On the discrete Bak-Sneppen model of self-organized criticality

Jérémy Barbay[1] Claire Kenyon [1]

October 19, 2000

Abstract

We propose a discrete variant of the Bak-Sneppen model for self-organized criticality. In this process, a configuration is an n-bit word, and at each step one chooses a random bit of minimum value (usually a zero) and replaces it and its two neighbors by independent Bernoulli variables with parameter p. We prove bounds on the average number of ones in the stationary distribution and present experimental results.

1 Introduction

1.1 Background How does one model rare catastrophic events such as avalanches, volcanic eruptions, and extinctions of species? Self-organizing criticality is a name common to such models. It refers to the tendency of slowly-driven dissipative systems with many degrees of freedom to evolve intermittently in terms of bursts spanning all scales up to system size. These systems traverse "rugged landscapes" in the space of configurations in search of their optimal configuration, with extremely slow relaxation dynamics.

One of the paradigms of self-organizing criticality is the Bak-Sneppen model for coevolutionary avalanches of different species in an ecology [1]. Different species in the same eco-system are related through, for instance, food chains, and co-evolve: the mutation or extinction of one species affects the species which are related to it. The evolution is guided by Darwinian principles, with mutation or extinction of the least fit species. This translates into the following clean mathematical model. The current state of the Bak-Sneppen model is completely defined by n^d "fitness" numbers f_i arranged on a d-dimensional lattice of size n (here each vertex i represents a species, the lattice is the eco-system, and the edges represent relations between species). The rules of its dynamics are very simple: At every time step the smallest of the numbers f_i and its $2d$ nearest neighbors are replaced with new uncorrelated random numbers drawn from some fixed distribution, usually the uniform distribution in $[0, 1]$ (this represents a mutation of the least fit species and coevolution of related species).

This dynamics results in a remarkably rich and interesting behavior, even in one dimension; the parameters describing the system (s.a. aging) typically are not Gaussian but follow power laws. The system organizes itself into a highly correlated state where:

- most species have reached a fitness above a certain threshold,
- but chain reactions, called avalanches, produce large, non-equilibrium fluctuations in the configuration of fitness values.

Although the Bak-Sneppen model is extremely simple, it has not yet been solved in spite of numerous analytical and numerical investigations [16, 8, 2, 11, 15, 12, 4, 13, 14]. In one dimension a real space renormalization group approach was taken [11, 15], but mathematically rigorous results are still lacking.

Inspired by such self-organized models, Boettcher and Percus [5] proposed a new general-purpose optimization heuristic, "extremal optimization", a competitor of paradigms such as genetic algorithms [9] and simulated annealing [10]. Extremal optimization heuristics transform the current suboptimal solution by focusing on its extremely undesirable components and randomly perturbating them (rather than "breeding" better components, which is the strategy of genetic algorithms). Such heuristics are inherently difficult to analyze but experimentally seem promising (see [3] for an application to graph bisection).

1.2 Our model and results Motivated by the difficulty of analyzing rigorously even the one-dimensional version of the Bak-Sneppen model, in this paper we analyze a still simpler model proposed by R. Kenyon[1] with discrete fitness values. In this model, each species has fitness 0 or 1, and each new fitness is drawn from the Bernoulli distribution with parameter p. Since there are typically several least fit species, the process then repeatedly chooses a species for mutation uniformly at random among the least fit species.

This paper only studies the one-dimensional case, whose state is simply an n-bit word $x_0 x_1 \dots x_{n-1} \in$

[1] LRI, CNRS URM 8623, Université Paris-Sud, 91405 ORSAY Cedex, FRANCE. Part of this work was done while the authors were visiting the University of Washington at Seattle.

[1] Personal commünication.

$\{0,1\}^n$. Assume that the process is initially in an arbitrary state: At each step a random $x_i = 0$ is chosen if there is one (or a random $x_i = 1$ otherwise), and x_i and its neighbors x_{i-1} and x_{i+1} are erased and replaced by new independent Bernoulli variables. We assume toroidal boundaries, i.e. x_0's left neighbor is x_{n-1} and x_{n-1}'s right neighbor is x_0.

We focus on the number of ones in the stationary distribution as a function of p. The result of simulations is plotted on figure 1. Here $n = 1000$, and for each value of p from $1/100$ to $100/100$, we take the result of 100 runs of the process (computing the expectation and standard deviation of $(\#1)'s/n$), where each run performs 10^6 steps, starting from the state in which every variable equals 1. We plot the average proportion of ones with an error-bar of radius one standard deviation. Looking at this data, one can make the following observations:

- There is a critical value $p_c < 1$ beyond which the zeroes do not survive, *i.e.* $p_c = \inf\{p | \lim_{n\to\infty} E(\#1)'s/n = 1\}$. If $p > p_c$, then in the stationary distribution there are $n - o(n)$ ones. From our simulation, it appears that $.3 < p_c < .4$. In Theorem 2.1, we prove that $p_c \leq .5436890125....$ This Theorem and its proof are presented in Section 2.

- If the process picked a random x_i, the expected fraction of ones would asymptotically equal p. Picking a least fit species has a drastic effect on the stationary distribution, even when p is quite small. In fact, it appears that for every $p \in (0,1)$, the average proportion of ones is always significantly greater than p. In fact, in Theorem 3.1, we prove that in the stationary distribution there are at least $n/(1+2/3p) - o(n)$ ones on average. This lower bound is experimentally tight for p small, as can be seen on figure 2, which is just a close-up of figure 1 and also plots the function $1/(1+2/3p)$: we see that for $p < .05$, the experimental data almost agrees with the lower bound. The proof of the lower bound is presented in Section 3.

- Some complex, high-variance phenomenon occurs in the middle range of p, between .35 and .4. We performed additional experiments in tht range: figure 3 shows snapshots of the configuration at several points in time, for a single run of the process. In this simulation, $p = .365$ and $n = 1000$. We plot the proportion of ones as a function of time, starting from the state where every variable equals 1. We see that the proportion of ones varies widely, between 100% and 70% or so.

  We take snapshots of the configuration the first time the proportion of ones reaches 90%, 80%, and 70%. To visualize the configuration at those instants, we plot for each $0 < i < n$ the average value of the fitness of x_i and of its 100 closest neighbors $(x_{i-50} + \ldots + x_{i-1} + x_i + x_{i+1} + \ldots + x_{i+50})/101$.

Many open questions remain on this model. One first goal would be proving that there is a phase transition, i.e. that for small p the species do not all achieve fitness 1: we are currently working on this. Another interesting question would be relating the model rigorously to the classical Bak-Sneppen model: for example, is the critical value of p in our model related to the threshold fitness of the Bak-Sneppen model ?

2 Behavior when p is large

The main result of this section is the following theorem, which says that when p is large enough, the ones invade everything and there are no avalanches.

THEOREM 2.1. *Assume* $p > \rho = K/3 - 1/3 - 2/3K$ *with* $K = (17+3\sqrt{33})^{1/3}$ *(so* $\rho \sim 0.5436890125...$*). Then as n goes to infinity, the expected number of zeroes in the stationary distribution is bounded above by a constant.*

Proof. The proof uses a potential function argument. Define the *diameter* D_t of the process at time t as the minimum size of an "interval" containing all the zeroes of the current distribution: in other words, $D_t = n - \max\{j \mid \exists i, x_i = x_{i+1} = \ldots = x_{i+j-1} = 1\}$, where again all indices are modulo n. Note that D_t, if it is greater than 1, can increase by at most 1 at each step. Assume that the process chooses the rightmost zero of D_t (the leftmost zero case is symmetrical).

- If $2 \leq D_t \leq n-1$ then:

$$D_{t+1} - D_t \leq \begin{cases} 1 & \text{with pr.} & (1-p) \\ 0 & \text{with pr.} & p(1-p) \\ -1 & \text{with pr.} & p^2(1-p) \\ -2 & \text{with pr.} & p^3 \end{cases}$$

  where the transitions considered are (in order) $*01 \to **0, *01 \to *01, *01 \to 011$, and $*01 \to 111$.

- If $D_t = 1$ or 0 then $D_{t+1} \leq 3$.

- If $D_t = n$ then D_{t+1} still equals n with probability $(1-p)^3$, and equals at most $n-1$ with the complementary probability $1-(1-p)^3$.

If the process chooses a zero which is not rightmost or leftmost of D_t, then $D_{t+1} \leq D_t$. Let Z_t denote the number of zeroes at time t. Then the process chooses the rightmost or leftmost zero of D_t with probability $2/Z_t$. Since $Z_t \leq D_t$, we always have

$$\frac{2}{D_t} \leq \frac{2}{Z_t} \leq 1.$$

Aside from these constraints, we have no control over Z_t, so in order to prove an upper bound for the diameter we will take the worst case on Z_t and analyze instead the following Markov chain X_t with state space $\{0, 1, \ldots, n\}$. Assume $X_t = i$ and let α_i be defined later.

- If $2 \leq i \leq n-1$ then

$$X_{t+1} - X_t = \begin{cases} 1 & \text{with pr.} \quad (1-p)\alpha_i \\ 0 & \text{with pr.} \quad p(1-p)\alpha_i + (1-\alpha_i) \\ -1 & \text{with pr.} \quad p^2(1-p)\alpha_i \\ -2 & \text{with pr.} \quad p^3\alpha_i \end{cases}$$

- If $i = 0$ or 1 then $X_{t+1} = 3$ with probability α_i and 0 with probability $1 - \alpha_i$.

- If $i = n$ then

$$X_{t+1} = \begin{cases} n & \text{with pr.} \quad (1-p)^3\alpha_i + (1-\alpha_i) \\ n-1 & \text{with pr.} \quad (1-(1-p)^3)\alpha_i \end{cases}$$

Observe that $\lim_t E(X_t) \geq \lim_t E(D_t)$ as long as the holding probabilities $(\alpha_i)_{0 \leq i \leq n}$ are defined to maximize $\lim_t E(X_t)$ subject to the constraint $2/i \leq \alpha_i \leq 1$.

LEMMA 2.1. *Let $\mathcal{M}$ be a Markov chain on $\{0, 1, \ldots n\}$ and let $\mathcal{M}'$ be the Markov chain on the same state space whose transitions from i are: with probability α_i take one step according to $\mathcal{M}$, and with probability $1 - \alpha_i$ stay in state i. Let (π_i) denote the stationary distribution of $\mathcal{M}$. Then the stationary distribution (π_i') of $\mathcal{M}'$ is:*

$$\pi_i' = \frac{\pi_i/\alpha_i}{\sum_j \pi_j/\alpha_j}.$$

The proof is straightforward and ommitted.

Let M denote the maximum value of $\lim_t E(X_t)$. From the above Lemma, clearly M is reached for (α_i) satisfying: $\alpha_i = 2/i$ for $i > \lceil M \rceil$ and $\alpha_i = 1$ for $i \leq \lceil M \rceil$. It only remains to analyze the Markov chain thus defined.

Since this would entail doing some boring calculations, instead we study the Markov chain Y_t which is the same as X_t, except that all holding probabilities α_i are equal to 1 in Y_t. For that very simple chain, a straightforward computation yields $E(Y_{t+1} - Y_t | 2 \leq Y_t \leq n-1) = 1 - p - p^2 - p^3$ and $E(Y_{t+1} - Y_t | Y_t = n) = -1 * (1 - (1-p)^3)$. Both of these are negative as long as $p < \rho = K/3 - 1/3 - 2/3K$ with $K = (17 + 3\sqrt{33})^{1/3}$. We now apply the following Lemma, adapted from [7]:

LEMMA 2.2. *Let $\mathcal{M}$ be an irreducible, aperiodic Markov chain with state space N, and $p_{ss'}$ be the transition probability from s to s' in $\mathcal{M}$, so that*

$$\exists C, \ \forall s, s' \ |s - s'| > C \Rightarrow p_{ss'} = 0$$

and

$$\exists S, \varepsilon \text{ s.t. } \sum_{s' \in N} p_{ss'} * (s' - s) < -\varepsilon \ \ \forall s > S.$$

Then $\mathcal{M}$ is ergodic with stationary distribution π satisfying $\pi(s) < C'e^{-\delta s}$ for all $s \in N$, where C' and δ are positive constants.

Thus the stationary distribution of Y_t is exponentially decreasing: $\Pr(Y = i) \leq Ac^i$ where c is some constant less than 1. From Lemma 2.1, the stationary probability of X must thus have $\Pr(X = i) \leq Aic^i$ and the expectation of X is at most $\sum_i Ai^2c^i$, a bounded quantity. Hence $E(D_t) = O(1)$, and there are on average $n - O(1)$ ones in the stationary distribution. This ends the proof of the Theorem.

3 A lower bound to the number of ones

Even when p is small, the fact that the process always chooses an $x_i = 0$ as the center of its window is already sufficient to bias the number of ones and make it significantly larger than pn (see figure 2).

THEOREM 3.1. *As n goes to infinity, the expected number of ones in the stationary distribution is greater than or equal to:*

$$\frac{3pn/2}{1 + 3p/2} - O(\sqrt{n}).$$

Proof. Let O_t be the number of ones and Z_t the number of zeroes at time t. Let $O_t^{(r)}$ be the number of ones whose immediate right neighbor is a zero, and $O_t^{(l)}$ be the number of ones whose immediate left neighbor is a zero. Then:

$$\begin{aligned} E(O_{t+1} - O_t | \text{history up to } t) &= -\frac{O_t^{(l)}}{Z_t} - \frac{O_t^{(r)}}{Z_t} + 3p \\ &\geq -\frac{2O_t}{n - O_t} + 3p \end{aligned}$$

, where the inequality follows from $O_t^{(l)} \leq O_t$, $O_t^{(r)} \leq O_t$, and $Z_t = n - O_t$. Thus, if we let $A_\epsilon = (3p - \epsilon)n/(2 + 3p - \epsilon)$ and $X_t = A_\epsilon - O_t$, we have that at each step X_t changes by at most 3, and

$$E(X_{t+1} - X_t | \text{history up to time } t, X_t \geq 0) \leq -\epsilon < 0.$$

Now, we use the following version of the Optional Stopping theorem adapted from [6]:

THEOREM 3.2. *If $T_0 \leq T$ are stopping times and $Y_{\min(N,n)}$ is a supermartingale whose absolute value converges in expectation to some finite value, then $EY_{T_0} \geq EY_T$.*

Let t_0 be a time when $X_{t_0} > 0$. Define the random stopping time T as the first time when $X_T \leq 0$ after t_0. The random variable $Z_t = X_t + \epsilon(t - t_0)$ is a supermartingale [17] up to time T, i.e. $E(Z_{t+1} | \text{history up to } t) \leq Z_t$. By the Optional Stopping theorem, $E(Z_T) \leq Z_{t_0}$, or in other words, $\epsilon E(T - t_0) \leq X_{t_0}$. Taking $t_0 = 0$ yields that X_t first becomes negative after $E(T) \leq A_\epsilon/\epsilon$ steps on average.

To examine subsequent excursions of X_t in the positive domain, let t_0' be any time at which $X_{t_0'} > 0$ and $X_{t_0'-1} \leq 0$. We then have $\epsilon(T - t_0') \leq X_{t_0'} \leq 3$, so that $T - t_0' \leq 3/\epsilon$ and the excursions of X_t in the positive domain have average duration at most $3/\epsilon$. Since X_t varies by at most 3 at each step, for every $i > 1$ we have

$$\Pr\{X_t \geq i\} \leq \Pr\{X_t \text{ in excursion of duration } \geq \lceil 2i/3 \rceil\}.$$

Then

$$\begin{aligned} E(X_t) &\leq \sum_{i \geq 1} \Pr\{X_t \geq i\} \\ &\leq 2 \sum_{i \geq 1} \Pr\{X_t \text{ in excursion of duration } \geq i\} \\ &= 2E(\text{length of excursion}) \\ &\leq 6/\epsilon. \end{aligned}$$

Translating this into our original random variable, we obtain:

$$E(O_t) \geq \frac{(3p - \epsilon)n}{(2 + 3p - \epsilon)} - \frac{6}{\epsilon}.$$

Letting $\epsilon = 1/\sqrt{n}$ yields the Theorem.

We leave open the following tantalizing conjecture.

CONJECTURE 1. *There exist $p_0 > 0$ and a constant $c < 1$ such that for $p < p_0$, in the stationary distribution the expected number of ones is at most cn.*

Acknowledgments: We would like to thank Richard Kenyon for suggesting the model and Anna Karlin and Dana Randall for fruitful suggestions.

References

[1] P. Bak and K. Sneppen, Phys. Rev. Lett. 71, 4083 (1993).

[2] J. de Boer, B. Derrida, H. Flyvbjerg, A.D. Jackson, and T. Wettig, Phys. Rev. Lett. 73, 906 (1994)

[3] S. Boettcher, *Extremal Optimization: Heuristics via Co-Evolutionary Avalanches*, Computing in Science and Engineering, to appear.

[4] S. Boettcher and M. Paczuski, Phys. Rev. Lett. 76, 348 (1996)

[5] S. Boettcher and A. Percus, "*Extremal optimization: Methods derived from co-evolution*", in GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference, Morgan Kaufmann, San Francisco, 1999, pp. 825-832

[6] R. Durrett. *Probability: Theory and examples*. Wadsworth and Brooks/Cole, 1991, Chapter 4, Section 7.

[7] G. Fayolle, V.A. Malyshev and M.V. Menshikov. *Constructive Theory of Countable Markov Chains, Part I.* Manuscript, May 1992.

[8] H. Flyvbjerg, K. Sneppen, and P. Bak, Phys. Rev Lett. 45, 4087 (1993).

[9] J.H. Holland. *Adaptation in Natural and Artificial Systems*. Ann Arbor, University of Michigan Press, 1975.

[10] Kirkpatrick S., Gelatt C. D., and Vecchi M. P. Optimization by simulated annealing. *Science*, 220(4598): 671–680, May 1983.

[11] M. Marsili, Europhys. Lett. 28, 385 (1994).

[12] S. Maslov, Phys. Rev. Lett. 74, 562 (1995).

[13] S. Maslov, Phys. Rev. Lett. 77, 1182 (1996).

[14] M. Marsili, P. De Los Rios, and S. Maslov, Phys. Rev. Lett. 80, 1457 (1998).

[15] B. Mikeska, Phys. Rev. E 55, 3708 (1997).

[16] M. Paczuski, S. Maslov, and P. Bak, Phys. Rev. E 53, 414 (1996).

[17] R. Motwani and P. Raghavan, *Randomized Algorithms*: 83–91, Cambridge University Press, 1995.

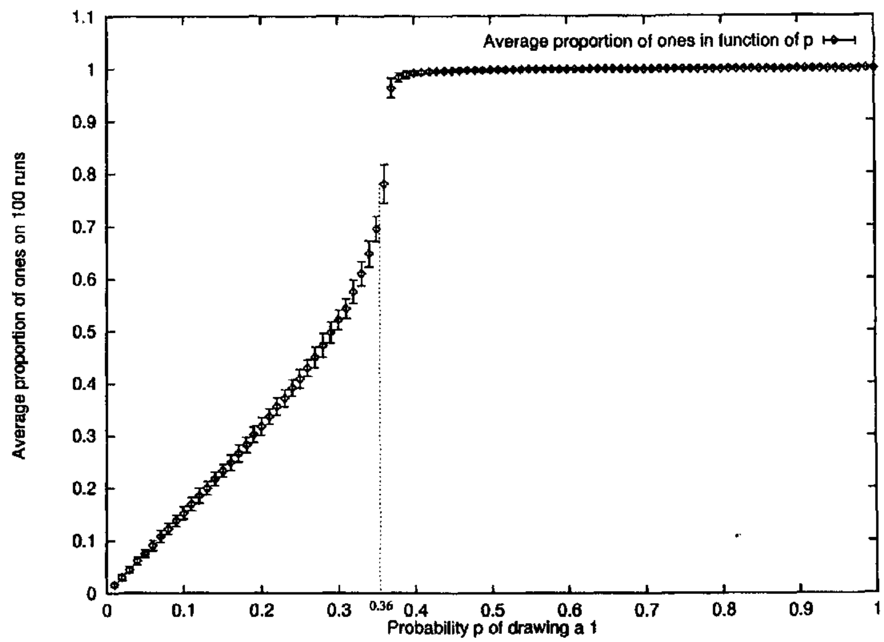

Figure 1: Experimental study of the asymptotic proportion of ones of the discrete Bak-Sneppen process as a function of p. Observe the large variance in the middle range of p. Here $n = 1000$, $T = 10^6$, we take the average of 100 runs, and the error bar has radius equal to one standard deviation.

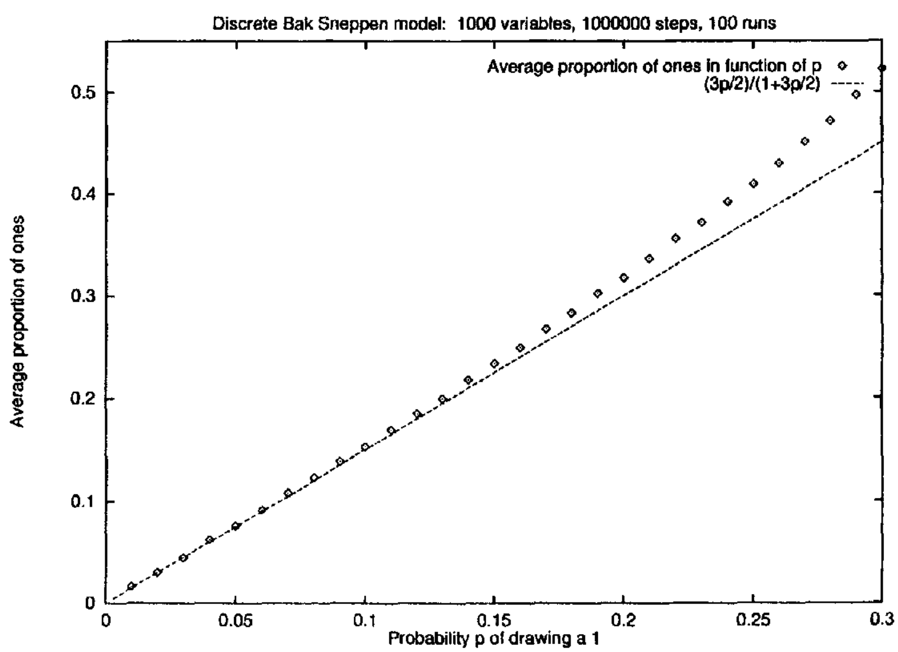

Figure 2: Experimental study of the asymptotic number of proportion of the discrete Bak-Sneppen process as a function of p in the range where p is small. Observe that the lower bound is tight for $p < .05$.

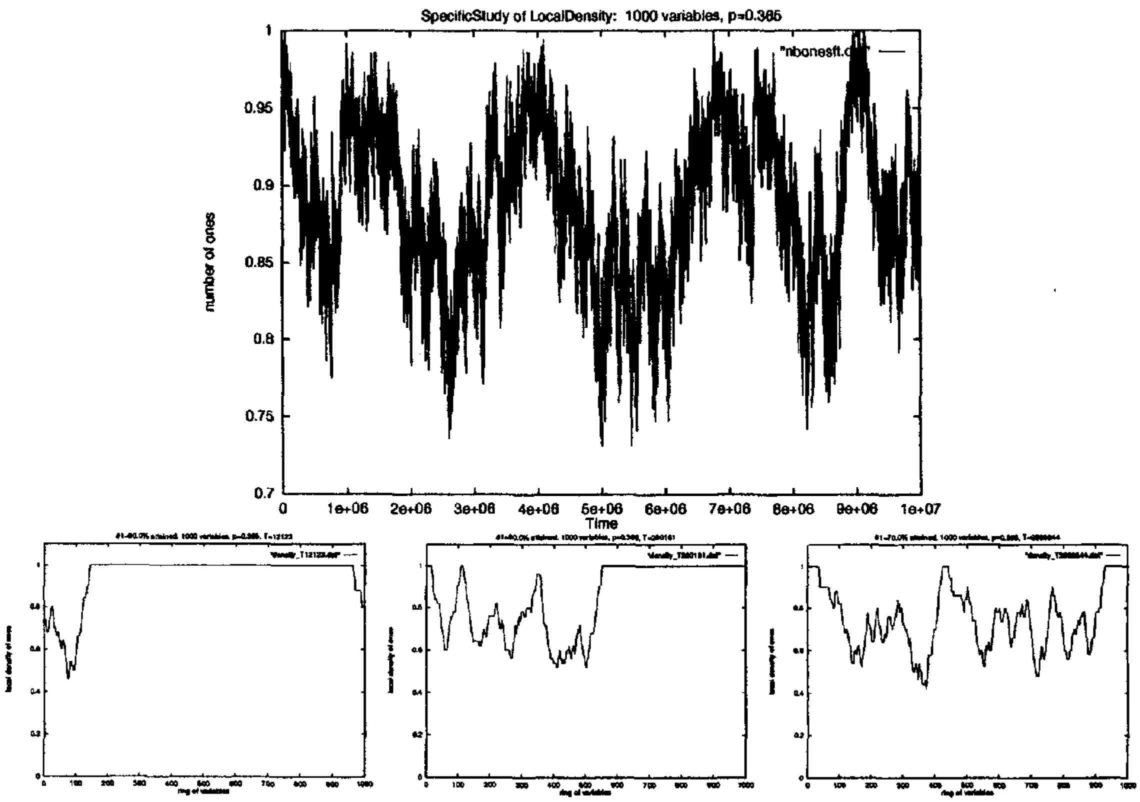

Figure 3: Study of a single run of the process for $n = 1000$ and $p = .365$. On the first line the evolution of the proportion of ones in function of time. Observe that the system returns periodically to the all-one state. On the second line, the snapshots of the first time the system attains a proportion of ones of 90%,80% and 70%. On these snapshots the x-coordinate ranges from 0 to n, and for each i we plot the average fitness of x_i and of its neighbors at distance less than 50 in the configuration.