighbors that have a fitness value below p_c when the lowest fitness particle is below p_c. The first right numerator $1 + \langle S \rangle$ corresponds to the number of particles that have a fitness below p_c from a total of $1 + m$ particles, knowing that the $\tilde{X}$ is below p_c. The second numerator term is zero because it corresponds to the case where $\tilde{X}$ is greater or equal to p_c and therefore none of the neighbors can have a fitness value below p_c. Finally, since at the thermodynamic limit $\langle \tau \rangle = \infty$ ($q = 1$), we obtain

$$p_c = \frac{1 + \langle S \rangle}{1 + m}. \quad (3)$$

Another way to think of Eq. (3) is the following: Let us suppose we have a permeable wall at $X = p_c$. At the thermodynamic limit and at the steady state, the lowest fitness particle, $\tilde{k}$, will be (with probability 1) below p_c. Some of the m neighbors of $\tilde{k}$ can be below p_c and some above it. The number of neighbors that are below are equal to $1 + \langle S \rangle$, and some of these will cross up the barrier. On average, $(1 + \langle S \rangle)(1 - p_c)$ will cross up the p_c barrier. On the other hand, above p_c there are $(m - \langle S \rangle)$ particles and on average, $(m - \langle S \rangle) p_c$ will cross down the barrier. If we equal both the number of up cross and down cross particles, imposing that the system maintain equilibrium, we obtain Eq. (3).

It is important to mention that the local equilibrium equation [Eq. (2)] is valid for any connected graph, not just for the ring configuration. In the general case, each site has some particular connectivity pattern and this linking structure determines $\langle S \rangle$. If one is able to calculate $\langle S \rangle$, then the p_c value is obtained. In the next section we will use this equation to obtain bounds for p_c from Bak-Sneppen-like models that only differ in the connectivity between sites.

B. Bounds for p_c

Now, based on Eq. (2), we show that the critical value of the Bak-Sneppen model can be bounded by the critical value of two different models. These models present a slight modification of the original BS model (Fig. 2). The modification is the following: Once the $m + 1$ particles of the BS model are updated, we proceed to change the position of the updated particles.

1. Lower bound

In one case, the m updated neighbors are exchanged with other m random particles. This modification makes the model
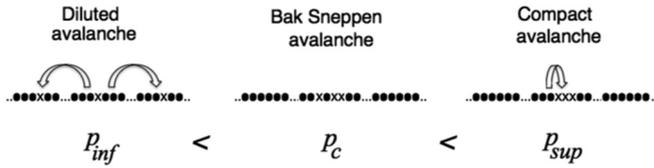
FIG. 2. Scheme of the three models studied: random neighbors (left), Bak-Sneppen (middle), and compact neighbors (right). Crosses represent active particles and points represent inactive ones.

more tractable. This model also presents a critical value, called here p_{inf}, that can be calculated exactly and has been previously studied under the names of modified or random BS model in Refs. [11,12]. In Ref. [11] the authors calculated the value p_{inf}, obtaining Eq. (4), by using mean-field theory,

$$p_{\mathrm{inf}} = \frac{1}{1+m}. \tag{4}$$

Here we show that the critical value for this model can be obtained in a simpler way by using Eq. (3). This model gives rise to avalanches that are named here "diluted" (left panel, Fig. 2) also with infinite mean duration, $\langle \tau \rangle = \infty$ ($q = 1$). At the thermodynamic limit, the neighbors selected by the lowest fitness particles previous to the update are always above p_{inf} (since the mean number of active particles at any time is finite) obtaining in this case $\langle S \rangle = 0$. Therefore replacing q by 1 and $\langle S \rangle$ by 0 in Eq. (3) we obtain Eq. (4).

Note that for this model we can also use a branching process argument to obtain the same result. Let $Z(t)$ equal the number of active sites at time t, i.e., $Z(t) = \#\{i : X_i(t) < p_{\mathrm{inf}}\}$. If we start from a unique "active particle," $Z(1) = 1$, then at the next discrete time point there may be $0, 1, 2, \ldots,$ or $m + 1$ active particles (born from the first particle that died). This process continues, and each offspring can in turn have anywhere between zero and $m + 1$ offspring. Since we are studying the system in the stationary regime and at the thermodynamic limit, the probability of selecting the same offspring twice before producing more offspring is zero [13], then we have a true branching process. It is well known that a branching process is critical if the expected number of offspring is equal to 1. Therefore, in terms of the Bak-Sneppen model, we obtain the following equation: $(m + 1)p_{\mathrm{inf}} = 1$, which is equivalent to Eq. (4).

2. Upper bound

Now we introduce the model that gives rise to an upper bound for p_c. In this case, nonconsecutive active (below the critical value, now called p_{sup}) particles are rearranged so that there are no inactive particles among active ones. In this case, "trapped" inactive particles are moved to the border between active and inactive particles. That is why we say that the model generates "compact" avalanches (right panel, Fig. 2). As far as we know this model has never before been investigated.

In the thermodynamic limit and at the steady state, particles have a fitness value that is uniformly distributed between p_{sup} and 1 ($U[p_{\mathrm{sup}}, 1]$). Unlike the two previous models, the duration of the avalanches (τ) follows an exponential distribution. In this case $\langle S \rangle$ cannot be easily calculated. Nevertheless, we found a superior bound for p_{sup}.

We show how to calculate an upper bound for p_{sup} for the case $m = 2$. Let Z_t be the number of active particles (below p_{sup}) at time t with $Z_0 = 1$, and let $Z = \lim_{t \to \infty} Z_t$ be the stationary version of the process with a mean value $\langle Z \rangle$. In the stationary condition, any of the $Z = k$ particles that are below p_c can have the lowest fitness, i.e., all have probability $1/k$ of being the lowest fitness particle ($\tilde{k}$). Now, since these k particles are all together (compact), then the number of neighbors of lowest particle, S, can be 1 or 2 (or 0 if $k = 1$). The value 1 corresponds to $\tilde{k}$ at the edge of active and inactive particles, while the value 2 corresponds to $\tilde{k}$ somewhere "inside." Therefore, the mean number of neighbors of lowest particle values verifies

$$\langle S \rangle = \sum_{k \geqslant 2} \left(\frac{2}{k} 1 + \left(1 - \frac{2}{k}\right) 2 \right) P(Z = k)$$

$$= 2 - 2\left\langle \frac{1}{Z} \right\rangle < 2 - 2\frac{1}{\langle Z \rangle}. \tag{5}$$

Note that if the probability law of Z [$P(Z = k)$] is known, then no upper bound for $\langle S \rangle$ is needed. For $m > 2$ the calculation is straightforward and we obtain:

$$p_{\mathrm{sup}} = q\frac{1 + \langle S \rangle}{1+m} < \frac{1 + \langle S \rangle}{1+m} < \frac{1 + m - m(2 + m)(4\langle Z \rangle)^{-1}}{1+m}. \tag{6}$$

Unfortunately, we do not know how to calculate $\langle Z \rangle$, but we believe it can be calculated since one advantage of this last model is that Z_t can be expressed by a simple birth and death equation.

C. The fitness of the neighbors of $\tilde{k}$

What else can we say about $\langle S \rangle$ for the Bak-Sneppen model? As mentioned above, S is the number of interacting neighbors that have a fitness value below p_c when the minimum fitness particle is below p_c. One can see that since each neighbor is independent, S has a binomial distribution with parameters m and $\tilde{p}$, i.e., $P(S = k) = \binom{m}{k}\tilde{p}^k(1 - \tilde{p})^{m-k}$. Then

$$\langle S \rangle = m\tilde{p} \quad \text{with} \quad \tilde{p} = \int_0^{p_c} h(x)dx, \tag{7}$$

where $h(x)$ is the fitness probability density function (p.d.f.) of a randomly selected neighbor of the particle with the lowest fitness when this last value is less than p_c.

In order to gain some intuition about $h(x)$, we study the Bak-Sneppen at the limit opposite the thermodynamic limit ($N \to \infty$), that is, the limit of few particles. Specifically, we study the BS model for a closed system of only $m + 1$ interacting particles. The periodic boundary condition ensures that the $m + 1$ particles are always interacting. In this case, the BS model does not have a critical value p_c, but we assume an arbitrary "fictitious critical value," p^*, just for studying the neighbor's fitness distribution when the lowest fitness value is smaller than p^* [14]. We want to understand how the fitness of neighbors of the minimum fitness particle is distributed when the lowest fitness particle is below a certain value, p^*. Let us call $h^{\mathrm{closed}}(x)$ the corresponding probability density. This probability density can be easily computed from the order statistics distribution. Considering that the fitness X_i of
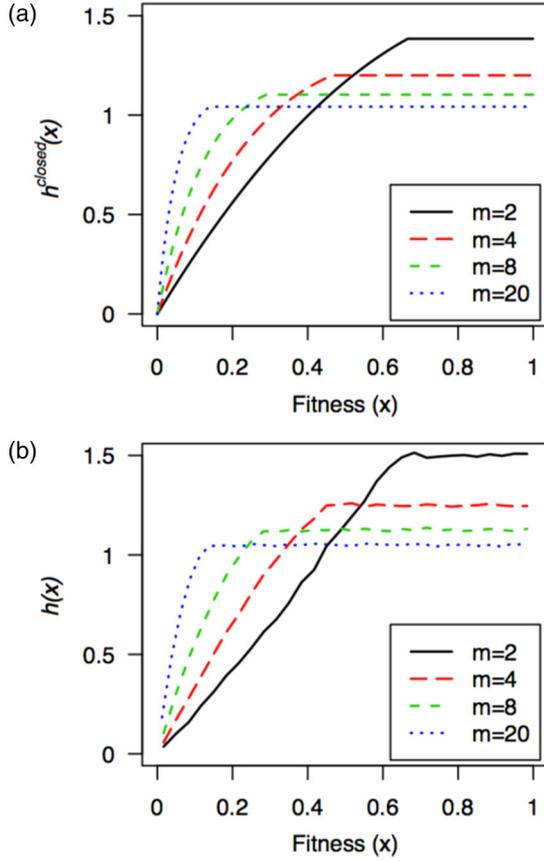
FIG. 3. Probability density of the fitness of neighbors of $\tilde{k}$ when $\tilde{k} < p_c(m)$ and when considering (a) a system of $N = m + 1$ particles and (b) a system of $N = 4000$ ($\sim$ infinity) particles. $h^{\text{closed}}(x)$ is calculated from Eq. (8). $h(x)$ is estimated by simulations.

particle i $(i = 1, 2, \ldots, m + 1)$ is a uniform $(0,1)$ random variable and that these are independent, if we call $X_{(1)}$ the lowest value, $X_{(2)}$ the following order statistic, and $X_{(m+1)}$ the maximum fitness value, then it is straightforward to verify that

$$h^{\text{closed}}(x) = \frac{1}{m \, P(X_{(1)} < p^*)} \frac{\partial}{\partial x} \left[\sum_{i=2}^{m+1} P(X_{(1)} < p^*, X_{(i)} < x) \right].$$
(8)

Finally, considering that the joint order statistics p.d.f. for uniform random variables is

$$f_{X_{(i)}, X_{(j)}}(u, v) = N! \frac{u^{i-1}}{(i-1)!} \frac{(v-u)^{j-i-1}}{(j-i-1)!} \frac{(1-v)^{N-j}}{(N-j)!}$$

with $0 \leqslant u < v \leqslant 1$ and N is the number of particles (in the closed system $N = m + 1$), it is possible to calculate $h^{\text{closed}}(x)$ from Eq. (8).

For example, for $m = 2$ neighbors, we obtain

$$h^{\text{closed}}(x) = \begin{cases} \frac{3x(2-x)}{2(1-(1-p^*)^3)} & \text{if } x < p^* \\ \frac{3p^*(2-p^*)}{2(1-(1-p^*)^3)} & \text{if } x \geqslant p^* \end{cases}.$$
(9)

Figure 3(a) shows $h^{\text{closed}}(x)$ for different values of $m = \{2, 4, 8, 20\}$ considering $p^*(m)$ as the true $p_c(m)$ observed at the thermodynamic limit.
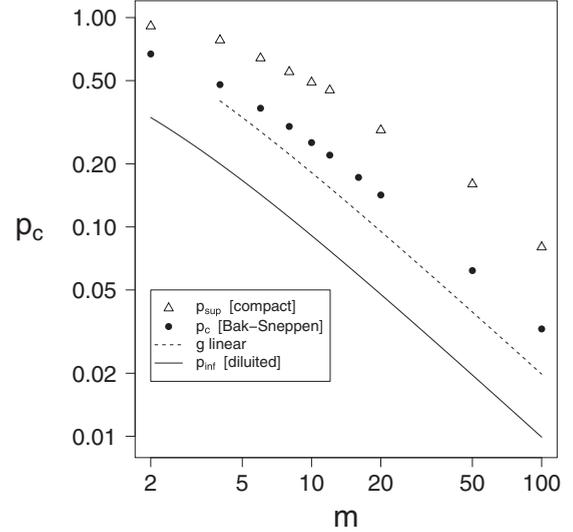
FIG. 4. Critical value of the BS model as a function of m. Upper and lower bounds are also represented.

On the other hand, for large N ($N \gg m$), an avalanche can be considered an open system. More particles can be incorporated into the avalanche as time evolves. This is not possible in the closed system where only the $m + 1$ fixed particles can be part of the "avalanche." This difference has an impact on the number of particles from which the minimum fitness is selected. Nevertheless, simulations show that $h(x)$ for large N [Fig. 3(b)] is similar to the one for $N = m + 1$ [Fig. 3(a)]. The probability density of the fitness of the neighbors, $h(x)$, is a smooth function that is partitioned into two sides and can be described as

$$h(x) = \begin{cases} g(x) & \text{if } x < p_c \\ g(p_c) & \text{if } x \geqslant p_c \end{cases},$$
(10)

with $g(0) = 0$.

For $m = 2$ the probability densities corresponding to $N = m + 1$ and to $N = \infty$ present some differences. But for $m > 2$ densities are similar. Next we focus on the case $m > 2$. In this case, $h(x)$ (or g) is a concave function, and therefore a lower bound for p_c can be obtained by proposing a linear g function. Under this hypothesis, it is straightforward to verify that

$$\frac{2}{1+m} < p_c,$$
(11)

just by using Eqs. (3) and (5).

Finally, in Fig. 4, we show the empirical critical value p_c as a function of m together with the empirical upper bound (p_{sup}) and the theoretical lower bounds [Eqs. (4) and (11)].

III. BAK-SNEPPEN-LIKE MODELS

Given the difficulty of obtaining exact results in the BS model, we study the model under different modifications that make it more manageable. First, we study the model under a nonuniform update distribution. This first result is useful for studying the subsequent models. Second, we study the random neighbors update case. Finally, we study the binary fitness model. The particle fitness notation of these BS-like models is

represented by Y. The fitness of the original BS model is still represented by X. Throughout this section we relate Y with X for each BS-like model.

A. Nonuniform fitness

First, we discuss what happens when the uniform updates assumption is broken. Let us replace U_k in Eq. (1) by $W_{k,t}$, where $W_{k,t}$ are now independent and identically distributed continuous random variables with arbitrary p.d.f. $f(w) > 0$ for $w \in \mathbb{R}$. The initial uniform condition is also replaced by $W_{k,0}$. Let $F(w) = \int_0^w f(h)dh$ be the cumulative probability function. For facilitation purposes, we use Y to denote the fitness value, and the dynamics remain as before,

$$Y_k(t+1) = \begin{cases} Y_k(t) & \text{if } d(k, \tilde{k}_t) > a \\ W_{k,t} & \text{if } d(k, \tilde{k}_t) \leqslant a \end{cases}, \quad (12)$$

where now $\tilde{k}_t$ is the lowest fitness particle at time t. As expected, this model also exhibits self-organized criticality, since no major modifications are applied to the original model. In fact, it is straightforward to see that the joint probability of the vector $(Y_1(t), Y_2(t), \ldots, Y_N(t))$ is the same as that of $(\Psi_{1/2}(X_1(t)), \Psi_{1/2}(X_2(t)), \ldots, \Psi_{1/2}(X_N(t)))$ for all t. The proof relies on two facts: (i) one way to satisfy that $W_{k,t}$, and $Y_k(0)$ $\forall k$ have p.d.f. f is to write $W_{k,t} = F^{-1}(U_{k,t})$ and $Y_k(0) = F^{-1}[X_k(0)]$ $\forall k$, and (ii) the only important thing for the dynamics is the fitness *order* of the particles, which is preserved by applying monotonic function F^{-1}, and then F^{-1} can be applied directly to all particles of the BS model at the time t studied. Therefore, at the stationary regime ($t \to \infty$), it is enough to understand the uniform fitness case [Eq. (1)] to extrapolate to an arbitrary fitness distribution [Eq. (12)]. In the first case, once the system reaches the stationary state, particle fitness X is uniform($p_c, 1$). Hence, at the thermodynamic limit, particles that evolve with nonuniform updates converge to a situation where the fitness is greater than a critical value p_c^{nu} and it satisfies

$$p_c^{\text{nu}} = F^{-1}(p_c) = F^{-1}\left(\frac{1 + \langle S \rangle}{1 + 2a}\right). \quad (13)$$

Moreover, applying F^{-1} one can see that the fitness of the particles converge to a p.d.f. h equal to

$$h(y) = \begin{cases} 0 & \text{if } y < p_c^{\text{nu}} \\ f(y)/(1 - p_c^{\text{nu}}) & \text{if } y \geqslant p_c^{\text{nu}} \end{cases}, \quad (14)$$

just by applying F^{-1} to the uniform case. We emphasize that $\langle S \rangle$ does not depend on $f(w)$, it depends only on a. This result implies that if one chooses a nonuniform update distribution that favors small values of fitness [e.g., $F(1/2) > 1/2$], then the critical value p_c^{nu} will be smaller than p_c. Figure 5 highlights the difference between uniform and nonuniform update distributions.

Next, we discuss some simplified BS models and study the impact of using Eq. (3) and transformations similar to the one presented above for the nonuniform update on those models.

B. Model 1: Random interacting neighbors

The most well-known modification to the BS model is to break the assumption that the nearest neighbors are the
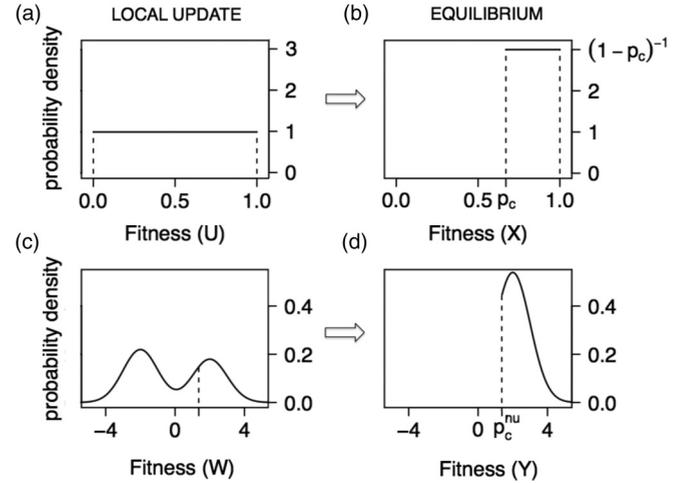
FIG. 5. Bak–Sneppen model. The update distribution is shown on the left and the equilibrium fitness distribution on the right. Two cases are shown: the (a) uniform (0,1) update distribution and (c) a two-modes update distribution example, $f(x)$. The dashed line in (c) corresponds to p_c^{nu}, which verifies $\int_{-\infty}^{p_c^{\text{nu}}} f(x)dx = p_c$.

ones that interact with the lowest fitness particle [8,11,12]. A simplified hypothesis is that the interacting particles are chosen randomly between all possible particles at each time step. This model was introduced in the previous section ("diluted avalanche"); let us call this model the random Bak-Sneppen model (rBS model). It is well known that this model also presents a critical value, p_c^r [11]. In the previous section, we found that

$$p_c^r = \frac{1}{1 + 2a}, \quad (15)$$

just replacing $\langle S \rangle$ by zero in Eq. (3). The value of $\langle S \rangle$ is zero because at the thermodynamic limit, all particles have a fitness value greater than p_c except for the ones that are part of the avalanche, which are not each others' neighbors. Fitness converges to a uniform ($p_c^r, 1$) distribution, and thus the mean fitness is equal to $\frac{1 + p_c^r}{2}$. Figure 6(a) shows the mean fitness as a function of the number of interacting neighbors.

Note that the number of interacting neighbors can be odd in this model; one just needs to replace $2a$ by the number of interacting neighbors, m. This result can be extended to the case of nonuniform fitness. The new critical value $p_c^{r,\text{nu}}$ verifies

$$p_c^{r,\text{nu}} = F^{-1}(p_c^r) = F^{-1}\left(\frac{1}{1 + m}\right). \quad (16)$$

C. Model 2: Random interacting neighbors and binary fitness

A discrete fitness version of this model can be introduced just by considering values of fitness, now called Y, that can only take the values 0 or 1. The dynamics is the following: At each discrete time, a random site with fitness value 0 is updated, as are m other random sites. If there is no site with 0 fitness, then a random site with a fitness of 1 is selected and updated with other m random sites. The sites are always updated with a Bernoulli variable with probability $1/2$. Note that, as there may be ties in fitness, we are obligated to randomly select the lowest fitness particle. The goal is the same as before: We want
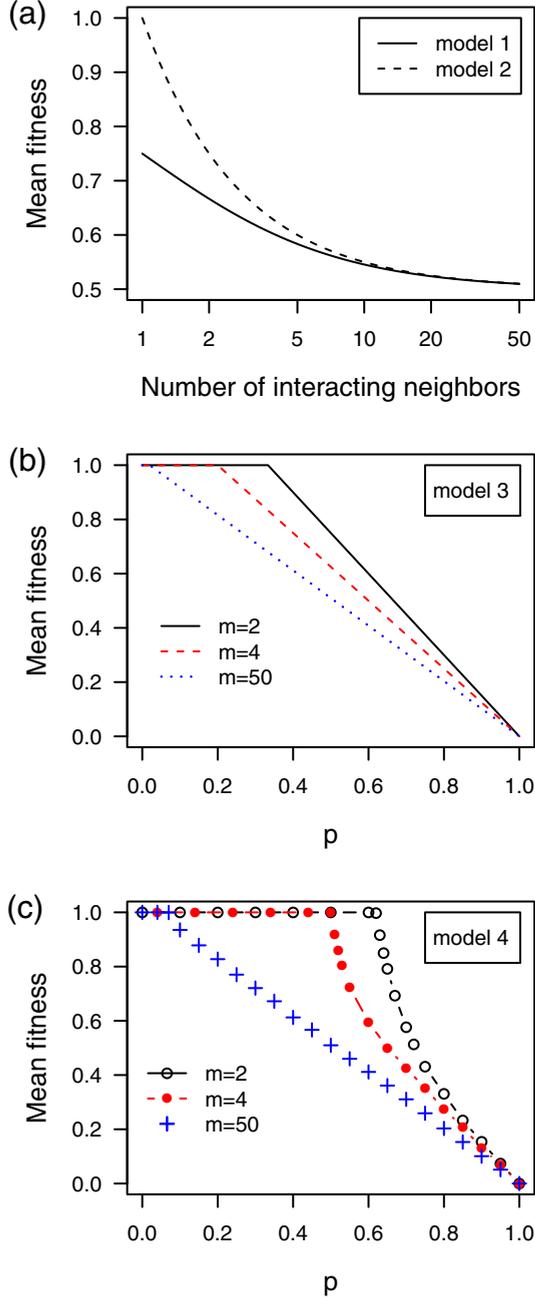
FIG. 6. (a) Mean fitness of models 1 and 2 as a function of the number of interacting neighbors. Mean fitness of (b) model 3 and (c) model 4 as a function of p (probability of fitness equal to zero immediately after the update).

to understand the limit probability law of the particles' fitness. Since fitness in this case is binary, we study the proportion of particles with fitness equal to 1, $P(Y = 1)$, which is equivalent to $\langle Y \rangle$.

This model can be described in terms of the rBS model. Let us take the original rBS model with fitness values X and apply the following function over the fitness:

$$\Psi_{1/2}(X) = \begin{cases} 1 & \text{if } X > 1/2 \\ 0 & \text{if } X \leqslant 1/2 \end{cases}. \quad (17)$$

Again, the joint probability of $(Y_1(t), Y_2(t), \ldots, Y_N(t))$ is the same as that of $(\Psi_{1/2}(X_1(t)), \Psi_{1/2}(X_2(t)), \ldots, \Psi_{1/2}(X_N(t)))$, i.e., the $\Psi_{1/2}$ function converts the rBS model into the discrete version introduced above. At the thermodynamic limit, once the system reaches the steady state, we know that the fitness of the rBS model particles obeys a uniform distribution from p_c^r to 1. Therefore, the proportion of particles that have a discrete fitness equal to 1, $\langle Y \rangle$, verifies

$$\lim_{N \to \infty} \langle Y \rangle = \lim_{N \to \infty} \langle \Psi_{1/2}(X) \rangle = P(X > 1/2) = \frac{1+m}{2m}. \quad (18)$$

Note that for $m = 1$ all particles have a discrete fitness equal to 1. Figure 6(a) shows the behavior of the mean fitness as a function of the number of interacting neighbors [Eq. (18)].

D. Model 3: Random interacting neighbors, binary fitness, and Bernoulli updates

A new model can be introduced if we consider that the updates of the discrete fitness values obey a Bernoulli variable with probability $1 - p$ of having fitness equal to 1. This model, unlike the Bak-Sneppen model, has a parameter. It resembles a percolation model more than the *self-organized* Bak-Sneppen model. Nevertheless, since part of the Bak-Sneppen dynamics is conserved (dynamics governed by the minimum fitness value), we anticipate some unexpected behavior at a particular value of p. Fortunately, as before, this model can be obtained from the rBS model. Just by applying the function Ψ_p [interchange $1/2$ by p in Eq. (17)] to the particle fitness that evolves according to the rBS model, we can obtain the new binary fitness.

At the thermodynamic limit, the fraction of sites with fitness equal to 1 behaves in the following way with the parameter p:

$$\lim_{N \to \infty} \langle Y \rangle = \lim_{N \to \infty} \langle \Psi_p(X) \rangle = \begin{cases} \frac{1-p}{1-p_c^r} & \text{if } p > p_c^r \\ 1 & \text{if } p \leqslant p_c^r \end{cases}. \quad (19)$$

which is equivalent [using Eq. (15)] to

$$\lim_{N \to \infty} \langle Y \rangle = \begin{cases} (1 + \frac{1}{m})(1 - p) & \text{if } p > \frac{1}{1+m} \\ 1 & \text{if } p \leqslant \frac{1}{1+m} \end{cases}. \quad (20)$$

Figure 7(b) shows the average fitness as a function of p for three different values of m.

E. Model 4: Nearest-neighbor interactions, binary fitness, and Bernoulli updates

Finally, we discuss a model similar to Model 3 where the nearest neighbors are the ones that are updated. This model was introduced by Barbay and Kenyon (BK) [15]. Fitness values are binary, but now, once the lowest fitness particle is selected (randomly, since there are ties), the m nearest neighbors are updated with independent Bernoulli variables with parameter $1 - p$ [16]. In Ref. [15] the authors show that for $m = 2$ a critical value p_c^{BK} exists. Moreover, they prove [15] that $0.4563 < p_c^{BK}$ and show by simulations that $p_c^{BK} \approx 0.635$.

Here we present a better lower bound as well as an upper bound for the case of an arbitrary number of neighbors (m). The result is the following:

$$p_c^r < p_c^{BK} < p_c. \quad (21)$$
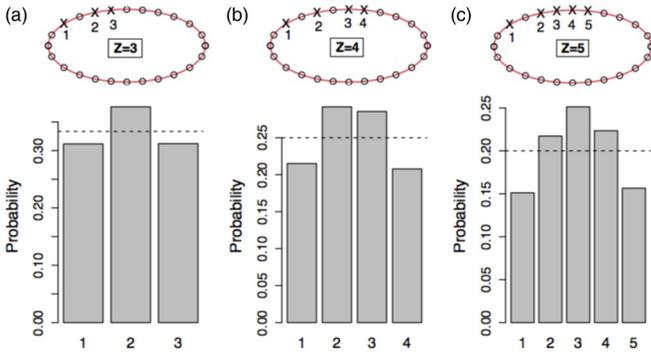
FIG. 7. Simulations from the BS model for $m = 2$. Upper panels: Representation of the BS model where circles represent inactive particles and crosses represent active ones. Lower panels: Proportion of times the lowest fitness particle is the one in the ith position going clockwise when there are (a) three, (b) four, or (c) five active particles.

The critical value of the BK model is bounded by the critical values of the rBS and BS models. For $m = 2$ [Eq. (21)], we state that $1/2 < p_c^{\text{BK}} < 2/3$. The lower bound $p_c^r < p_c^{\text{BK}}$ is easy to understand based on the results presented in the previous section. We have already shown that more "compact" avalanches give rise to larger critical values. Now if we compare the avalanches of Model 3 (critical value p_c^r) with the ones generated with Model 4 (critical value p_c^{BK}), then we see that the latter are more compact, since Model 4 evolves through the updates of *nearest* neighbors. To understand the upper bound, $p_c^{\text{BK}} < p_c$, the argument is more complex but is based on the same idea as avalanche compaction. The avalanches of the discrete version of the BS model are more compact than the ones of the BK model. We explain this argument next.

We first describe the BK model mathematically. Let $N_{0,t}$ be the set that contains all particles with fitness equal to zero at time t, and $N_{1,t}$ the set that contains the rest of the particles (with fitness equal to 1). Let us randomly choose one particle in each set (with equal probability), which we call $h_{0,t}$, and $h_{1,t}$ to the randomly selected particle from set $N_{0,t}$ and $N_{1,t}$, respectively. The lowest fitness particle at time t, $\tilde{k}_t$, is defined as

$$\tilde{k}_t = \begin{cases} h_{0,t} & \text{if } N_{0,t} \neq \varnothing \\ h_{1,t} & \text{if } N_{0,t} = \varnothing \end{cases}. \quad (22)$$

Particles obey the following dynamics:

$$Y_k(t+1) = \begin{cases} Y_k(t) & \text{if } d(k, \tilde{k}_t) > a \\ W_{k,t} & \text{if } d(k, \tilde{k}_t) \leqslant a \end{cases}, \quad (23)$$

where $W_{k,t}$ and $Y_k(0)$ are independent and identically distributed Bernoulli random variables with parameter $1 - p$. For $p = 1/2$, although the model seems to be the binary fitness version of the Bak-Sneppen model, it is not. In order to create its discrete version, there must exist a function $\Psi : \mathbb{R} \to \{0, 1\}$ (or $\Psi : \mathbb{R}^N \to \{0, 1\}^N$) which, when applied individually to each fitness particle (or to its vector), converts the fitness given by Eq. (1) into the discrete version given by Eq. (23). Specifically, if the BK model with $p = 1/2$ were the discrete version of the BS model, then the joint probability of $(Y_1(t), Y_2(t), \ldots, Y_N(t))$ must be equal to the joint probability of $(\Psi(X_1(t)), \Psi(X_2(t)), \ldots, \Psi(X_N(t)))$ [or

$\Psi(X_1(t), X_2(t), \ldots, X_N(t))$], where X_k is the fitness of particle k given by the BS model. To understand why there is no Ψ function that verifies the previous conditions, we focus on the selection mechanism of the lowest fitness particle. In the rBS (Model 1), in its discrete version (Model 2), and also in Model 3, each particle below p_c^r has the same probability of being the lowest one (the same happens above p_c^r). The BS model behaves in a different way. Each particle below p_c does not have the same chance of being the lowest one. At equilibrium, knowing that there are Z_t particles below p_c at time t, the ones at the edge of the avalanche have a lower chance than $\frac{1}{Z_t}$ of being the lowest one, which is one of the reasons why avalanches in the BS model diffuse so slowly. Figure 7 shows empirical evidence of this nonequiprobable law. Once the system is in steady state, we take snapshots and study the distribution of the fitness of active particles. In these snapshots, there are different numbers of active particles (Z). We only analyze the fitness of the snapshots that verify $Z = \{3, 4, 5\}$. Once we have the fitness values for $Z = 3$, for example, we construct a vector (X_a, X_b, X_c) with those fitness values. In the vector's first position, X_a, we put the fitness of the first clockwise particle; in the second position, we put the fitness of the particle in the middle, and in the third position, we put the last particle's fitness. The same procedure is done for $Z = 4$ and 5. Figure 7 shows the fraction of times each particle was the one with the lowest fitness when $Z = 3$ (a), $Z = 4$ (b), and $Z = 5$ (c). As one can observe, the distribution is symmetric as it must be, but the particles in the middle have a greater probability of being the lowest fitness particle. Moreover, the distance between the active particles shapes this distribution (data not shown), i.e., the probability of being the lowest particle depends on the relative positions of the active particles (i.e., the order and distance between them). In the BK model, all active particles have the same chance of being the lowest fitness particle, and in the BS model, we just showed that the particles in the middle have the greatest probability. This last observation is the key argument for why there is no Ψ function that converts the BS model into the BK model preserving the dynamics. There is no way to convert the nonequiprobable selection mechanism in a equiprobable one [Eq. 22].

Nevertheless, if we apply the Ψ_p function previously defined to the fitness of the BS model, then we obtain the true discrete version of the BS model when $p = 1/2$. This discrete model evolves by Eq. (23) but with $\tilde{k}_t$, obeying a much more complicated probability law than the one defined in Eq. (22). In particular, $h_{0,t}$ is not selected with equal probability from the $N_{0,t}$ set. The specific law for $\tilde{k}_t$ exceeds the scope of this work, but we can say that this true discrete version of the BS model will verify the same type of behavior found in Model 3,

$$\lim_{N \to \infty} \langle Y \rangle = \lim_{N \to \infty} \langle \Psi_p(X) \rangle = \begin{cases} \frac{1-p}{1-p_c} & \text{if } p > p_c \\ 1 & \text{if } p \leqslant p_c \end{cases}. \quad (24)$$

On the other hand, it is difficult for the BK model [Eqs. (22) and (23)] to calculate the exact behavior of $\lim_{N \to \infty} \langle Y \rangle$ as a function of p. We can only say that p_c^{BK} must be smaller than the critical value of the true discrete BS model, which is equal to the continuous case (p_c), since the true discrete BS model

generates more compact avalanches than the BK model, and more compact avalanches give rise to larger critical values.

Finally, we study $\lim_{N\to\infty}\langle Y \rangle$ by simulations. Figure 6(c) shows the average fitness as a function p for cases $m = \{2,4,50\}$. We can say that

$$\lim_{N\to\infty}\langle Y \rangle = \begin{cases} g(p) & \text{if } p > p_c^{\text{BK}} \\ 1 & \text{if } p \leqslant p_c^{\text{BK}}, \end{cases} \qquad (25)$$

where g is a nonlinear decreasing function. We believe that there is an interesting and challenging problem in describing function g in detail.

IV. DISCUSSION

In summary, in this paper we presented a local equilibrium equation [Eq. (2)] that allows one to obtain information about the critical value p_c of the Bak-Sneppen model. This equation relates the critical value p_c with the proportion of local neighbors of the lowest fitness particle, $p_{\text{local}} := \frac{\langle S \rangle}{m}$, that have a fitness lower than p_c. We believe that this local equilibrium equation and similar ideas may be applied to advance in the description of other self-organized criticality models. Also, we showed how transformations of the original BS model can be done without altering the model's complex dynamics. First, we showed how to compute the p_c value for the BS model with nonuniform updates. Although perhaps not surprising, this first result gave us the mathematical tools to analyze four different Bak-Sneppen-like models in detail. The average fitness and critical values were studied in these models (see Fig. 6). We believe that there is still a lot to be learned from the original BS model and from BS-like models. How to calculate the value p_c for different topologies, and the function g for the BK model are some of the challenging problems.

[1] P. Bak and K. Sneppen, Punctuated Equilibrium and Criticality in a Simple Model of Evolution, Phys. Rev. Lett. **71**, 4083 (1993).

[2] J. de Boer, B. Derrida, H. Flyvbjerg, A. Jackson, and T. Wettig, Simple Model of Self-Organized Biological Evolution, Phys. Rev. Lett. **73**, 906 (1994).

[3] K. Sneppen, P. Bak, H. Flyvbjerg, and M. Jensen, Evolution as a self-organized critical phenomenon, Proc. Natl. Acad. Sci. USA **92**, 5209 (1995).

[4] P. Grassberger, The Bak-Sneppen model for punctuated evolution, Phys. Lett. A **200**, 277 (1995).

[5] M. Paczuski, S. Maslov, and P. Bak, Field theory for a model of self-organized criticality, Europhys. Lett. **27**, 97 (1994).

[6] S. Boettcher and M. Paczuski, Exact Results for Spatiotemporal Correlations in a Self-Organized Critical Model of Punctuated Equilibrium, Phys. Rev. Lett. **76**, 348 (1996).

[7] M. Felici, G. Caldarelli, A. Gabrielli, and L. Pietronero, Perturbative Approach to the Bak-Sneppen Model, Phys. Rev. Lett. **86**, 1896 (2001).

[8] H. Flyvbjerg, Simplest Possible Self-Organized Critical System, Phys. Rev. Lett. **76**, 940 (1996).

[9] S. Maslov, Infinite Series of Exact Equations in the Bak-Sneppen Model of Biological Evolution, Phys. Rev. Lett. **77**, 1182 (1996).

[10] A. Gillett, R. Meester, and P. van der Wal, Maximal avalanches in the Bak-Sneppen model, J. Appl. Prob. **43**, 840 (2006).

[11] H. Flyvbjerg, K. Sneppen, and P. Bak, Mean Field Theory for a Simple Model of Evolution, Phys. Rev. Lett. **71**, 4087 (1993).

[12] R. Meester and A. Sarkar, Rigorous self-organised criticality in the modified Bak-Sneppen model, J. Stat. Phys. **149**, 964 (2012).

[13] If the number of particles is infinity, then a particle dies only when it has offspring. There is no chance of being "killed" by the offspring of another particle that "fell" at the same site.

[14] In this case, we verify that $p^* = q\frac{1+\langle S \rangle}{1+m}$ where $q = \frac{\langle \tau_1 \rangle}{\langle \tau_1 \rangle + \langle \tau_0 \rangle}$ with $\langle \tau_1 \rangle = \frac{1}{(1-p^*)^3}$ and $\langle \tau_0 \rangle = \frac{1}{1-(1-p^*)^3}$. Using Eq. (2), we obtain that p^* can be any value between 0 and 1, which makes sense considering there is no real critical value

[15] J. Barbay and C. Kenyon, On the discrete Bak-Sneppen model of self-organized criticality, in *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001), pp. 928–933.

[16] In Ref. [15] the authors updated with Bernoulli variables with parameter p and studied the case $m = 2$.