Closing probabilities in the Kauffman model: an annealed computation.

U. Bastolla and G. Parisi

February 20, 1996

Department of Physics, University "La Sapienza", P.le Aldo Moro 2, I-00185 Roma, Italy

Keywords: Disordered Systems, Genetic Regulatory Networks,
Random Boolean Networks, Cellular Automata

Abstract

We use a probabilistic approach to compute the distributions of periods, transients and weights of attraction basins in Kauffman networks. The results for these quantities are obtained in the framework of the annealed approximation, first introduced by Derrida and Pomeau. Numerical results for the average periods are in good agreement with the computed values of the exponents. We report also on some interesting features which cannot be explained within the annealed approximation.

1 Introduction

Kauffman networks are disordered dynamical systems. They were introduced in 1969 as a simplified model of the genetic regulatory system acting in cell differentiation [1]. The debate about cellular automata caused physicists to be interested in the model [2]. It was soon recognized [3] that these networks have deep analogies with the disordered systems studied in statistical mechanics, *e.g.* the infinite range spin glass [4].

Kauffman networks are defined as follows. A configuration of the system consists of N binary variables, $\sigma_i = 0, 1,\ i = 1 \cdots N$. In the biological interpretation, the variable σ_i represents the state of activation of the i-th gene in the cell.

The configuration space is provided of a metric structure through the Hamming distance (normalized to one). The Hamming distance of two configurations C and $C'$is defined as the fraction of variables which are different between two configurations:

$$d\,(C, C') = \frac{1}{N} \sum_{i=1}^{N} \mid \sigma_i - \sigma_i' \mid . \tag{1}$$

Although the system is disordered, the dynamics are deterministic, with time independent rules chosen at random in the following way.

For each gene i, K genes are extracted randomly; these genes send a signal to the i^{th} gene [1]. In the following these genes will be denoted by $j_l(i)$, $l = 1 \cdots K$.

For each gene i and each of the 2^K configurations of the controlling genes one defines a functions $f_i(\sigma_1 \cdots \sigma_K)$ whose values are random extracted: they are 0 with probability p and 1 with probability $1 - p$. These dynamic rules (i.e. the list of the controlling genes and the functions f_i) are *quenched*: they do not change during the evolution of the system.

At each time step all the N variables are updated simultaneously, according to the signal received (*i.e.* the values of the K controlling genes). The boolean function $f_i(\sigma_1, \cdots \sigma_K) \in \{0, 1\}$ codes the response to the signal. The dynamic rules are thus the following:

$$\sigma_i(t+1) = f_i\left(\sigma_{j_1(i)}, \cdots \sigma_{j_K(i)}\right) . \tag{2}$$

In a more compact way, the evolution law can be put in the form $C(t+1) = \Phi_\eta\,(C(t))$, where C denotes a configuration of the variables σ at time t and η a choice of the dynamic rules.

Due to the fact that configuration space is finite, every trajectory ends up, after a transient time, on a periodic orbit. Configuration space can be thus partitioned into a certain number of attraction basins of weight W_α, where α labels the periodic orbit, W_α being proportional to the number of different initial conditions which end in the periodic orbit α.

[1] It is normally assumed that the same gene can be extracted more than once in the list of the K genes which control the i^{th}. A different choice, e.g. all the K genes different is likely to produce the same effects in the infinite volume limit.

We define three kinds of average, as is customary in the theory of disordered systems:

- The time average for a fixed initial configuration of a given network, which in the infinite time limit is equivalent to the average over the corresponding limit cycle.
- The average over all possible initial configurations of the same network. In the large time limit this average corresponds to a sum over all different cycles weighted with their own W_α.
- The average over different realizations of the network.

It is interesting to study what happens in the limit where $N \to \infty$. It was soon discovered that in this limit Kauffman Networks have two different kinds of dynamical behavior: a *chaotic* regime where the average cycle length grows exponentially with N, and a *frozen* regime, where it goes to a constant limit. It was rather surprising to discover that in the chaotic regime there are cycles whose weight is not going to zero and that the weights W have a well defined probability distribution in the limit where $N \to \infty$ [3, 5].

Analytic results are available for two systems that represent extremal situations and are not at the transition between one regime and another: the point with $K = 1$, which was studied analytically by Flyvbjerg and Kjaer [6], and the point with $K = \infty$, which is equivalent to a model previously known as the Random Map [7], as Derrida and Flyvbjerg noted in [5].

At the boundary between the two regimes the average cycle length grows as a power of N. There is one point in this critical region of the model, (i.e. the point with $K = 2$ and $p = 1/2$), which is has been the object of careful numerical investigations because it shows characteristics very reminiscent of the typical properties of real cells [8]. Numerical simulations [1] have shown that in networks with these parameter values the median period length grows as $\sqrt{N}$. This behavior corresponds, in the biological metaphor proposed by Kauffman, to the fact that the median cell cycle time grows as a power law of the number of genes in the cell, independent of the *taxa* considered.

Moreover, the same simulations show that the average number of limit cycles in one of such networks grows as $\sqrt{N}$. Kauffman suggests that each cycle in the activities of the genes corresponds to one of the different cell types that can be realized by the integrated system of the genes. Within this interpretation the results of simulations is in striking agreement with the fact that the average number of cell types in an individual of a given *species* appears to grow just as the square root of the number of genes contained in its genome.

In contrast to the interest devoted to networks with $K = 2$ and $p = 1/2$, other systems which belong to the the critical line which separate the two regimes have not received attention either in numerical simulations or in theoretical works. In this article we suggest, using some information coming from the annealed approximation, that such quantitative properties are common to the whole critical line. Numerical results that will be reported elsewhere seem to confirm this point.

Another characteristic of the critical and frozen networks which has a biological interpretation is much better, though not completely, understood: such networks share remarkable properties of stability with respect to external perturbations, in the sense that changing at random the state of a macroscopic portion of the genes after the system has attained a limit cycle does not alter the cycle where the system is. This property has a strong analogy with the homeostatic stability of true cells.

It is possible to characterize the transition between the two dynamical regimes as a phase transition in statistical mechanics. Two order parameters have been used to describe it.

- Derrida and Pomeau [9] considered the distance between the equal time configurations on two randomly chosen trajectories in the same network. In the framework of the annealed approximation (which we also use in the present work) they showed that the average distance reaches (when the time goes to infinity) a limit value independent on its initial value.

  The limit value of the the average distance is non zero in the chaotic phase and zero elsewhere, thus justifying its identification as order parameter. Moreover in the chaotic phase the effect of a perturbation, however small, grows to a macroscopic size. On the contrary in the frozen phase and at the critical point, the effect of a small perturbation tends to zero when the size N of the system grows. In this last case a perturbation however large becomes microscopic after a transient time.

  The annealed approximation for the evolution of the average distance has been shown to be exact, in the limit of large systems, up to times of order $\log N$[10, 11].

- Some years later Flyvbjerg introduced the concept of *stable core*, defined as the set of variables that evolve to a constant state not depending on the initial configuration.

  The fraction of variables belonging to the stable core may be regarded as an order parameter. In the limit of an infinite system, it tends to a value less than 1 in the chaotic phase while it tends to 1 in the frozen phase and in the critical line. In this last case, when the system reaches a cycle nearly all the elements of the network freeze to a constant state independent on the cycle and they do not take part in the dynamics. This fact explains the remarkable stability of the frozen phase.

Both the approaches outlined above agree on the position of the critical line in parameter space, which turns out to be given by the equation[9, 13, 12]:

$$2p_c(1-p_c) = \frac{1}{K} \tag{3}$$

Despite of the success of this approach, the connection between the order parameters and dynamical properties (*e.g.* the average cycle length or the weights of attraction basins) is still lacking. The present work is intended to help to fill in this gap.

Our approach will be a probabilistic one. We will define a probabilistic scheme that allows us to compute the dynamic properties starting from the distribution of the overlap $q(t,t') = 1 - d(t,t')$. The introduction of closing probabilities (defined in section 2) will play a crucial role.

In the framework of the annealed approximation (introduced by Derrida and Pomeau [9]) the overlap distribution will be obtained in section 3. In the following sections we will derive from it all the relevant dynamic properties (at least in the chaotic phase where the annealed approximation works better), *i.e.* the distribution of cycle lengths (in section 4), the distribution of attraction basins weights [2] (in section 5) and the average number of cycles in a generic network (in section 6).

In section 7 we present some numerical results. This section is subdivided into three subsections: in the first one we show data concerning the distribution of the overlap; they support the validity of the annealed approximation, but only in some range of the temporal variables. In the second subsection we report some numerical simulations of the closing probabilities. These simulations show deviations from the annealed approximation at later times after an initial agreement at short times. Finally in subsection 7.3 for different values of the parameter K we estimate the exponents which control the large N behavior of the average cycle length. We find that there is a good agreement between our computations and the numerical results as far as the exponents are concerned. Unfortunately the distribution of cycle lengths that we measured is quite different from the predicted one.

In the last section we discuss the range of validity of the approximation used and the physical origin of discrepancies that we noticed between the theoretical findings and the simulations. We finally point out the direction of future work.

2 Closing probabilities

Let us start with a few definitions. In the following, we will use the overlap $q(C,C') = 1 - d(C,C')$ to describe the metric of configuration space.

We will follow a probabilistic strategy, with the aim of computing the probability distribution of the overlap $q(t,t')$ between the configurations $C(t)$ and $C(t')$, $t' > t$, on the same trajectory. We find it convenient to only consider trajectories not yet closed at time $t' - 1$. We denote by the symbol $P_N(t,t';q)$ the probability that $q(t,t') = q$, computed on the restricted set of of trajectories which are not yet closed at time $t' - 1$.

We define the closing probability $\pi_N(t,t')$ as the probability to have $q(t,t') = 1$ (this condition implies that that the two configurations at time t and t' are equal), the average being done over the trajectories not yet closed at time $t' - 1$:

[2] In the chaotic phase the distribution of attraction basins weights turns out to be identical to the one computed by Derrida and Flyvbjerg in the limit case $K \to \infty$, i.e.the so called Random Map [5].

$$\pi_N(t,t') = P\left\{q\left(t,t'\right) = 1 \mid q(t_1,t_2) < 1, t_1 < t_2 < t'\right\}. \quad (4)$$

We will have also to deal with the total closing probability, $\tilde{\pi}_N(t)$, which is defined as the probability that a trajectory not yet closed at time $t-1$ closes at time t regardless of the length of the cycle reached:

$$\tilde{\pi}_N(t) = \sum_{l=1,t-1} \pi_N(t-l,t). \quad (5)$$

With these definitions it is easy to evaluate $F_N(t)$, i.e. the probability that a trajectory has not yet closed at time t. It satisfies the equation

$$F_N(t+1) = F_N(t)\left(1 - \tilde{\pi}_N(t+1)\right). \quad (6)$$

In the limit of large systems we can approximate $\log(1-\tilde{\pi}_N(t))$ as $-\tilde{\pi}_N(t)$, using the fact that this quantity is small even at times when nearly every trajectory has closed. Introducing a continuous time variable, we find the simple expression

$$F_N(t) = \exp\left(-\int_1^t \tilde{\pi}_N(\tau)d\tau\right), \quad (7)$$

so that the probability to find a trajectory that, after a transient time t, reaches a cycle of length l, is given by

$$P\left\{T=t, L=l\right\} = \exp\left(-\int_1^{t+l} \tilde{\pi}_N(\tau)d\tau\right) \pi_N(t,t+l). \quad (8)$$

3 Annealed approximation: the overlap distribution

The goal of computing the overlap distribution becomes easier if one can reduce the evolution of the overlap to a stochastic process.

The deterministic dynamics of Kauffman networks depends on two kinds of quenched random disorder: the connections in the network, and the response functions, whose binary value must be extracted for any of the 2^K possible configurations of the input genes. For each gene i the configuration of the corresponding input genes will be called a signal and we will use for it the symbol $S_i(t)$:

$$S_i(t) = \{\sigma_{j_1(i)}(t), \cdots, \sigma_{j_K(i)}(t)\}. \quad (9)$$

We can transform the Kauffman model into a stochastic process by extracting the values of the response functions to the signal not at the beginning, but every time a variable receives a signal that it had never received, and using the values already extracted every time the same signal comes back.

Note that, if every gene receives signals from the whole system (i.e. when $K \to \infty$), we do not have to worry about this last eventuality, which never occurs before the trajectory closes. In this limit the annealed approximation, that we are going to define later, is exact.

The limit of infinite K is equivalent to the Random Map model (see, *e.g.*, [7]) and was studied analytically in great detail by Derrida and Flyvbjerg [5]. We will generalize the results of [5] to finite K values, in the framework of the annealed approximation.

This approximation was introduced by Derrida and Pomeau, in order to obtain the equations for the evolution of the average Hamming distance between two randomly chosen configurations; we will use it to treat the overlap as a Markovian stochastic process, and so to compute the closing probabilities. We will argue later from our simulations that the overlap distribution obtained through the annealed approximation is very close to the quenched one.

In order to compute the overlap between two configurations at times $t+1$ and $t'+1$, for every gene we distinguish two cases: either the gene received the same signal at time t and at time t', and so its state at $t+1$ and at $t'+1$ is the same, or it didn't:

$$q(t+1,t'+1) = \frac{1}{N}\sum_i \delta_{S_i(t)S_i(t')} + \frac{1}{N}\sum_i \left(1-\delta_{S_i(t)S_i(t')}\right)\delta_{f(S_i(t))f(S_i(t'))}. \tag{10}$$

The annealed approximation consists of the following procedure: we extract randomly *at every time step* all the connections in the network, keeping memory only of the value of the overlap $q(t,t')$, so that the probability that the two signals $S_i(t)$ and $S_i(t')$ are the same is $q(t,t')^K$; if they are not equal, we extract at random the values of the response function, and they will be equal with probability $1-\rho$, where $\rho = 2p(1-p)$ [9].

Combining these two steps, we get an expression for $\gamma(x)$, which represents the annealed probability that a gene is in the same state at two different times t and t', under the condition that the overlap between the two preceding configurations (i.e. $q(t-1,t'-1)$) was x:

$$\gamma(x) = (1-\rho) + \rho x^K. \tag{11}$$

When using the annealed approximation, it is not possible to impose the condition that the trajectory where we measure $q(t,t')$ was not closed at time $t'-1$, because we are keeping memory only of the value of $q(t-1,t'-1)$. We can only impose the weaker condition that it did not close on a cycle of length $l = t'-t$. As we shall see later, this fact is likely to be the source of the main discrepancies between the annealed approximation and our simulations.

If we impose the condition that the trajectory does not close on a cycle of length $l = t'-t$ we obtain the master equation

$$P_N\left(t+1,t+l+1;q_n\right) = \binom{N}{n}\frac{\sum_{m=0}^{N-1} P_N\left(t,t+l;q_m\right)\left(\gamma(q_m)\right)^n\left(1-\gamma(q_m)\right)^{N-n}}{1-P_N\left(t,t+l;1\right)}, \tag{12}$$

where $q_m = m/N$ and $\gamma(x)$ is given by (11).

Now we need to find the solutions of such an equation. Towards this end it is convenient to put the overlap distribution into an exponential form:

$$P_N\{q(t,t+l)=q_n\} = C_N(q_n;t,l)\exp\left(-N\alpha_{l,t}(q_n)\right), \tag{13}$$

so that, using the Stirling expansion of the binomial coefficient and transforming the sum into an integral, we get the following form of the master equation:

$$C_N(x;t+1)e^{-N\alpha_{t+1}(x)} = \frac{N}{\sqrt{2\pi N x(1-x)}}\int_0^{1-\frac{1}{N}} dq C_N(q;t)\frac{\exp\left(-Nf(q,x;t)\right)}{1-C_N(1;t)e^{-N\alpha_t(1)}}, \tag{14}$$

where

$$f(q,x;t) = \alpha_t(q) + x\log\left(\frac{x}{\gamma(q)}\right) + (1-x)\log\left(\frac{1-x}{1-\gamma(q)}\right). \tag{15}$$

We have omitted the l dependence of $\alpha_{l,t}(x)$ and $C_N(x;l,t)$. It is consistent to assume that the l dependence can be neglected because the variable l doesn't appear explicitly in the equation.

Asymptotically in N, we can perform the integral in equation (14) with the saddle point method, neglecting everything but the contribution of the stationary point of the function $f(q,x;t)$, which, from now on, will receive the name $q_t(x)$. In this limit the value of the prefactor $C_N(q_n;t,l)$ disappears. Moreover, in the same limit, we can approximate the normalization factor $1-P_N(1)$ by 1.

Thus, taking the logarithm of both terms of equation (14), dividing by N and neglecting terms which are infinitesimal when N goes to infinity, we get the following equation for the unknown function $\alpha_{t+1}(x)$:

$$\alpha_{t+1}(x) = \alpha_t(q_t(x)) + x\log\left(\frac{x}{\gamma(q_t(x))}\right) + (1-x)\log\left(\frac{1-x}{1-\gamma(q_t(x))}\right). \tag{16}$$

The function $q_t(x)$ must be determined taking the derivative of $f(q,x;t)$ respect to q and equating it to zero, with the additional condition $q_t(x)<1$:

$$\alpha_t'(q_t(x)) - \gamma'(q_t(x))\ \left(\frac{x}{\gamma(q_(x))} - \frac{1-x}{1-\gamma(q_t(x))}\right) = 0. \tag{17}$$

The process defined by (12) is ergodic, because we exclude the overlap $q=1$ which is an absorbing point, and, given an initial distribution $P_N(0,l;q)$, it evolves to a distribution stationary in t and independent on l, which appears only in the initial distribution but not in the transition probability.

We look for a stationary distribution $\widetilde{P}_N$ in an exponential form,

$$\widetilde{P}_N(q=x) = C_N(x) e^{-N\alpha(x)}. \tag{18}$$

In the saddle point approximation, the equation for $\alpha(x)$ is just eq. (16) without the time label,

$$\alpha(x) = \alpha(q(x)) + x \log\left(\frac{x}{\gamma(q(x))}\right) + (1-x)\log\left(\frac{1-x}{1-\gamma(q(x))}\right), \tag{19}$$

and the unknown function $q(x)$ can be determined selfconsistently imposing the saddle point condition, with the additional condition $q(x) < 1$:

$$\alpha'(q(x)) - \gamma'(q(x)) \left(\frac{x}{\gamma(q(x))} - \frac{1-x}{1-\gamma(q(x))}\right) = 0. \tag{20}$$

Taking the derivative of (19) and inserting the result into (20) we get a non-local transcendent equation for the function $q(x)$:

$$\gamma'\left(q(x)\right)\left(\frac{x}{\gamma\left(q(x)\right)} - \frac{1-x}{1-\gamma\left(q(x)\right)}\right) = \log\left(\frac{q(x)}{1-q(x)}\right) - \log\left(\frac{\gamma\left(q\left(q(x)\right)\right)}{1-\gamma\left(q\left(q(x)\right)\right)}\right). \tag{21}$$

We were not able to solve this equation analytically, so we had to solve it numerically. This can be done by considering a discrete version of (16).

Before turning to this point, however, there are some important features of the overlap distribution that can be easily understood.

It is not difficult in principle to compute all the derivatives of $\alpha(x)$ and $q(x)$ in the point Q^* where the overlap is concentrated

Taking the derivative of (16), we can see that the point $Q(t)$ where the overlap distribution is concentrated satisfies the equation

$$Q(t+1) = \gamma\left(Q(t)\right), \tag{22}$$

which, at stationarity, becomes

$$Q^* = \gamma(Q^*) = 1 - \rho + \rho {Q^*}^K. \tag{23}$$

This equation was first derived by Derrida and Pomeau as a mean field solution of the annealed model [9]. The condition $\gamma'(1) = K\rho = 1$ separates parameter space in two different regions:

- $K\rho \leq 1$: *frozen phase* and critical line. The only fixed point of the map (22) is $Q^* = 1$. This implies, as we shall see, that $\alpha(1)$ goes to zero and the typical cycle length increases less than exponentially with system size.

- $K\rho > 1$: *chaotic phase*. The fixed point $Q^* = 1$ becomes unstable and the overlap is concentrated around another fixed point of (22), $Q^* < 1$. In this case $\alpha(1)$ is finite and the typical cycle length increases exponentially with system size.

The variance of the distribution can also be easily obtained, in the Gaussian approximation, taking the second derivative of (16) computed in $x = Q(t+1)$ and imposing the saddle point condition.

One finds that the variance multiplied by N, $V(t)$, follows the recursive equation

$$V(t+1) = (\gamma'(Q(t))^2\, V(t) + Q(t+1)(1 - Q(t+1)), \tag{24}$$

whose fixed point is

$$V^* = \frac{Q^*(1-Q^*)}{1 - (\gamma'(Q^*))^2}. \tag{25}$$

From here it is clear that the stationary overlap distribution can not be concentrated near $Q^* = 1$ if $\gamma'(1) = K\rho$ is greater than one, otherwise the variance of the distribution would be negative.

It is apparent that the stationary overlap distribution is a non binomial one: the presence of correlations between the genes are made evident by the fact that V^* is larger than $Q^*(1 - Q^*)$. Moreover, it is also easy to see that $\alpha'(1)$ is approximately equal to ρK and does not diverge, while in the binomial distribution it does.

On the other hand, when K tends to infinity destroying the correlations between consecutive overlaps, the stationary overlap distribution becomes a binomial one and Q^* tends exponentially to $1 - \rho$.

The exponent $\alpha(1)$ is an increasing function of K and tends exponentially to $1 - \rho$ when K grows.

In the frozen phase and on the critical line, where $\gamma'(1) \le 1$, $Q^* = 1$ is the true solution of equation (23), and $\alpha(1)$ vanishes.

3.1 Numerical computation of $\alpha(x)$

Let us now turn to the numerical computation of the exponent $\alpha(x)$; we are particularly interested to its value at $x = 1$. We shall see later that the exponent controlling the closing probability is given by $\alpha(1)$.

In order to discretize the equations for $\alpha(x)$, we will consider the situation where N is finite and we will eventually send N to infinity. In this computation we will consider the equation (16) for the evolution of the function $\alpha_t(x)$, which is now function of the discrete overlap $q_n = n/N$. As usual, we do not specify the dependence on l, which only appears in the initial condition $\alpha_0(\{q_n\})$. Equation (16) then becomes

$$\alpha_{t+1}(q_n) = q_n \log(q_n) + (1-q_n)\log(1-q_n) + \\ \min_{m<N}\left[\alpha_t(q_m) - (q_n)\log(\gamma(q_m)) - (1-q_n)\log(1-\gamma(q_m))\right], \tag{26}$$

where the saddle point condition has been substituted by the prescription to take the minimum value of the function in square brackets.

For *finite* N this map does not have a fixed point $\{\bar{\alpha}(q_n)\}$, as can be seen numerically: for instance, $\alpha_t(1)$ reaches a minimum at a certain time and then increases uniformly.

This increase is not in contradiction with the existence of a stationary distribution; in fact, for finite N, equation (26) misses the normalization factor $(1 - P_N(1;t))^{-1}$, and so the probability computed by means of it has to decrease when the exact probability has reached the stationary value.

The increase of $\alpha_t(1)$ in one time step is of order $1/N$. In fact, when the minimum of $\alpha_t(x)$, $Q_N(t)$, reaches the stationary value Q_N^*, $\alpha(Q_N^*)$ grows by a term equal to the difference between Q_N^* and $\gamma(Q_N^*)$, which comes from the discretization and is of order $1/N$. So it is of the same order of the terms neglected in the saddle point approximation.

There are no problems in obtaining the solution of the equation in the limit $N \to \infty$. One possibility consists in stopping the iterations immediately after Q_N^* has become stationary, for instance when $\alpha_t(1)$ reaches its minimum value as a function of t; that is what we did in our algorithm. In this way we find a solution of the saddle point equations with an accuracy of order N^{-1}.

There is another finite N effect in our numerical computation: because of the condition $m < N$ (this means that the overlap $q_N = 1$ cannot be taken as a starting point), the stationary value of $\alpha_N(q_n)$ decreases with N. In fact, the larger N, the lower the minimum in (26) can be, and this effect, of order $1/N$, is especially important for $n = N$, because in this case in the stationary state the minimum in (26) would be attained for $m = N$.

The finite size effects become more relevant on the critical line $\rho = 1/K$. As a discretization effect, the most likely value of the annealed distribution is not $Q^* = 1$, because the corresponding value $m = N$ in the equation (26) is not allowed. At stationarity $\alpha(q_n)$ has a minimum for $Q_N^* = 1 - \epsilon_N$, where Q_N^* is such that $\gamma(Q_N^*) - Q_N^*$ is of order $1/N$ and hence ϵ_N is of order $1/\sqrt{N}$, as we expect it to be on the grounds of considering the exact equations for the moments of the overlap.

3.2 Beyond the saddle point approximation

In the frozen phase and on the critical line equation (16) is no longer a good approximation, because we are not allowed to neglect $P_N(1)$ in the denominator of equation (12). If we keep equation (12) in the complete form, it is easy to obtain, through the generating function, the equations for the moments of the overlap distribution. To simplify some calculations, however, it is convenient in this case to consider the moments of the distance $d = 1 - q$, which go to zero when N grows. We get the following system of equations:

$$(1 - P_N) \left\langle d \left(d - \frac{1}{N} \right) \cdots \left(d - \frac{l-1}{N} \right) \right\rangle = \tag{27}$$

$$= \left(1 - \frac{1}{N} \right) \cdots \left(1 - \frac{l-1}{N} \right) \left\langle (1 - \gamma(1-d))^l \right\rangle , \tag{28}$$

$$P_N \, (1 - P_N) = \left\langle (1 - \gamma(1-d))^N \right\rangle - P_N, \tag{29}$$

where $P_N \equiv \Pr\{d = 0\} = \Pr\{q = 1\}$ and l is arbitrary.

The first of these equations has the form

$$(1 - P_N) \langle d \rangle = \langle 1 - \gamma(1-d) \rangle . \tag{30}$$

Using the expansion $1 - \gamma(1-d) \approx cd + (K-1)/2d^2$, where $c = K\rho$, only the two moments of lowest order appear in this equation. In the frozen phase ($c < 1$) it is enough to consider the terms of order $\langle d \rangle$, and one gets $P_N = (1-c)$, independent of N, and one discovers that $\langle d \rangle$ is a quantity proportional to $c/(1-c)N^{-1}$.

On the critical line, on the other hand, equation (30) becomes

$$P_N \langle d \rangle = \frac{K-1}{2} \left\langle d^2 \right\rangle , \tag{31}$$

which suggests, coherently with the results of last subsection, that both P_N and $\langle d \rangle$ are of order $N^{-1/2}$ along the critical line.

These equations are not valid in the case $K = 1$, $c = 1$ (critical), where $\gamma(q) \equiv q$ and the average distance is for all times equal to $1/2$, as one can see for instance by using the saddle point approximation (22). After a time of order N the approximation fails, because the variance diverges linearly with time. Beyond the saddle point approximation this case must be treated separately because another absorbing point, $d = 1$, is now present beside $d = 0$ (the distance $d(t, t+l) = 1$ corresponds in this case to a system which reaches a cycle of length $2l$ after a transient time t). We will not get into further details.

4 Period distribution within the Annealed Approximation

4.1 Closing probabilities

Although the closing probability (*i.e.* the probability that $q(t, t')$ is equal to 1 with the condition that the trajectory has not closed before time t') is ill defined in the annealed approximation, we can obtain results on its behavior by assuming that it has the exponential behavior shown in (13) and that the value of the function $\alpha(1)$ is not too different from the annealed approximation. In this way, if we assume that the function $\alpha(x)$ is continuous at one, we obtain that the closure probability is given by:

$$\pi_N(t,t') \propto \exp\left(-N\alpha(1)\right), \tag{32}$$

provided that the smallest time t is large enough for the overlap to reach the stationary distribution.

In the chaotic phase $\alpha(1)$ has a finite limit when N goes to infinity and the decrease of the closing probability is exponential in N. But outside the chaotic phase $\alpha(1)$ goes to zero. To compute the closing probability in the frozen phase and in the critical line it is important to see how the term of $O(1)$ vanishes with time. Let us see what happens at the critical line $\rho = 1/K$.

The asymptotic evolution of the most probable overlap in the critical line can be described by a differential equation, identifying the derivative with the increase in one step:

$$Q'(t) = \frac{1}{K}(Q-1)^2 \sum_{i=1}^{K-1} (K-1-i)Q^i, \tag{33}$$

whence we get, neglecting terms of higher order in $1/t$,

$$1 - Q \propto \left(\frac{(K-1)t}{2}\right)^{-1}. \tag{34}$$

When $Q(t)$ is close to 1 the closing probability is well approximated by a Gaussian:

$$\begin{aligned} \pi_N(t,t') &\propto \frac{1}{\sqrt{N}} \exp\left(\frac{N}{2}\frac{(1-Q(t))^2}{V(t)}\right), \\ &= \frac{1}{\sqrt{N}} \exp\left(\frac{-4N}{(K-1)t^2}\right), \end{aligned} \tag{35}$$

where we have used the fact that the variance $V(t)$ in the critical line is given by $1/2(K-1)$ plus terms of order $1-Q(t)$, and we have inserted the factor $1/\sqrt{N}$ to take into account the fact that the asymptotic closing probability is a quantity of this order.

In the critical case, $K=1$, $\rho=1$, this formula is no longer valid, because, as we stressed in last section, the most probable overlap is always $Q(t)=1/2$, while the variance, computed in the saddle point approximation, diverges linearly with time, thus destroying the validity of the approximation used. This case needs a more careful treatment, and we will not present it here.

Finally, in the frozen phase the most probable overlap tends to one exponentially with t and the closing probability becomes of order 1 after a time of order $\log N/(1-c)$, where $c = K\rho$.

4.2 Transient times and cycle lengths

We now use the closing probabilities obtained in the previous section to compute the asymptotic distribution of transient times T and cycle lengths L.

In the chaotic phase, inserting the expression for the asymptotic closing probability into the equation (8), we obtain

$$\Pr\{T=t, L=l\} \simeq \exp\left(-\frac{1}{2}(t+l)^2 e^{-\alpha N}\right) e^{-\alpha N}, \tag{36}$$

(from now on, we will shortcut $\alpha(1)$ with α).

So, for finite K, we obtain the same form of the distribution as in the Random Map Model, differing only in the value of α.

For infinite K the overlap distribution is a binomial one with mean value $1-\rho$ and the value of α is exactly given, in the framework of the annealed approximation (without making use of saddle point equations), by $\alpha = -\log(1-\rho)$, which coincides, in the case $\rho = 1/2$ (the so called Random Map), with the value known in the literature (see for instance [5]). This coincidence is obvious because in the case $K \to \infty$ one knows the annealed approximation to be exact. For finite values of K the value of α is smaller and the agreement between the annealed prediction and the simulations is satisfactory.

Introducing the continuous variables $x = t/\tau$ and $y = l/\tau$, with

$$\tau = e^{\frac{1}{2}\alpha N}, \tag{37}$$

the period and transient distribution becomes a continuous distribution with density

$$f(x,y) = e^{-\frac{1}{2}(x+y)^2}. \tag{38}$$

While the form of the distribution predicted by the annealed approximation is not confirmed by our simulations, the N behavior of the time scale τ is in good agreement with the results of the simulations, as it is shown in figure 10 where we compare the value of α obtained by the iteration of equation (26) with twice the exponent of the average cycle length or with the exponent of the variance of the period obtained by a fit of experimental data, for some K value. Other numerical results will be presented later.

On the critical line and in the frozen phase α vanishes, and the cycle length scale grows less than exponentially with N.

In the following we only consider the critical line $\rho_c(K)$ for $K > 1$. In this case, the closing probability is given by

$$\pi_N(t, t+l) \propto \frac{1}{(K-1)\tau} \exp\left(-\left(\frac{2\tau}{t}\right)^2\right), \tag{39}$$

where the time-scale on the critical line is given by

$$\tau = \sqrt{N/(K-1)}. \tag{40}$$

This suggests a kind of universality of the critical line in the Kauffman model: the time scale of the dynamics diverges always as the square root of the number of the elements (except for the case $K = 1$, where the variance in the Gaussian approximation diverges linearly with t, thus destroying the validity of the approximation, and our method fails).

It is easy to see that the exponent α of the closing probability vanishes quadratically when ρ approaches the critical point (except for the case $K = 1$). In fact, putting $\rho = \rho_c + \epsilon$, one finds

$$Q^* = 1 - \frac{2K}{K-1}\epsilon + o(\epsilon), \tag{41}$$

$$\alpha \approx \frac{1}{2}\frac{(1-Q^*)^2}{V} = \frac{(2K)^2}{K-1}\epsilon^2 + o\left(\epsilon^2\right) \tag{42}$$

(of course these expressions are not valid in the limit case of infinite K, where α is equal to $-\log(1-\rho)$).

5 Distribution of weights

In Kauffman networks, configuration space breaks into a number of limit cycles with their own attraction basins. The statistical distribution of the weights of such basins, W_α (the ratio of the number of configurations in the basin to the total number), has been computed in the limit cases $K = 1$ and $K = \infty$ by Derrida and Flyvbjerg [5] and Flyvbjerg and Kjaer [6] respectively. Such distributions turned out to be non trivial, i.e. the average weight is different both from zero and from one in the limit $N \to \infty$.

In the framework of the annealed approximation it is possible to obtain the distribution of the weights for every value of the parameters, at least in the chaotic phase: in this case, and in the limit of an infinite system, it turns out to be the same distribution computed in [5] for the Random Map.

Following [5], it is convenient to start computing the quantities

$$\overline{Y}_p = \sum_\alpha \overline{W_\alpha^p}, \tag{43}$$

that are the moments (of order $p-1$) of the probability to find a basin of weight W. They can be computed by noting that they represent the probability that p configurations chosen at random ultimately meet.

This was done for the case of the Random Map in [5], and the method can be generalized without problems to finite values of K in the framework of the annealed approximation.

The result, as we will see, is that these quantities do not depend, in the chaotic phase and in the limit $N \to \infty$, on the time scale τ of cycle lengths distribution, which is, in our approximation, the only quantity that distinguishes systems with different parameter values in the chaotic phase.

To be more concrete, let us calculate $\overline{Y}_2$. We must generate randomly two trajectories on the same network. We extract the first one and follow it till the time T_1 when it closes on itself, as we have done in the previous section; then we generate another "replica" of the system and we study the overlap $q_{12}(t_1, t_2)$ between the configuration at time t_1 on the first trajectory and the one at time t_2 on the second one.

We define the closing probability of the second trajectory on the first one as the probability that $q_{12}(t_1, t_2)$ is equal to 1, with the condition that the first trajectory is composed of T_1 different configurations, the second of $t_2 - 1$ different ones and none of them has met the first trajectory before time t_2 (we express this condition by the symbol $O(T_1, t_2)$):

$$\pi_N^{(2)}(T_1; t_1, t_2) = P\left\{q_{12}(t_1, t_2) = 1 | O(T_1, t_2)\right\}. \tag{44}$$

The probability that the second trajectory meets the first one after T_2 time steps is then

$$F_N^{(2)}(T_1, T_2)\tilde{\pi}_N^{(2)}(T_1; T_2), \tag{45}$$

where $\tilde{\pi}_N^{(2)}(T_1; t_2)$ is the total closing probability, i.e. the sum over t_1 of $\pi_N^{(2)}(T_1; t_1, t_2)$, and $F_N^{(2)}(T_1, T_2)$ is the probability that the second trajectory is composed of T_2 different configurations and they do not touch the first trajectory before time T_2. One has, in the limit of continuous time variables,

$$F_N^{(2)}(T_1, T_2) = \exp\left(-\int_1^{T_2-1}\left(\tilde{\pi}_N^{(2)}(T_1; t) + \tilde{\pi}_N^{(2)}(t)\right) dt\right), \tag{46}$$

and

$$\overline{Y}_2 = \sum_{T_1, T_2} F_N(T_1) F_N^{(2)}(T_1, T_2) \tilde{\pi}_N(T_1) \tilde{\pi}_N^{(2)}(T_1; T_2), \tag{47}$$

where $\tilde{\pi}_N(t)$ and $F_N(t)$ are respectively the integral closing probability and the opening probability for only one trajectory, as defined in the second section.

In the framework of the annealed approximation, we can imagine that the overlap $q_{12}(t, 0)$ between the initial configuration of the second trajectory and a given configuration of the first one is a binomial variable, and that the distribution of $q_{12}(t+1, 1)$ is obtained from it through the annealed evolution equation.

In this hypothesis, asymptotically in t_1 and t_2, the closing probability in the chaotic phase is given by $\pi_N^{(2)}(T_1; t_1, t_2) = 1/\tau^2$, where $\tau = \exp(\alpha N/2)$ is the cycle length time scale.

Making the substitutions $x = T_1/\tau$, $y = T_2/\tau$ and transforming to continuous variables, one finds, for large values of τ,

$$\overline{Y}_2 \simeq \int_0^\infty \int_0^\infty x^2 e^{-\frac{(x+y)^2}{2}} dx dy = 2/3, \tag{48}$$

which is independent of τ and is the same result found in [5].

This computation generalizes easily to an arbitrary number of trajectories, and in this way one finds for each of the $\overline{Y}_p$ the same value obtained by Derrida and Flyvbjerg for the Random Map. Thus we can argue that in the chaotic phase and in the framework of the annealed approximation $f(W)$ ($f(W)dW$ is equal to the average number of cycles with weight between W and $W+dW$) is given by the same expression valid in the Random Map model:

$$f(W) = \frac{1}{2} W^{-1} (1-W)^{-1/2}. \tag{49}$$

This result is in good agreement with numerical simulations, as we will see in section 7.

6 Average cycle number

In this section we aim to compute the average number of cycles in a network. Following Kauffman, this quantity represents, in his biological metaphor, the typical number of possible cell types for a genome of given length. This quantity is difficult to observe in the simulations, because one must consider also cycles of very small weight, and has been computed only for the Random Map (infinite K, $\rho = 1/2$; [5]).

For this purpose, we have to compute the distribution of the overlap between a configuration at time t and the initial configuration, $P_N(0,t;q)$, which plays the role of the initial distribution for the stochastic process defined by the master equation (12).

It is easy to generalize the annealed approximation in order to get the transition probability of the joint distribution of the overlaps $q(0,t)$ and $q(t-1,t)$: every gene remains at time $t+1$ in the same state in which it was at the previous step with probability $\gamma\left(q(t-1,t)\right)$, where $\gamma(x)$ is defined in the previous section (11).

If the genes unchanged are Nz of the genes whose state was the same in the initial configuration and $N(q(t,t+1)-z)$ of the other genes, the overlap $q(0,t+1)$ is given by $1-q(0,t)-q(t,t+1)+2z$. One then obtains the transition probability

$$\begin{aligned} P\left\{q(0,t+1)=x_0, q(t,t+1)=x_1 \mid q(0,t)=q_0, q(t-1,t)=q_1\right\} = \\ \binom{Nx_0}{Nz}\binom{N(1-x_0)}{N(x_1-z)} \left(\gamma(q_1)\right)^{Nx_1} \left(1-\gamma(q_1)\right)^{N(1-x_1)}, \end{aligned} \tag{50}$$

with

$$z = \frac{1}{2}\left(x_1 + x_0 + q_0 - 1\right), \tag{51}$$

and with the bounds

$$|1-(x_0+x_1)| \leq q_o \leq 1 - |x_0 - x_1|. \tag{52}$$

At time $t = 1$, the variables $q(0,t)$ and $q(t-1,t)$ coincide and their distribution is a binomial one, but the correlations disappear after 2 time steps, at least in the Gaussian approximation, and the stationary joint distribution factorizes into the stationary distribution of $q(t-1,t)$, that we met in the previous section, and a binomial distribution of mean value $1/2$. The most likely value of $q(0,t)$ is always $1/2$, even before the stationary distribution has been reached, and the exponential part of the closing probability $\pi_N(0,t)$ (the probability to have $q(0,t) = 1$) is always $1/2^N$, but the prefactor is different from 1, in the initial steps. As usual, this can be seen putting the distribution in the exponential form and performing the integral with the saddle point method.

The average number of cycles of length l is related to the probability to find a configuration belonging to a cycle of length l, $P\{L = l, T = 0\}$, by the obvious formula

$$\bar{n}_l = \frac{2^N}{l} P\{L = l, T = 0\}. \tag{53}$$

In the framework of the annealed approximation, using the results of section 4 about period and transient distribution and the fact that $\pi_N(0,l) = 1/2^N c_l$ one finds, in the chaotic phase,

$$\bar{n}_l = \frac{c_l}{l} \exp\left(-\frac{1}{2}\frac{l^2}{\tau^2}\right). \tag{54}$$

Summing over l one gets the average number of cycles in a network. Apart from systems on the critical line, c_l is a quantity of order 1, and tends to 1 when l grows; so, substituting c_l by 1 and transforming the sum into an integral, one finds, asymptotically in τ, the following behavior:

$$\sum_{l=1}^{2^N} \bar{n}_l \approx \log \tau = \frac{1}{2}\alpha N, \tag{55}$$

where α is the exponent of the stationary closing probability introduced in the previous section.

The previous formula is valid only in the chaotic phase: in the frozen phase, where α cancels, we are not allowed to put c_l to 1.

For instance, for $l = 2$ one finds

$$\pi_N(0,2) = \frac{1}{2^N} \int e^{-Nf(x)} (x(1-x))^{-1/2} \frac{dx}{\sqrt{2\pi/N}}, \tag{56}$$

where

$$f(x) = x\log(x) + (1-x)\log(1-x) - x\log(\gamma(x)) - (1-x)\log(1-\gamma(x)). \tag{57}$$

($\gamma(x)$ is defined in (11)). The function $f(x)$ is concentrated around the same point Q^* where the stationary distribution of the overlap is also concentrated, *i.e.* the solution of the equation $Q^* = \gamma(Q^*)$. The saddle point method with Gaussian corrections yields the result

$$\pi_N(0,2) = \frac{1}{2^N}\left(1 - \gamma'(Q^*)\right)^{-1}. \tag{58}$$

On the critical line $1 - \gamma'(Q^*) = 1 - (Q^*)^{K-1}$ is of order $1/\sqrt{N}$, as we saw in section 3, and so, using the annealed approximation, we expect that the number of cycles of length 2 grows as $\sqrt{N}$.

This result is in interesting agreement with the simulations by Kauffman, which show that, in the critical networks with $K = 2$ and $\rho = 1/2$, the median of the distribution of the number of cycles in a network grows as

$$\sum_l \overline{n}_l \propto \sqrt{N}. \tag{59}$$

7 Numerical tests of the Annealed Approximation

The annealed approximation was not obtained in a rigourous way, nor we could estimate what it neglects, so we have to justify it by comparing its results with numerical simulations, that will be presented in this section.

Our numerical analysis of the overlap distribution confirms the soundness of the annealed approximation in a certain range of the time variables and reveals deviations from it in another range. The probability distribution of the periods turns out to be different from the one predicted by the annealed approximation, while the time-scale of the period and the first moment of the distribution of the weights are in good agreement with our computation.

All results reported here refer to Kauffman networks with unbiased response functions, *i.e.* the parameter ρ is fixed at the value 0.5, and we will no longer specify it. In such networks, the critical point is at $K = 2$; smaller values of the connectivity belong to the frozen phase, while larger values belong to the chaotic phase.

7.1 Overlap distribution

We measured the overlap $q(t, t+l)$ on the subset of the trajectories not yet closed at time $t + l - 1$, in order to sample with the conditional probability defined in section 2. In this way we also eliminate the "deterministic noise" due to the periodicity of the orbits.

For each value of the parameters, we generated 20,000 sample networks and on each of them we simulated a trajectory measuring the overlap $q\left(C(t), C(t+l)\right)$ till the time when the trajectory closed. So we obtained the histogram of the overlap $q(t, t+l)$ for chosen values of the temporal distance l and t large enough to suppose that the distribution is stationary.

From the annealed approximation we expect the stationary distribution not to depend on l, but in fact we could observe different behaviors for different l values. The data analysis is rather complex since for every point (K, ρ) in parameter space three variables, N, t and l, are involved and different limits, *e.g.* N large at fixed t and l or t and l large at fixed N, may give different behaviors.

We are interested in the exponential part of the stationary distribution, and we estimate it through the formula

$$\alpha_l(q) \approx \frac{1}{N}\left(-\frac{1}{2}\log N - \log P_{Nl}(q)\right). \tag{60}$$

Here the factor $1/\sqrt{N}$ coming from the Stirling expansion of the binomial coefficient, which was neglected in the saddle point computation of the annealed overlap distribution, has been taken into account (actually, this factor is absent for q close to 0 or to 1, which is the value at which we are mainly interested, so we used this formula only for $0.1 \leq q \leq 0.9$, while for $q = 0$ and 1 we do not subtract the term with $\log N$ and we interpolate linearly for intermediate values of q).

The function of q so obtained is compared with the function $\alpha(q)$ numerically computed in the annealed approximation, using the map (26).

Most of our simulations have been performed on systems with connectivity $K = 3$.

The agreement is very good when l, the temporal distance between the configurations, is large (in fig. 1 (b) data are shown for $l = 62$, in networks with $K = 3$ and $N = 50$), while the shape of the distribution is different, especially close to $q = 1$, for small values of l (in fig. 1 (a) we show data for $l = 2$, in networks with $K = 3$ and $N = 50$).

At the critical point the agreement between the annealed approximation and our data worsen, and we expect an agreement even worse in the deep frozen phase, which, with our choice $\rho = 1/2$, corresponds to the analytically studied case $K = 1$ [6].

In figure 2 it is shown the function $\alpha(x)$ measured at the critical point $K = 2$ for two different system sizes, $N = 60$ and $N = 90$. Each point was obtained generating $40,000$ networks and simulating one trajectory over each sample network. The temporal distance between the configurations is $l = 2$ in the left part of figure and $l = 26$ in the right one. Different values of t from 12 to 20 were averaged together. Then we obtained $\alpha(x)$ from the histogram through the formula (60) (while for $K = 3$ we didn't subtract the term with $\log N$ to get $\alpha(1)$, in this case we did, because, as we explained in section 3, we expect the closing probability to be proportional to $N^{-1/2}$ on the critical line).

The solid line is the annealed prediction, computed numerically through the iteration of the map (26). On the critical line, the results of such an iteration are sensitive to the number N of units in which the interval $[0, 1]$ is divided: for instance, the minimum value of $\alpha(x)$ is attained at stationarity at a point x^* such that $1 - x^*$ is of order $N^{-1/2}$. Thus, to compare better the results of the simulations, we used the same value $N = 90$ in the simulations and in our numerical computation. As a result of such a discretization, the minimum value of $\alpha(x)$ in the curve that represents the annealed prediction is attained for $x = 0.933$.

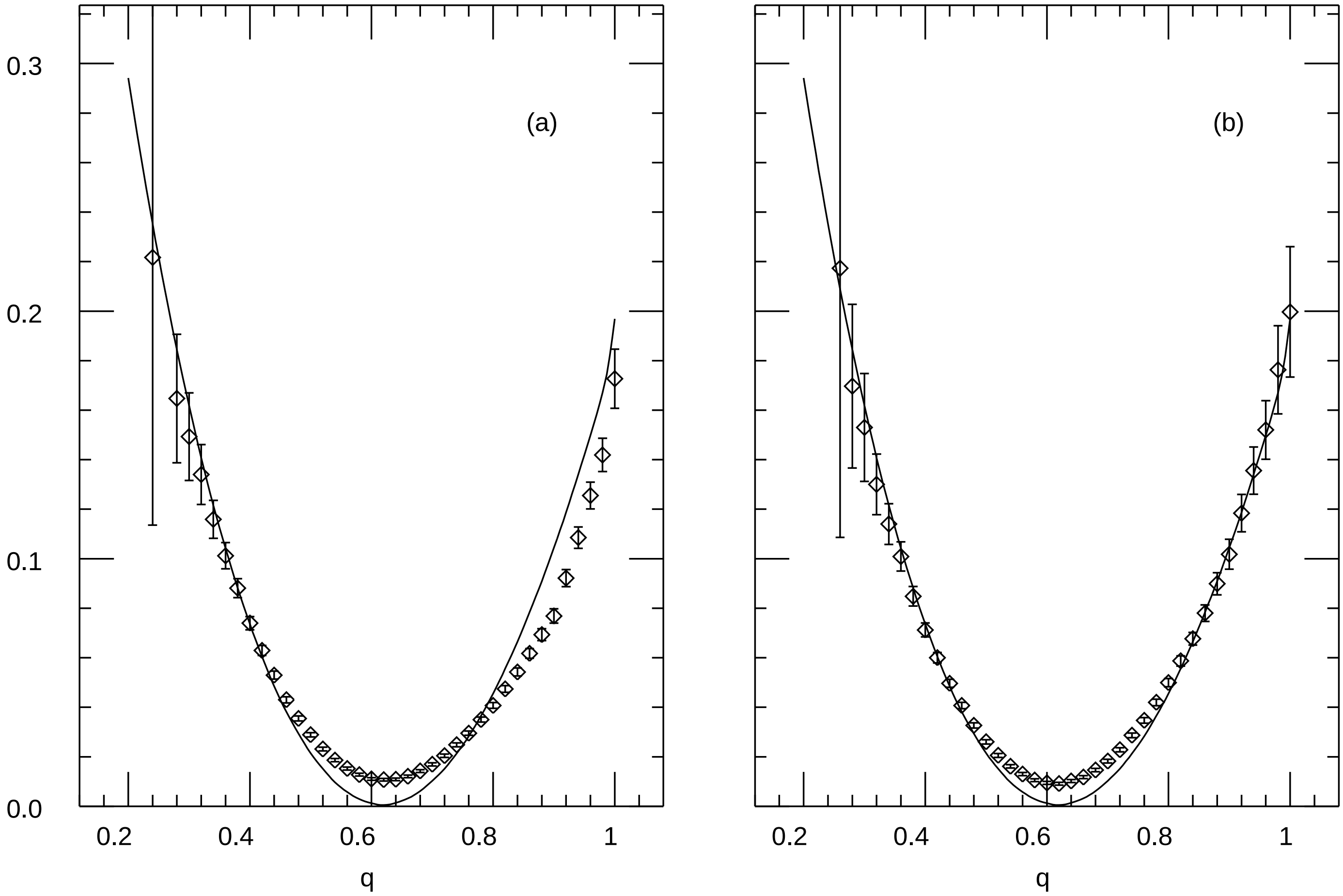

Figure 1: *The function* $\alpha_l(q)$*, which is the exponential part of the distribution of the overlap* $q(t, t+l)$*. The solid line is the annealed prediction, the points plotted come from the simulation of* 20000 *networks with* $K = 3$ *and* $N = 50$*. The distribution is obtained by considering different values of* t *between 50 and 80. (a):* $l = 2$*; (b):* $l = 62$ *(in this case the agreement for* q *close to 1 is much better). The condiction that the trajectory must not be closed at time* $t + l - 1$ *selects respectively 79 and 63 per cent of trajectories in the sample.*

The agreement between data and computation is again better when the temporal distance between the configurations compared is large, but the real distribution is always much broader than the annealed one and the small overlap values are much more likely in the simulation than in the annealed approximation.

We have tried to find out if the overlap between configurations on the limit cycles is statistically different from the overlap between transient configurations. We computed the function $\alpha_l(q)$ imposing two kind of conditions: first, that both configurations are on the cycle, second that both configurations are transient.

In the chaotic phase (we report data with $K = 3$ and $N = 50$) the functions so obtained fit within one or two standard deviations (which are large, because the selection of trajectories is very severe) to the annealed prediction, for the same values of parameters for which the overall distribution gives a good function $\alpha(q)$.

Some systematic differences are visible between the two distributions: the sample taken into the cycles is concentrated around an higher value, and the overlap values close to 1 are more likely in this case than in the other, while the contrary is true for the low values of the overlap, but such an effect is very small.

The situation is different at the critical point where one can see that the most probable value of the overlap between transient configurations is lower than the most probable overlap between configurations of a cycle.

7.2 Measure of the closing probabilities

In section 2 we defined the closing probability $\pi_N(t, t+l)$ as the probability that the configurations at times t and $t+l$ are equal, with the condition that the trajectory is not yet closed at the previous time steps.

We studied the closing probability for l fixed as a function of t, and we saw that in the quenched model the overlap distribution does not go to a stationary distribution as is predicted by the annealed approximation.

This fact means that the simple scheme of the annealed approximation, involving a Markovian stochastic process with time independent transition probability, is not very accurate to describe the quenched evolution of the overlap, as will be pointed out later.

For a fixed system size, the closing probability grows to a maximum value, usually larger than the annealed stationary value (but this depends on l!) and then decreases.

The rate of decrease appears to be independent of l, while the maximum value attained is not: it is larger if l is even and decreases with l for a given parity.

In figure 5 we report data for $K = 3$ in the chaotic phase, and $K = 2$, which, in our case $\rho = 0.5$ is the critical one.

The simulation results are compared to the annealed closing probability exactly computed

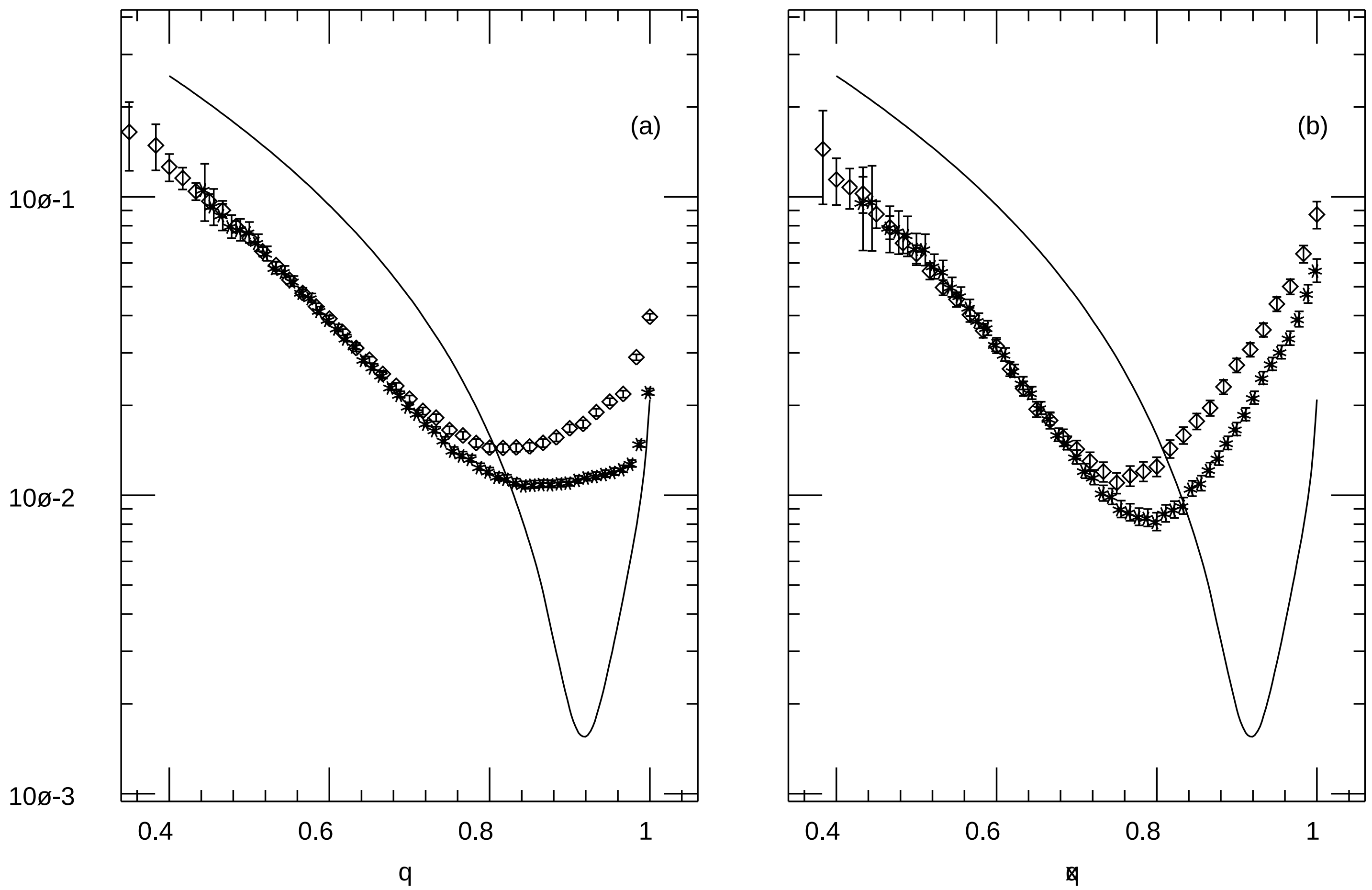

Figure 2: *The function $\alpha_l(q)$, which is the exponential part of the distribution of the overlap $q(t, t+l)$. The solid line is the annealed prediction for $K = 2$ and $N = 90$, the points plotted come from the simulation of $40,000$ networks with $K = 2$, and N respectively equal to 60 (diamonds) and 90 (asterisks). The distribution is obtained considering different values of t between 12 and 20. The vertical scale is logarithmic. (a): $l = 2$; (b): $l = 26$. The condiction that the trajectory must not be closed at time $t + l - 1$ selects respectively 59 and 12.5 per cent of trajectories in the sample for $N = 60$ and 76 and 19 per cent for $N = 90$.*

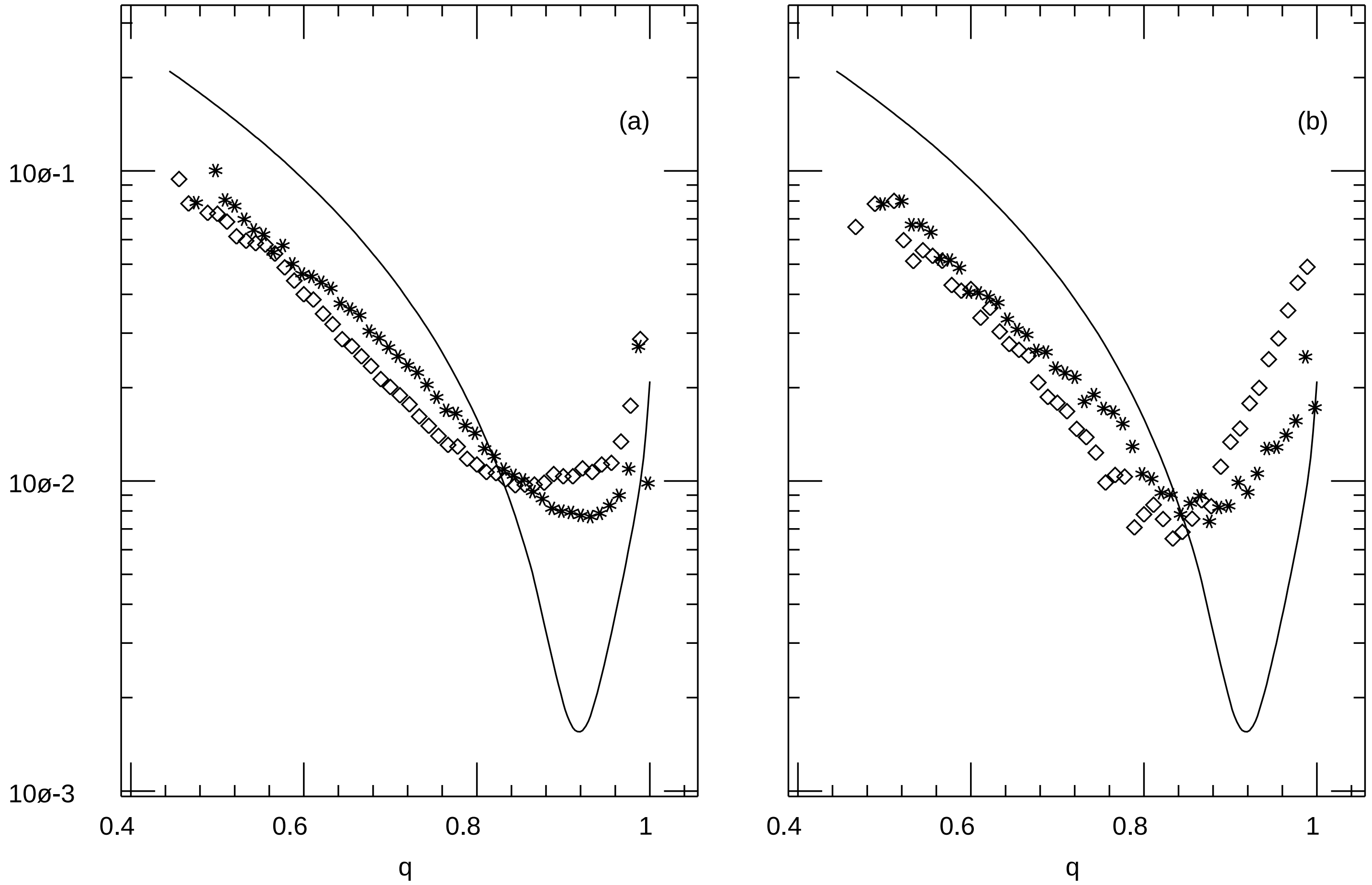

Figure 3: *The function $\alpha_l(q)$ again. Here $K = 2$ and $N = 90$. The solid line is the annealed prediction, the points plotted come from the simulation of $40,000$ networks. The distribution of the overlap $q(t, t+l)$ is obtained considering different values of t between 12 and 20. The vertical scale is logarithmic. The two graphs concern $l = 2$ (a) and $l = 14$ (b) respectively. The asterisks are refered to the overlap of two configurations in a limit cycle, the diamonds to the overlap between transient configurations.*

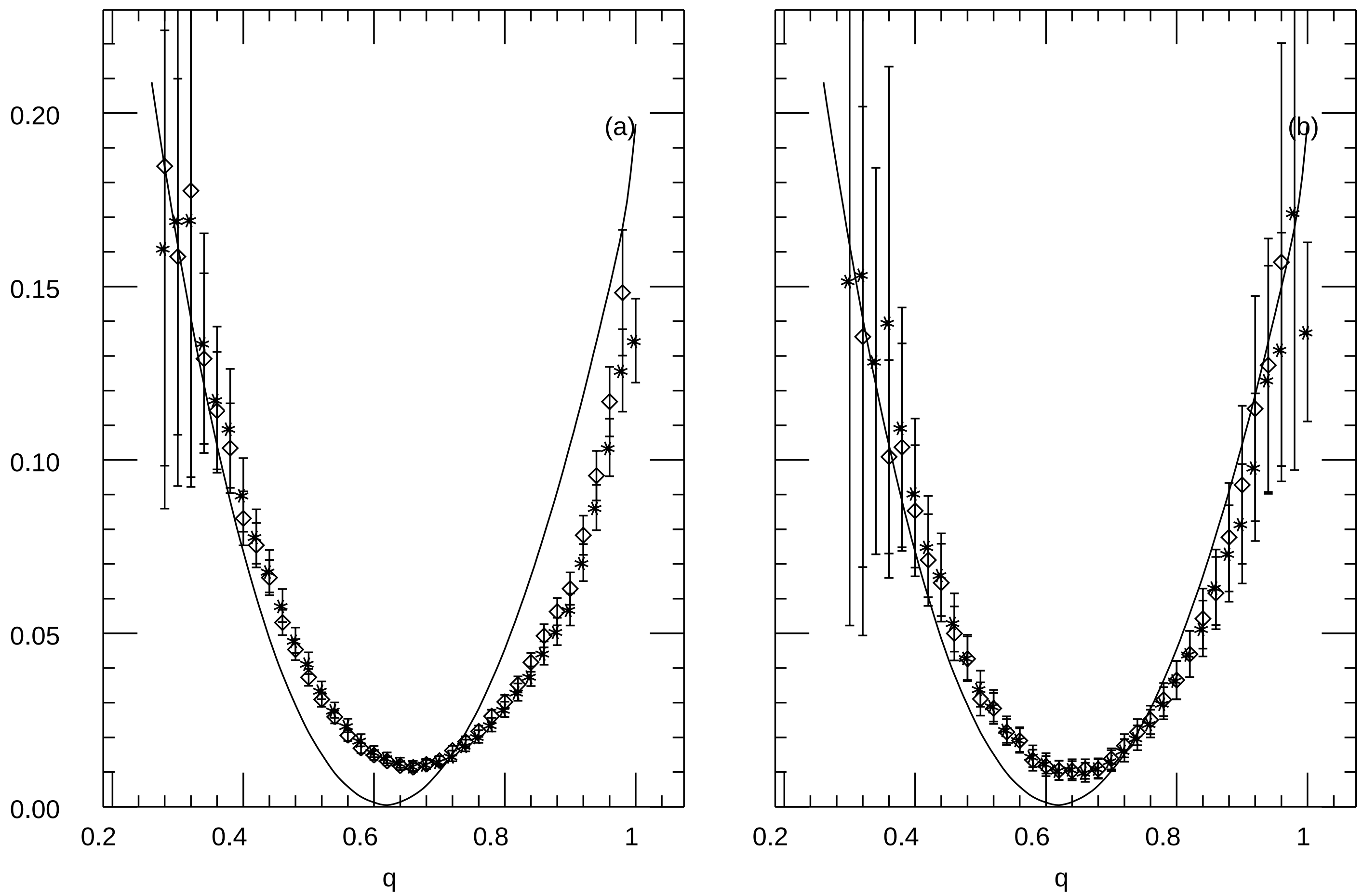

Figure 4: *The function $\alpha_l(q)$ again. Here $K = 3$ and $N = 50$. The solid line is the annealed prediction, the points plotted come from the simulation of $20,000$ networks. The distribution of the overlap $q(t, t+l)$ is obtained considering different values of t between 50 and 80. The two figures are concerned with $l = 2$ (a) and $l = 62$ (b) respectively. The asterisks are refered to the overlap of two configurations in a limit cycle, the diamonds to the overlap between transient configurations.*

iterating equation (12). The time behavior is at first in good agreement (especially for l large), over a fairly large range: in the networks with $K = 2$ and $N = 120$ the annealed and the quenched grow together from 10^{-36} to 10^{-2}. Then, the annealed remains constant while the quenched decreases.

On the other hand, when we compared with the annealed approximation the closing probability measured at fixed times varying system size, we found that the agreement, initially not very good (many standard deviations), becomes very satisfactory with larger systems.

On a graph, the N behavior of the closing probability appears to be an exponential decay with small corrections. It is possible to fit the N behavior with an exponential law,

$$\pi_N(t, t+l) \approx C(t, l) \exp\left(-Na(t, l)\right), \tag{61}$$

also when N is too small for the annealed approximation to hold, but the functions a and C obtained from the fit depend on the interval of N considered; nevertheless, it is interesting to note that, for a fixed interval of N values, a seems to go down to a constant value in the chaotic phase and to zero in the frozen one when t grows, consistently with the annealed picture.

The function $C(t, l)$ seems to decrease always in t as a power law, with an exponent which, in the cases analyzed, ranges from 1.5 to 2.5, and has an oscillating behavior in l (the closure is more likely when the period is even). The latter fact has a reaction on the distribution of cycle lengths, as we shall see.

We now study the integral closing probability, $\tilde{\pi}_N(t)$, which is the sum over t' of $\pi_N(t', t)$. As usual, we considered two values of K, 2 and 3, in and near to the frozen phase, and N ranging from 60 to 180 for $K = 2$ and from 20 to 50 for $K = 3$.

$\tilde{\pi}_N(t)$ also has a non monotone behavior for t not too large: it increases to a maximum value and then, surprisingly, it starts decreasing. Of course, it has to increase at large times, because it has to fulfill the normalization condition $\tilde{\pi}\left(2^N\right) = 1$, but in our simulations it keeps decreasing at all times accessible to measurement. On the other hand, in the framework of the annealed approximation, we expect that it grows linearly with t.

If one remembers the previous results, one realizes that this behaviour is just a consequence of the decrease in time of the closing probability. In a few paragraphs, we will present other data that allow us to interpret this result, which is the source of the main discrepancies between the behaviour of the quenched system and the predictions of the annealed approximation, above all for what concerns the distribution of cycle lengths.

We show the temporal behavior of $\tilde{\pi}_N(t)$ for different system sizes in the chaotic phase (fig. 7) and at the critical point (fig. 6). The two plots have qualitatively the same shape, but in the critical point the integral closing probability decreases very slowly with system size N, while in the chaotic phase there is a much stronger N dependence at fixed time, which is nearly an exponential decrease.

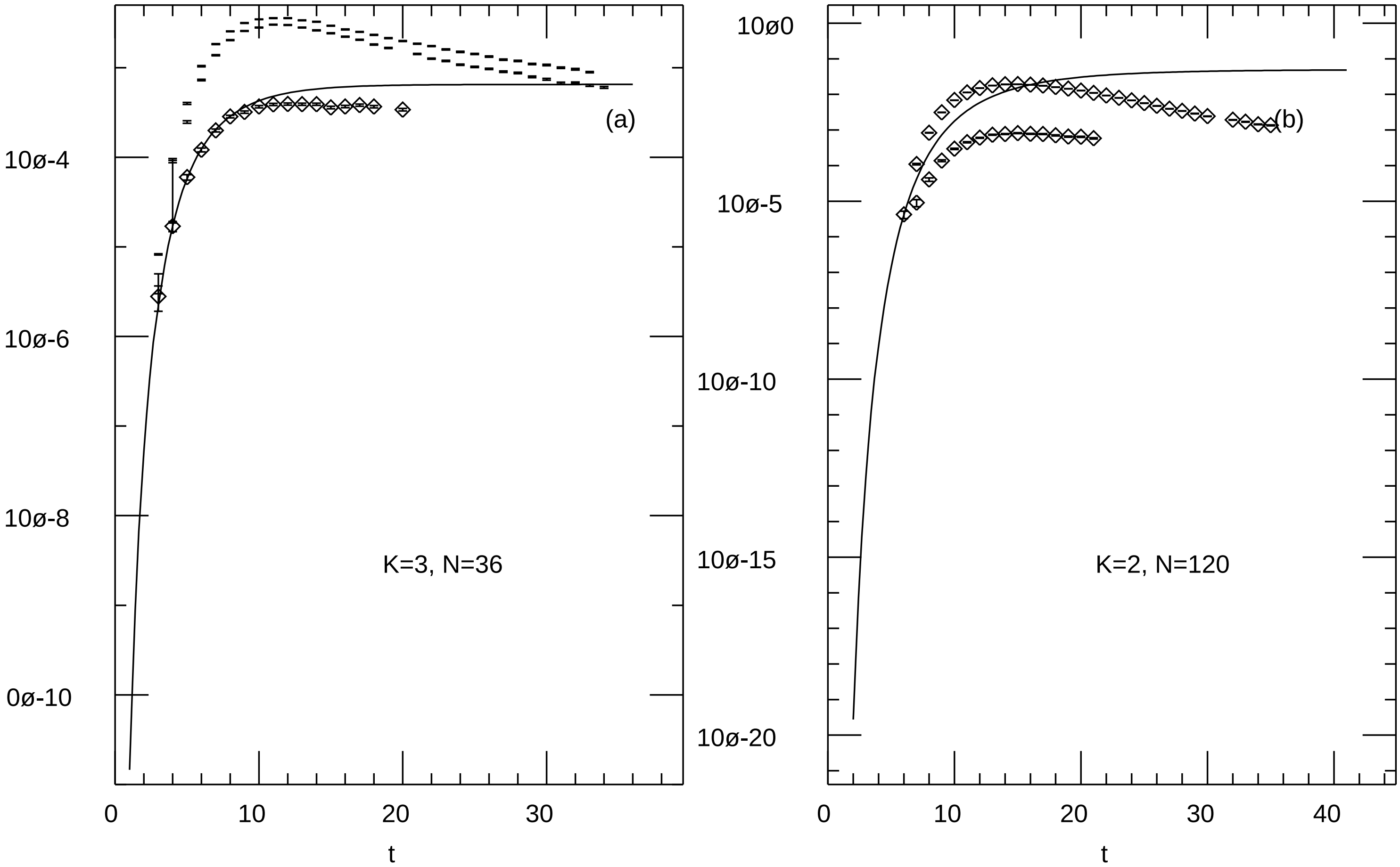

Figure 5: *Closing probability* $\pi_N(t, t+l)$. *The annealed value with* $l = 1$ *is shown in solid. The ordinate scale is logarithmic. (a): The system is in the chaotic phase,* $K = 3$ *and* $N = 36$. *Data correspond to* $l = 1$ *(middle),* $l = 2$ *(upper line: the closing probability is systematically higher) and to* $l = 15$ *(diamonds; the agreement with the annealed approximation is better). (b): The system is at the critical point,* $K = 2$ *and* $N = 120$. *Data correspond to* $l = 2$ *(upper line) and to* $l = 15$ *(diamonds). The first point of the annealed computation,* $P = 7.52 \cdot 10^{-37}$, *is not shown.*

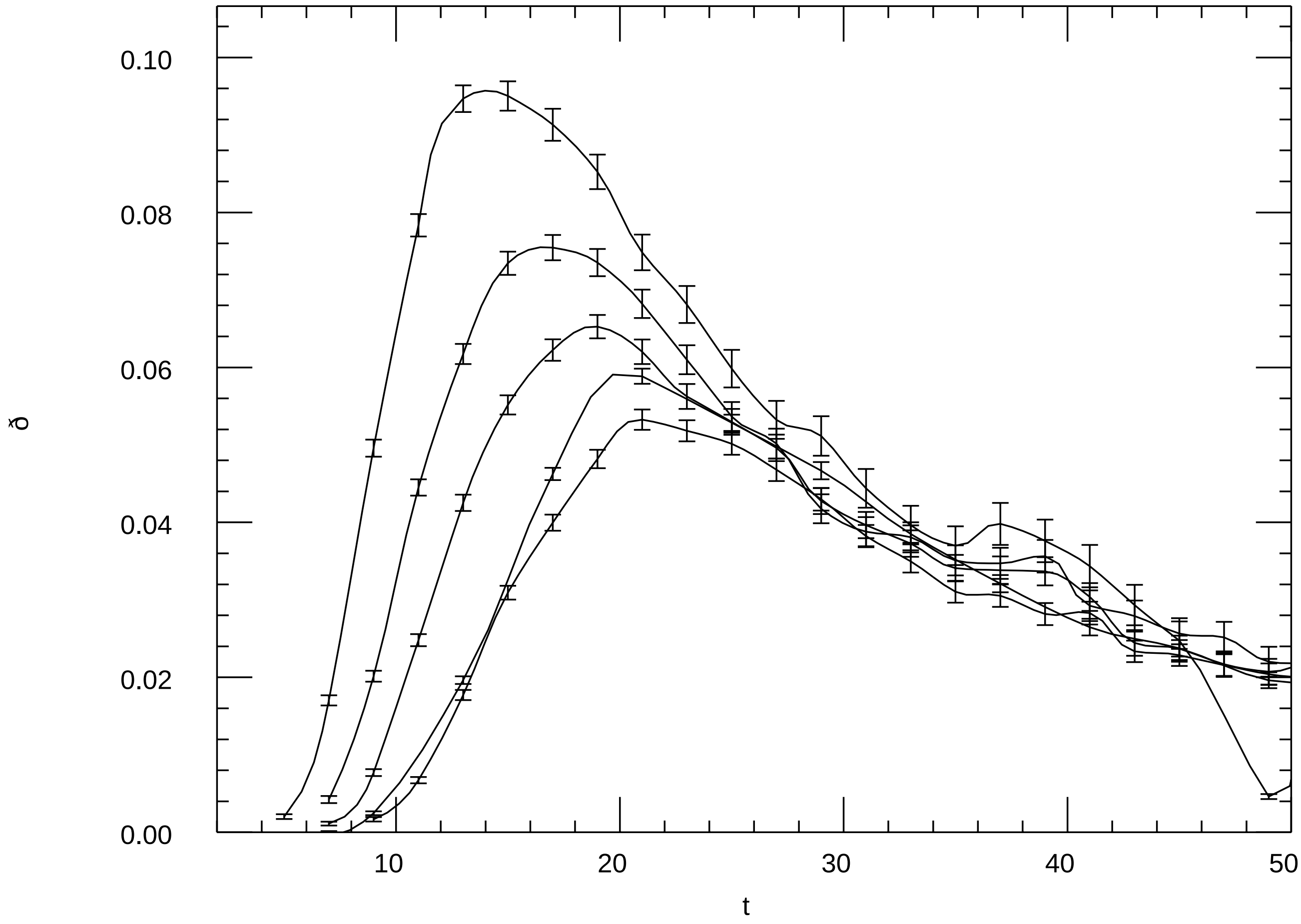

Figure 6: *Integral closing probability,* $\tilde{\pi}_N(t)$ *in the frozen phase* $(K = 2)$ *for different system sizes:* $N = 60, 90, 120, 150, 180$. $\tilde{\pi}_N$ *reduces when* N *grows.*

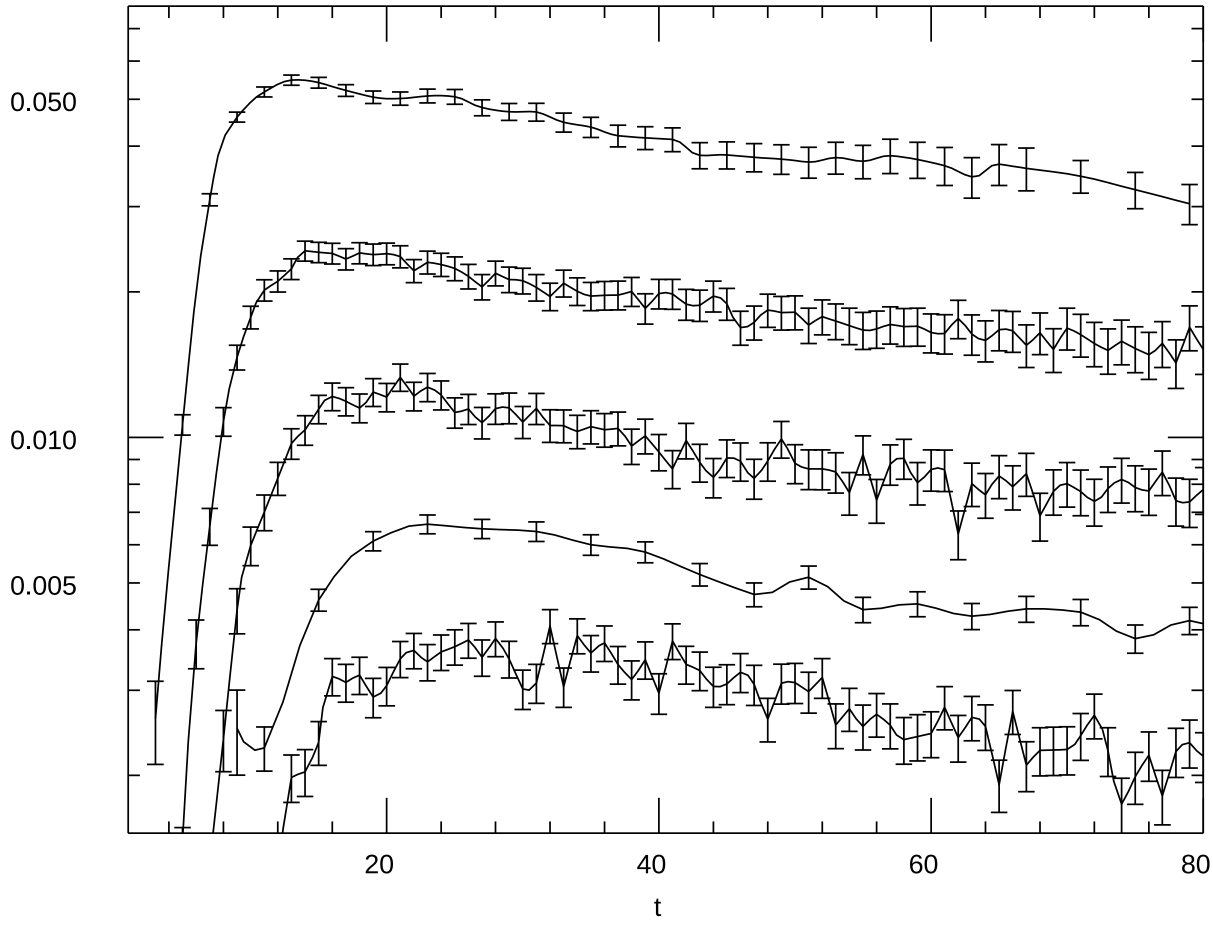

Figure 7: *Integral closing probability, $\tilde{\pi}_N(t)$ in the chaotic phase ($K = 3$) for different system sizes: $N = 20, 30, 40, 50, 60$. $\tilde{\pi}_N$ reduces when N grows.*

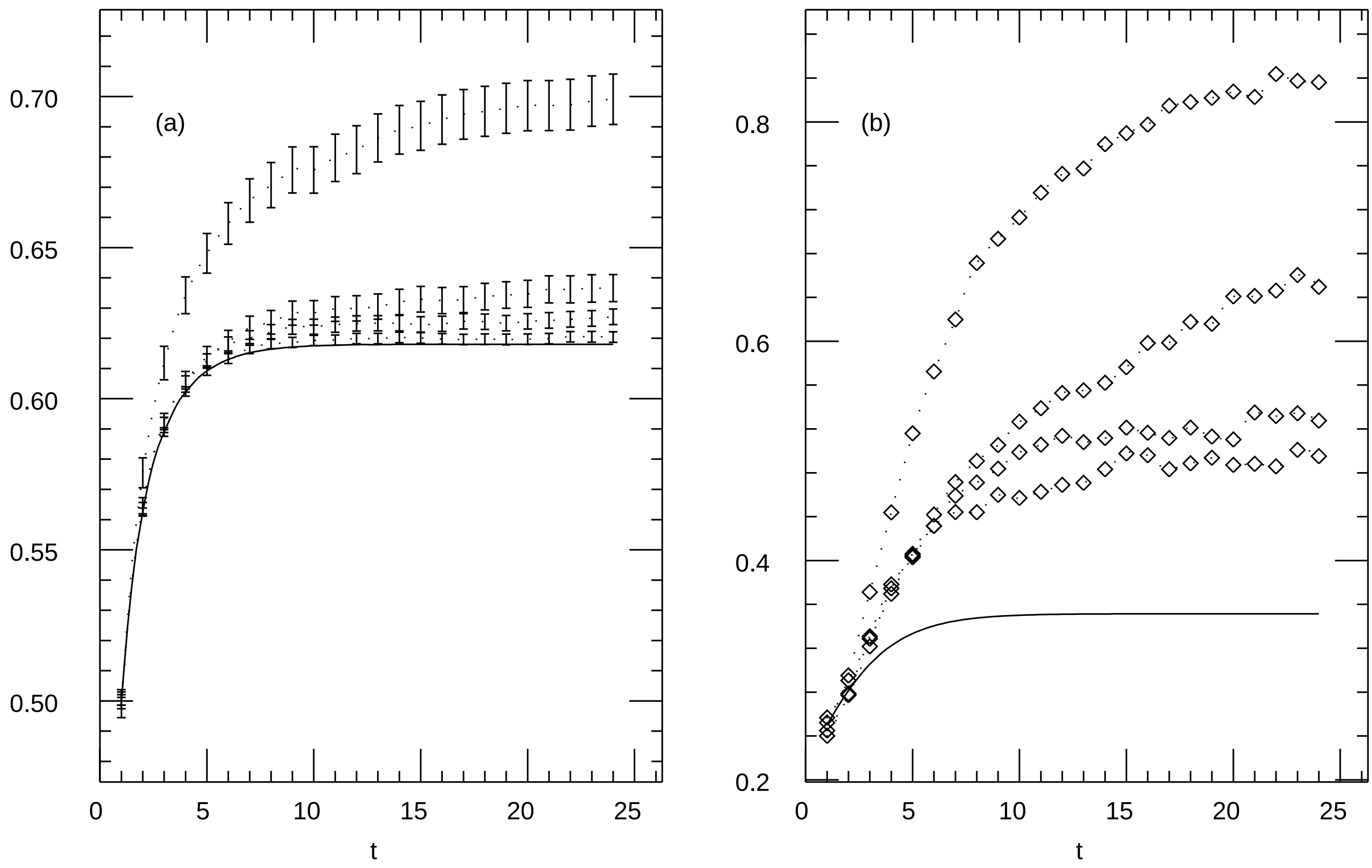

Figure 8: *(a): Average overlap between successive configurations The solid line shows the annealed equation,* $Q(t+1) = \frac{1}{2}\left(1 + Q(t)^K\right)$. *(b) Variance of the overlap between successive configurations. The solid line shows the annealed equation,* $V(t+1) = Q(t+1)\,(1 - Q(t+1)) + \left(\frac{K}{2}Q(t)^{K-1}\right)^2 V(t)$. *The system is in the chaotic phase,* $K = 3$*; data are shown for* $N = 30$*, 80, 200 and 400; the larger is* N*, the lower is the curve.*

Not only the closing probability, which is the value of the overlap distribution relative to $q = 1$, but also the mean value of the distribution ultimately decreases with time.

We measured the mean overlap between successive configurations, $\bar{q}(t, t + 1)$, on the subset of the trajectories not yet closed at time $t + l$, and observed that this quantity has a non-monotone behavior as a function of time t. Initially it increases from the value 1/2, reaches a maximum value at about the same time at which the integral closing probability has its maximum value, and then decreases. This behaviour is more evident on the critical line, but it is still present in the chaotic phase for low values of K.

Fig. 9 shows the mean and the variance of the distance $d = 1 - q$ of successors on trajectories not yet closed and the integral closing probability for $K = 2$ and $N = 90$. The asymptotic increase of the distance corresponds to the decrease of the overlap described above.

We also measured the mean overlap between successive configurations, $\bar{q}(t, t+1)$, without limiting the sample to the only trajectories not yet closed. We considered three different connectivities in the chaotic phase (K=3,4 and 5) and varied N.

As a function of t, this quantity increases to a stationary value which tends, when N grows, to the same value Q^* computed within the annealed approximation. This fact is not surprising: Derrida and Weisbuch [10] and Hilorst and Nijmayer [11] showed that the equation for the evolution of the mean overlap is the same in the annealed approximation and in the quenched system up to times of order $\log N$, and we see that in the large systems the mean overlap becomes stationary in time very early.

The corrections to the annealed value Q^* are positive and large even for system sizes of the order of hundreds. As one could expect, when the temporal distance between the configurations compared is larger, the corrections are smaller.

These results indicate that when we impose the condition that the trajectory is not yet closed at a certain time we select, from a point on, trajectories characterized by smaller and smaller overlaps: this condition, that we could not impose in the framework of the annealed approximation, is thus the cause of the leftward movement of the overlap distribution that is not expected on the grounds of our approximation.

7.3 Properties of the cycles

Let us turn now to the quantities that have attracted the major interest from the first investigators of Kauffman model: the statistical properties of the lengths of limit cycles and of the weights of their attraction basins.

We will mainly present here results concerning the chaotic phase, which agree better with the annealed approximation. Some characteristics of the critical line (which is the only region of parameter space susceptible to have a biological interpretation) seem also to agree with the annealed predictions, but we will present them elsewhere, because they need a

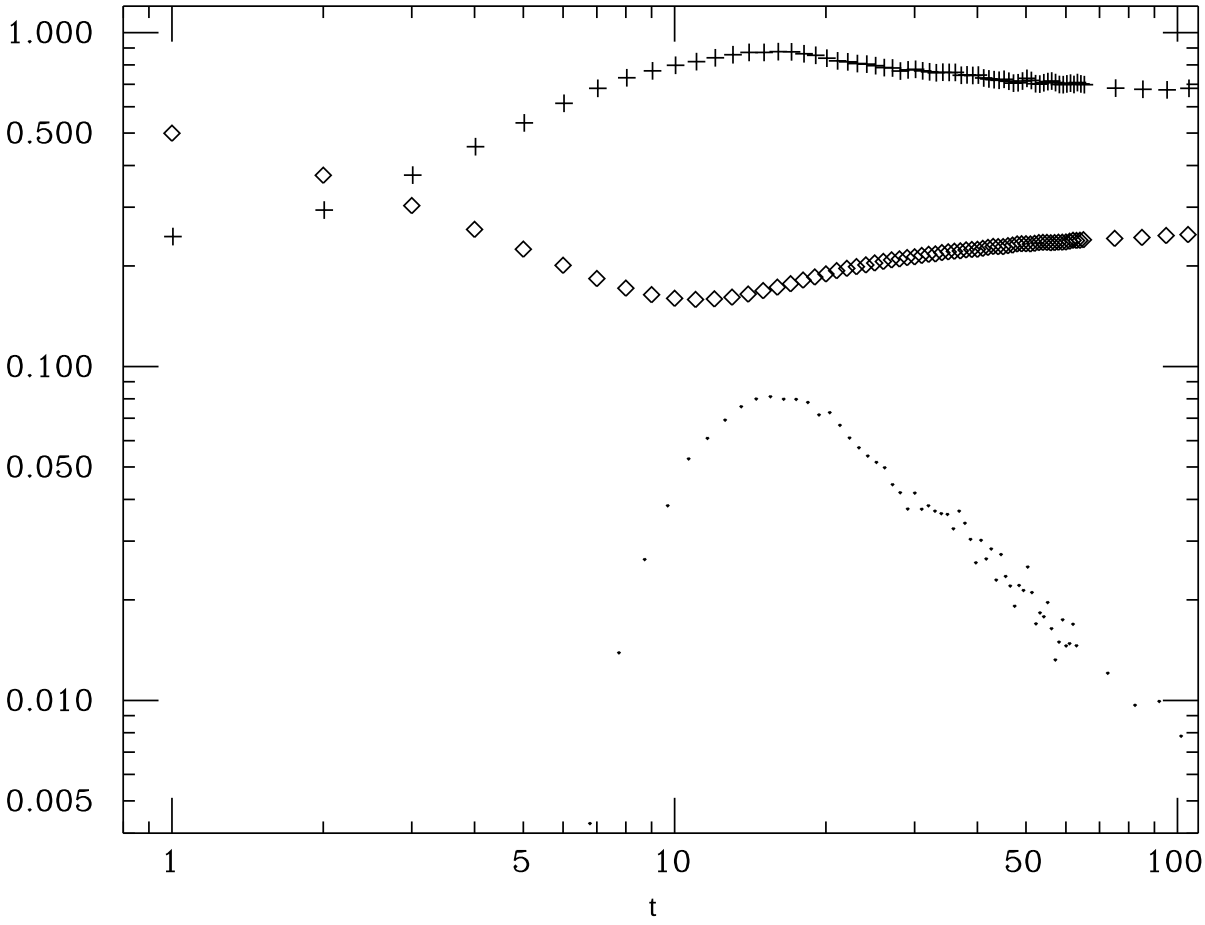

Figure 9: *Average ($\diamondsuit$) and variance (+) of the distance $d = 1 - q$ between successive configurations on trajectories not yet closed at time $t-1$, as a function of t. The integral closing probability, is also shown ($\cdot$). $K = 2$, $N = 90$, $10,000$ sample networks.*

somewhat wider treatment.

Our simulations confirm well the predictions of the annealed approximation concerning the time scale of the period and of the transient distribution.

In the chaotic phase these quantities have an exponential behavior, $\exp(\alpha N/2)$. We simulated systems with K ranging from 3 to 6 and several values of N (from 10 to 100 for $K = 3$, from 16 to 40 for $K = 4$, from 16 to 31 for $K = 5$ and from 6 to 28 for $K = 6$), generating at least 10,000 samples of each network and few configurations on each network (except for the case $K = 3$, where we generated hundreds of initial configurations in order to compare the variance of cycle length on a given network with the variance between different networks).

We fixed a maximum time for the simulation, exponentially increasing with N; this leads us to underestimate the average cycle length because some trajectories never completed their limit cycle; they amounted to about one percent in the worst case, and zero in many other cases. We do not think that this fact will significantly affect the determination of the exponents of the average period.

In fig.10 we plot the exponent α defined by $\tau = \exp(\alpha N/2)$, where α is the time scale of the system, and compare it with the annealed prediction, numerically computed.

It is possible to measure in the simulations four different exponents: α_L, which is twice the exponent of the mean period; α_V, which is the exponent of the variance of the period; α_T, which is twice the exponent of the mean transient time needed to reach the cycle; and, finally, α_P, the exponent of the time-scale of the period distribution obtained by fitting the probability to find a period larger than t with the function $\exp\left(-(t/\tau)^{\beta}\right)$.

These exponents were obtained through fits of data with different N values leading to values that coincide within a few percent, so the whole picture is coherent. In the figure they can hardly be distinguished, so we show only α_L, with no error bars, because the fits of our data are affected by finite size effects whose importance is difficult to estimate. The last point represents the analytical result valid for the Random Map ($K = \infty$), where $\alpha = \log 2$ [5]. The agreement is quite good, even with not very large systems.

The largest deviation was found for the case $K = 3$, where the annealed approximation gives $\alpha = 0.182$ while our simulations give $\alpha_L = 0.210$, $\alpha_T = 0.204$, $\alpha_P = 0.21$ and $\alpha_V = 0.206$ for the average variance in a given network and 0.205 for the variance between different networks. In this case, we performed the annealed computation with the map (26) and the same values of N used in the simulations, in order to obtain the corrections due to the discretization, whose effect is to increase α. In figure 11 we compare twice the logarithm of the average period (diamonds) with $N\alpha(N)$, where $\alpha(N)$ is computed with the annealed approximation. The discrepancies now are reduced with respect to the asymptotic α value, but the simulation data still seem to be systematically larger and the angular coefficients of the two asymptotic straight lines are different, respectively 0.21 for the simulation data and 0.18 for data from the annealed approximation.

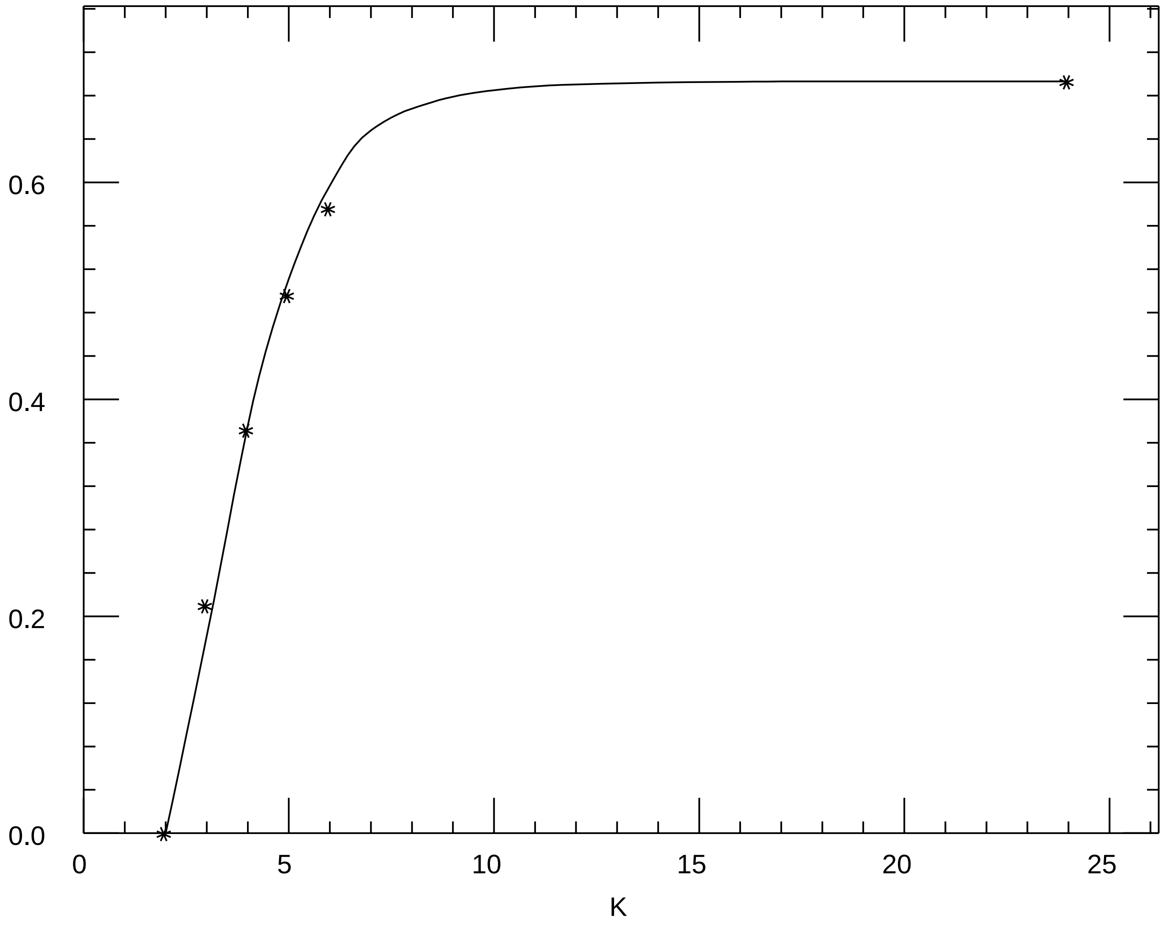

Figure 10: *The function $\alpha(K)$, where α is twice the exponent of the characteristic time-scale of the period and transient distribution, $\tau = \exp(\alpha N/2)$, for $\rho = 0.5$, as obtained from the annealed approximation. The asterisks are obtained from a fit of simulation data; the last one is the theoretical result for the Random Map, $\alpha = \log 2$.*

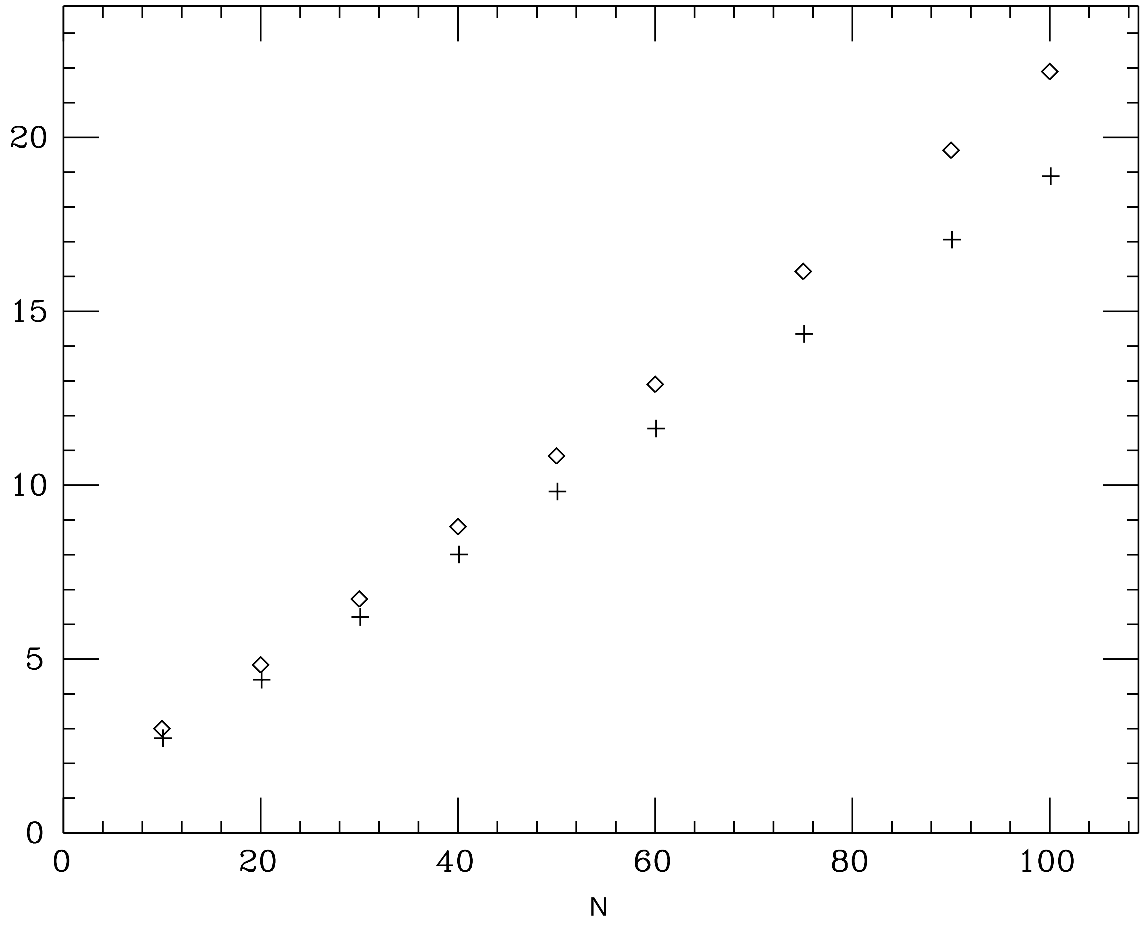

Figure 11: *($\diamondsuit$) : twice the logarithm of the average period in networks with $K = 3$ and N nodes (simulation); (+) : the annealed coefficient α computed in the saddle point approximation for a system of N nodes, times N.*

In the simulations with $K = 3$ we measured also $F_N(t)$, defined in the second section as the probability that a random chosen trajectory is not yet closed at time t, and related to the closing probability by equation

$$\log\left(F_N(t)\right) = \int_1^t \tilde{\pi}_N(\tau)d\tau. \tag{62}$$

According to the annealed approximation, one should find $\log\left(F_N(t)\right) \propto t^2$, but we found that this quantity is a power law, t^β, where the exponent β is less than one, so that $F_N(t)$ has a stretched exponential behavior.

This feature is just a consequence of the fact that the integral closing probability is a decreasing function of t from a certain point on, while we expect it to increase linearly in t in the framework of the annealed approximation.

The distribution of cycle length L and transient time T are simply related to $F_N(t)$:

$$Prob\{T = t, L = l\} = F_N(t + l)\pi_N(t, t + l). \tag{63}$$

The period distribution is much easier to measure in a simulation than $F_N(t)$: to find a cycle one only has to keep in memory a single reference configuration, while the measure of $F_N(t)$ requires the whole trajectory until it closes. For small values of K in the chaotic phase, the probability to find a cycle longer than t decays as a stretched exponential, $\exp\left(-(t/\tau)^\beta\right)$. We expect that the exponent β coincides with the one obtained through $F_N(t)$. The stretched exponential behavior of the cycle lengths distribution is thus a remarkable manifestation of the decrease in time of the closing probabilities.

We did a systematic study of the exponent β obtained through the period distribution. A striking result of this analysis is the fact that the exponent β is a slowly decreasing function of N (while, within the annealed approximation, we expect its value to be 2). With our data, we can not understand if β vanishes in the limit of an infinite system, or if it remains finite, but we favour this second hypothesis.

When K is equal to 3 the value of β is 0.91 for $N = 10$ and 0.39 for $N = 100$. For larger K values, β is always a decreasing function of N but its value is larger: for example, when K is equal to 5 the value of β is 1.06 for $N = 16$ and 0.91 for $N = 36$. In the limit $K \to \infty$ in which the annealed approximation is exact β must go to the annealed value, $\beta = 2$. At the critical point $K = 2$ β vanishes and the distribution function of cycle lengths decreases as a power law[3]. We will present elsewhere our results for critical systems.

In contrast, the time scale τ of cycle lengths distribution is in good agreement with the predictions of the annealed approximation, as we saw above. In the chaotic phase, it grows exponentially with N, with the same exponent α_L of the average cycle length.

[3] A recent preprint makes a careful study of the period distribution in Kauffman networks with $K = 2$ [17].

At the critical point the definition of the time scale of the period distribution gives some problems, because the distribution has a power law behaviour. Nonetheless, we will argue from our data that it is possible to define a time scale, at least in some range of the distribution, and that this time scale is proportional to $\sqrt{N}$, as predicted by the annealed approximation. With our data, however, we can not exclude the possibility that the time scale so defined makes sense only for N not too large, because we can not exclude that the exponent of the power law vanishes in the limit $N \to \infty$. All these subjects will be discussed in [14].

Another observation very useful to understand the limits of the annealed approximation is the following: when we measured the variance of the mean cycle length of a single network, $\overline{\langle L \rangle^2} - \overline{\langle L \rangle}^2$, we found a value much higher than the mean variance of cycle length in a network, $\overline{\langle L^2 \rangle - \langle L \rangle^2}$; in other words, we can say that there are networks characterized by long cycles and networks characterized by short ones. It is natural to wonder whether one can define some structural property of the network that accounts for this behaviour.

The most natural candidate for this role is the number of relevant elements in a network. We defined the relevant elements generalizing the definition due to Flyvbjerg and Kjaer valid for networks of connectivity $K = 1$ [6]. Flyvbjerg [12] showed that in Kauffman networks there are elements that evolve to a state constant in time and independent on the initial configuration (the fraction of this elements can be used as order parameter), and called them *stable elements*. The relevant elements are the ones that are not stable and that control some other relevant element.

There is a strong correlation between the number of relevant elements in a network and its average cycle length. Moreover, the study of the distribution of the number of relevant elements gives useful information for the study of the critical line and of the chaotic phase. We will present the results of our numerical study of relevant elements in [15].

The period distribution has another remarkable feature, unexpected in the context of the annealed approximation: cycles of even length are much more likely than odd ones (see figures 12 and 13). This fact can also be seen in terms of closing probabilities: the closing probability $\pi_N(t, t+l)$ is higher for even l than for odd l, for the same value of t.

We have seen in this fact a clue of a non trivial distribution of local periods in a cycle, in the sense that the relevant elements in the network (*i.e.* the ones which are not constant and which send signals to other non constant elements) are subdivided into independent modules each characterized by its own period, and the cycle of the network is the composition of the independent cycles of its modules.

This fact is very important to understand the structure of Kauffman networks on the critical line and in the frozen phase. We studied numerically the modular structure of critical Kauffman networks, obtaining interesting results that will be presented in a forecoming paper [16].

Finally, we measured the histogram of the transient time needed to reach the cycles, simulating on systems with $K = 2$ and N ranging from 30 to 90 and $K = 3$ and N ranging from 10 to 100. Such distribution has a peak corresponding to the time t where the closing probability reaches its maximum value, and then decreases approximately with the same law that we saw for the period distribution. Figures 12 and 13 compare period and transient distributions, respectively for a system in the chaotic phase ($K = 3$ and $N = 50$) and a system at the critical point ($K = 2$ and $N = 120$).

7.4 Distribution of weights

In the framework of the annealed approximation, it turns out that the distribution of weights in the chaotic phase of Kauffman networks is universal and coincides with the weight distribution of the Random Map, computed by Derrida and Flyvbjerg [5].

To compare this prediction with numerical data we measured $\overline{Y_2}$, the first moment of the weights distribution (equation 43), for the system with $K = 3$, which is the chaotic system closest to the critical line among the ones we studied, pursuing the simulations performed by Derrida and Flyvbjerg in [3] toward systems of larger size. We expect that the agreement with the annealed approximation should improve when the connectivity K grows (the approximation is exact for the completely connected networks), and the results presented in [3] for other K values are consistent with this conjecture.

For each system size we generated 10000 sample networks and two configurations at random on each of them and we measured $\overline{Y_2}$ as the probability that they end up on the same cycle, as it was done in [3].

In this way we obtained a curve of $\overline{Y_2}$ versus N which reaches a minimum value of $\overline{Y}_2 = 0.590$ at $N = 40$ (this was the largest size simulated in [3]) and then increases with N. The largest system that we simulated ($N = 100$) still looks far from having attained the stationary value, so we are unable to say if it agrees with the annealed prediction $\overline{Y}_2 = 2/3$, but it is very close to that value (figure 14).

This non monotonic behavior can be interpreted at the light of the modular structure of Kauffman networks [16].

We measured also the density of the probability to find a cycle of weight W, $g(W) = Wf(W)$. Figure 15 compares the histogram obtained generating at random 1000 networks with $K = 3$ and $N = 40$ and simulating 250 trajectories on each of them with the annealed prediction,

$$g(W) = \frac{1}{2}(1 - W)^{-1/2}, \tag{64}$$

which is valid in the infinite N limit. The agreement, as it is possible to see, is not so bad, even if we chose the N value where $\overline{Y_2}$ has the largest difference from the annealed value.

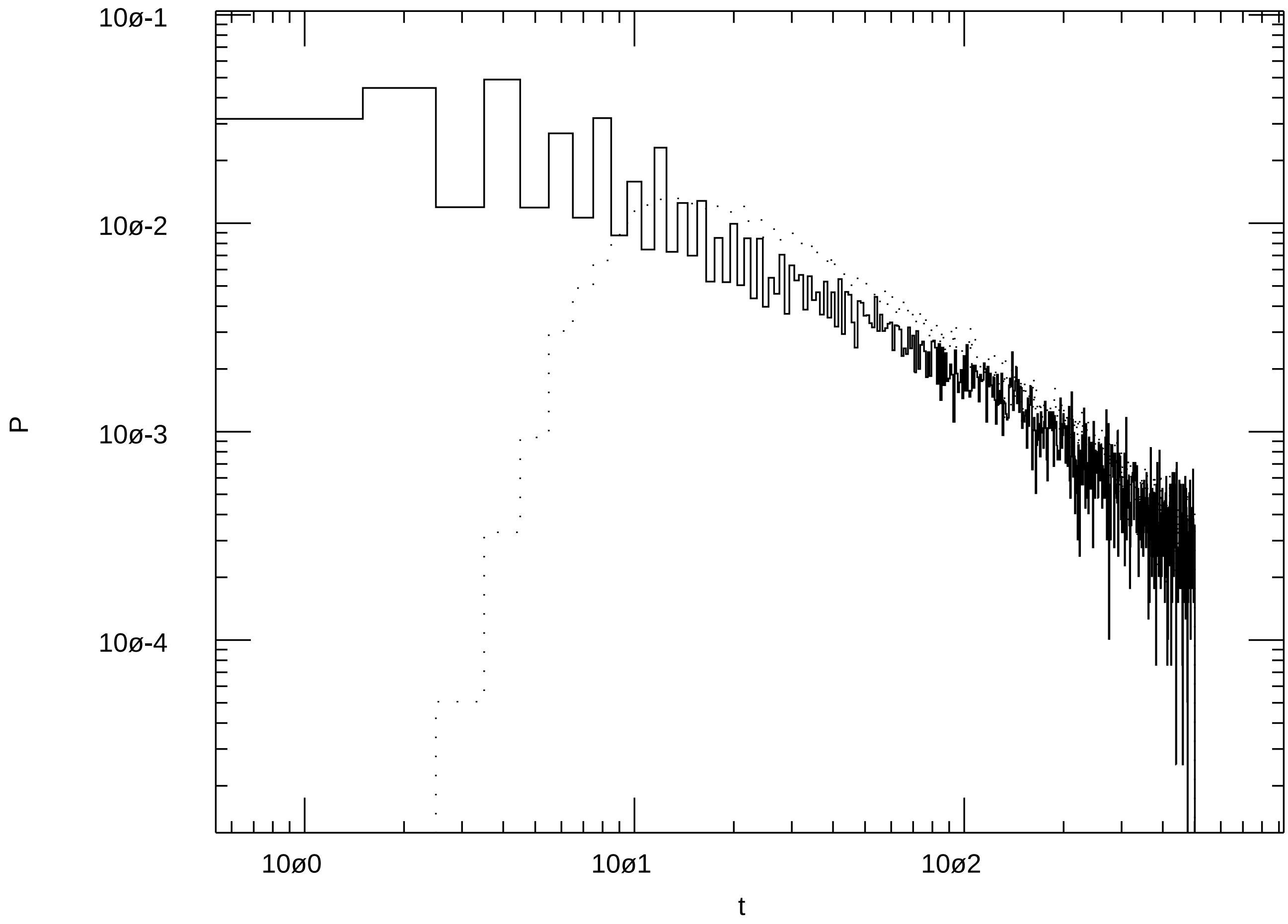

Figure 12: *Histogram of cycle (solid line) and transient (dots) length for a system in the chaotic phase,* $K = 3$ *and* $N = 50$. $10,000$ *sample networks were generated.*

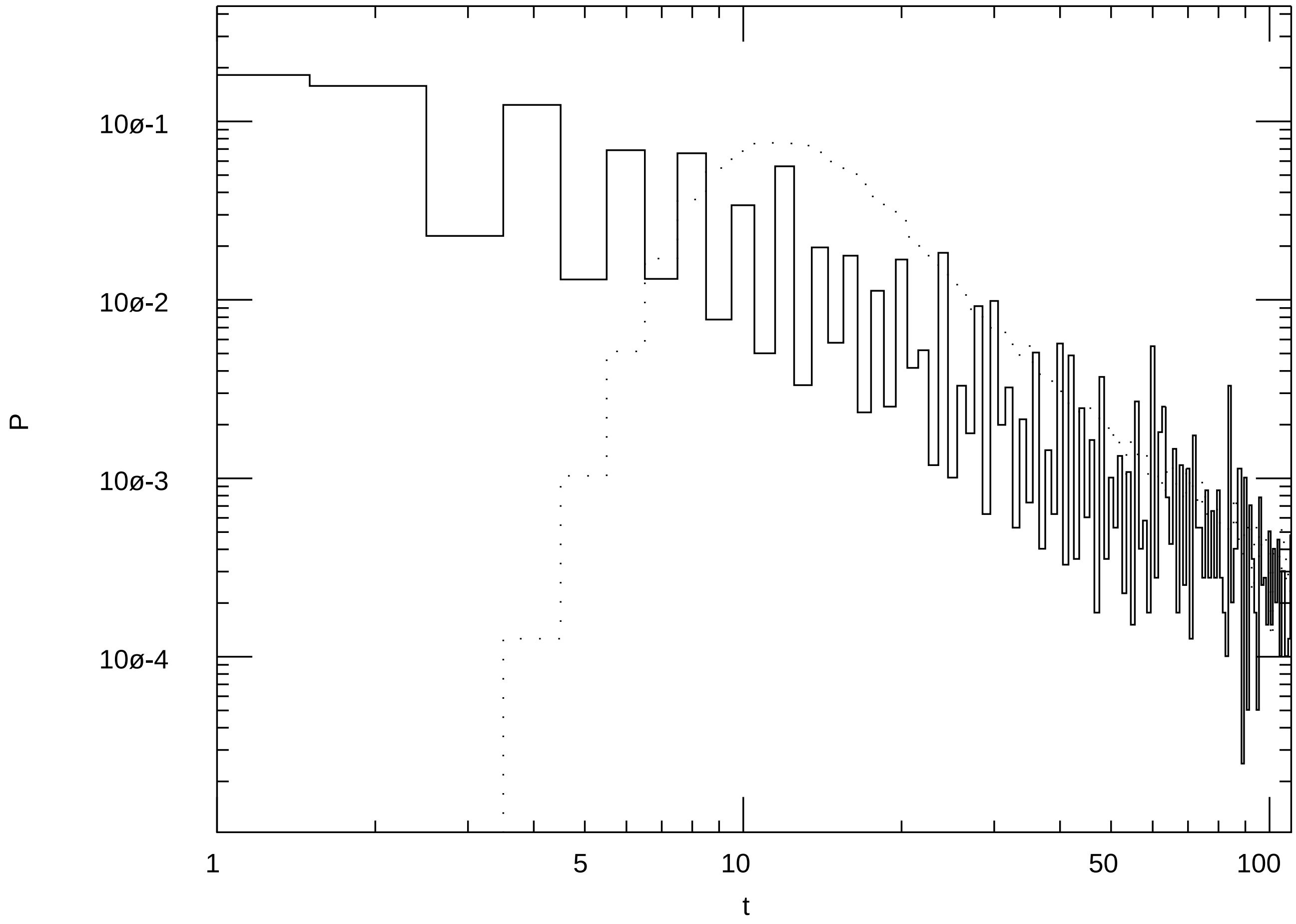

Figure 13: *Histogram of cycle (solid line) and transient (dots) length for a system in the frozen phase, $K = 2$ and $N = 120$. $10,000$ sample networks were generated.*

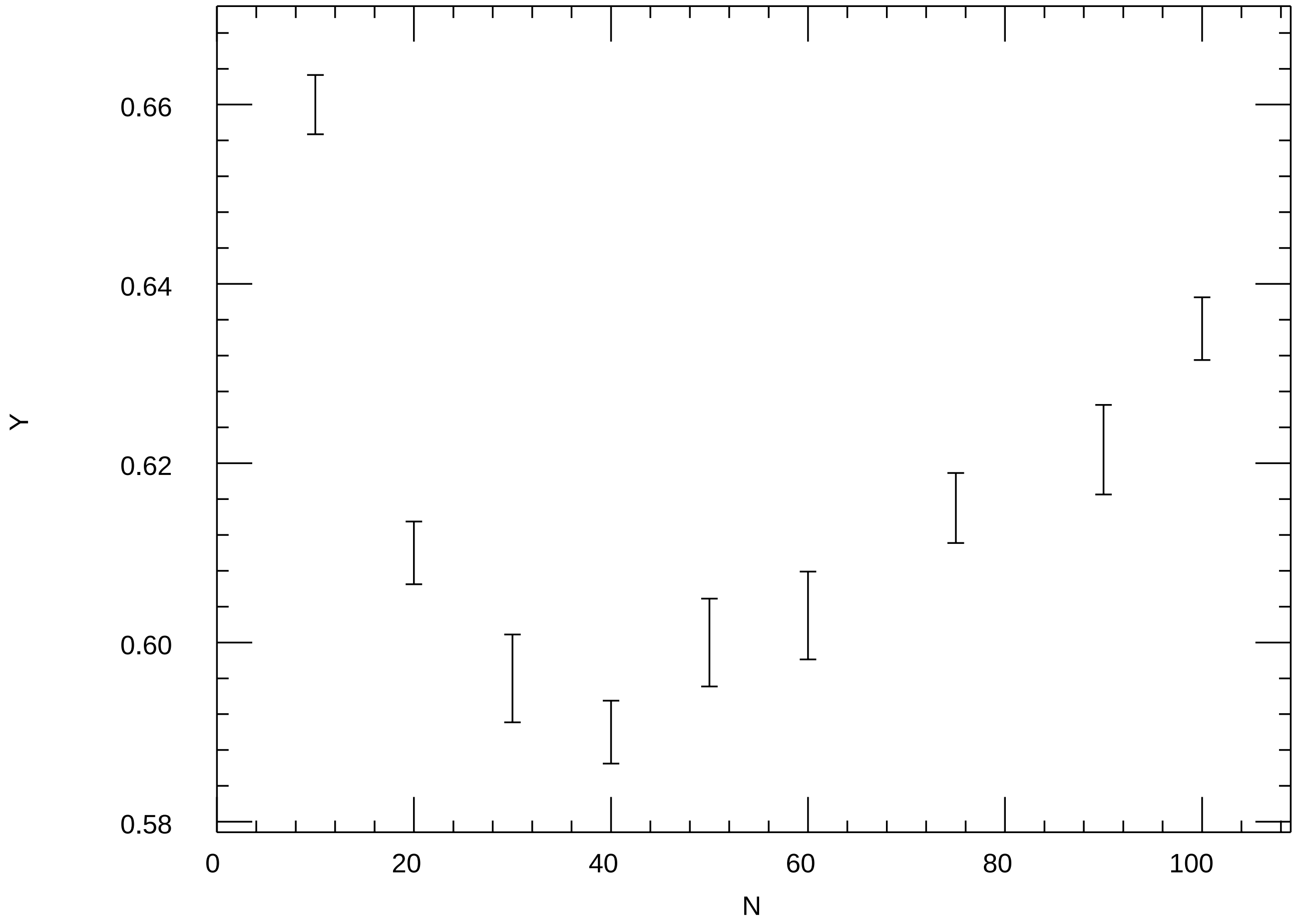

Figure 14: $\overline{Y_2}$ *versus* N *in networks with* $K = 3$*.The annealed prediction is* $\overline{Y_2} = 2/3$.

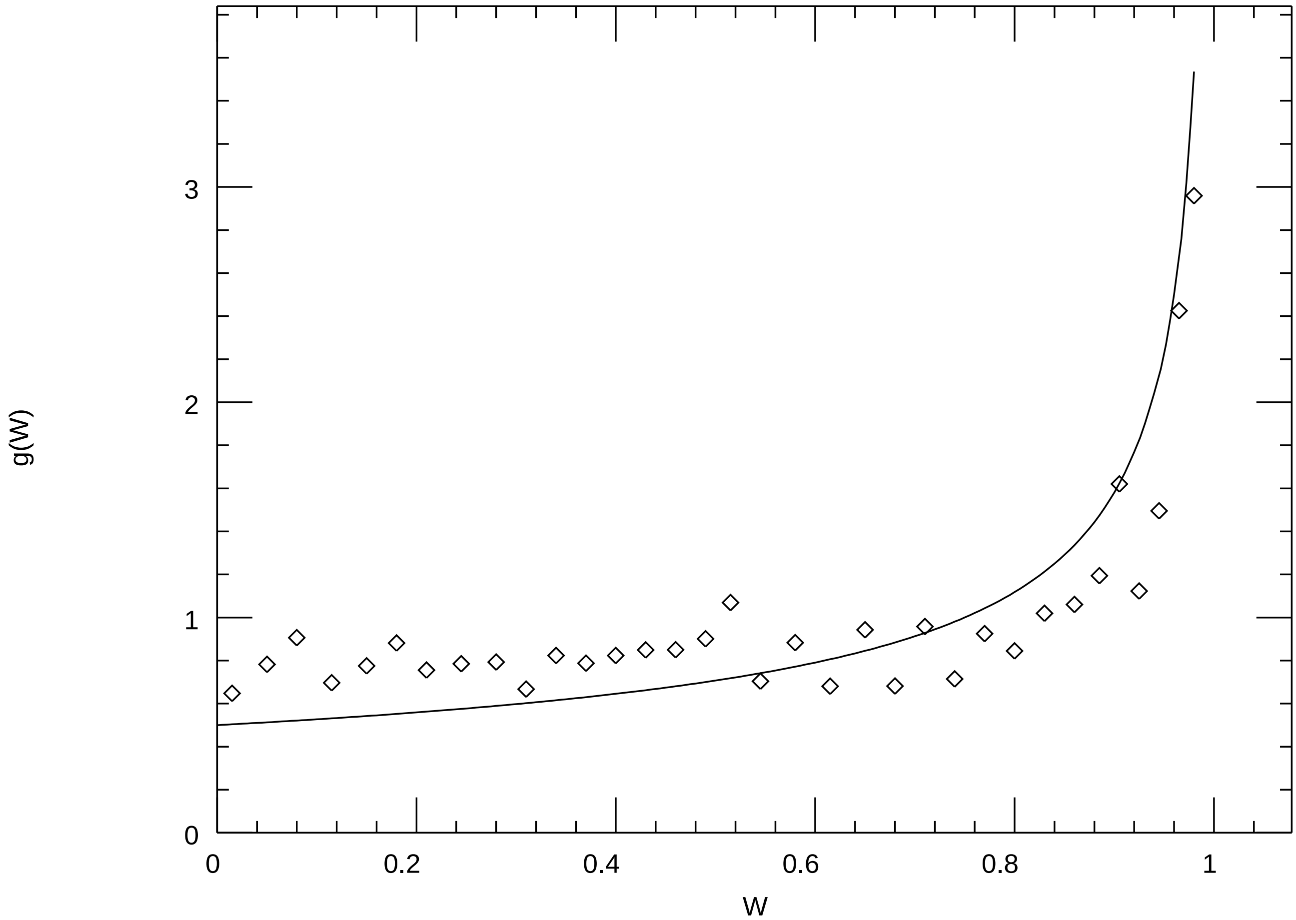

Figure 15: *Density of the probability for finding a cycle of weigth W in networks with $K = 3$ and $N = 40$, obtained by simulating 1000 networks and 250 trajectories on each of them. The dotted line is the same thing, computed in the framework of the annealed approximation.*

8 Discussion

In this work we have used a stochastic strategy, finding in the distribution of the overlaps between configurations at different time steps and, in particular, in the closing probabilities, the link between the statistical properties of our disordered dynamical system and a stochastic process. For this we had to impose the condition that the trajectory where we measure the overlap has not yet closed.

The simplest possibility is to consider a Markovian stochastic process, supposing that the system loses memory of all its past history but the previous value of the overlap. This choice leads naturally to an identification of the stochastic process that approximates the evolution of the overlap with the annealed Kauffman model, introduced in [9] by Derrida and Pomeau to approximate the evolution of the average overlap.

This very simple hypothesis leads to surprisingly good predictions: from the knowledge of the stationary distribution of the overlap in the annealed model we could easily obtain the time scale of period and transient distribution and the distribution of weights of attraction basins in very nice agreement with numerical simulations, for all parameter values in the chaotic phase. On the critical line our approximation predicts a phase transition were the scales of the distributions of cycle lengths and of the number of attractors grow as $\sqrt{N}$; numerical tests (that we will present elsewhere) seem to confirm this predictions, but it is not clear if these behaviours can be extrapolated to the limit of an infinite system.

However, there are quantities that have a completely different behaviour in the annealed approximation and in the quenched system. In particular, the distribution of cycle lengths decays much slower in the real quenched system than is expected in the annealed approximation; moreover, the decay becomes slower and slower when the number of elements increases, creating problems in the extrapolation of the numerical results to the infinite N limit. This extrapolation thus requires some more theoretical understanding.

The ambiguity of these numerical tests can be elucidated by looking directly at the overlap distribution and at the closing probabilities. Then one sees that the distribution of the overlap $q(t, t+l)$ (or, at least, its exponential part), is practically indistinguishable from the annealed prediction in a range of t growing with N[4]. But at a certain time t_N, which does not depend on l, the average overlap reaches a maximum value and starts to decrease, and the

[4]This result does not seem to have much to do with the work of Derrida and Weisbuch [10] and of Hilorst and Nijmayer [11] who proved that the annealed approximation gives the correct evolution of the average overlap for times of order $\log(N)/K$, first because here the whole distribution is involved and also because according to those considerations the smaller is K, the better is the validity of the annealed approximation, while the loss of memory that we postulated is better realized by large K values and, as we imposed the condition to measure the overlap over trajectories not yet closed before time $t+l$, the annealed approximation in our case is exact for infinite K.

overlap distribution moves leftward, contrary to our assumption of a Markovian stochastic process, with a time independent transition probability and a property of ergodicity, which should go to a stationary distribution.

We interpreted this discrepancy, which is the source of the unexpected behaviour of the period distribution, as a consequence of the opening condition, that selects trajectories where the typical overlaps are smaller and smaller.

It is not possible to impose the opening condition in the framework of the annealed approximation, where we suppose that the system loses memory of all its past evolution, except the value of the overlap $q(t-1,t'-1)$; but it is possible to find a probabilistic scheme in which the corrections due to the opening condition can be computed approximately with the minimum amount of information about the past evolution. We intend to pursue this in another paper.

If these corrections happen to improve substantially the annealed approximation, it would be interesting to justify the success more rigorously. We have not attempted to do this: our conjecture is that the validity of the annealed approximation is connected with the loss of memory of the evolution of the system. This fact is consistent with our numerical results, which show that the overlap of two configurations very far apart along the trajectory is in some sense "more annealed" than the overlap of two configurations temporally closer. How this loss of memory comes about and what are its limitations would be an interesting question to address to understand in a more analytical way dynamical systems with quenched disorder.

An alternative possibility, complementary in some sense to the annealed approximation, is to concentrate attention on the structural properties of Kauffman networks. We saw that the period fluctuates much more between different networks than in the same network: this is a first clue of the fact that the structural properties may play an important role.

The first of such properties to be studied analytically in a mean field theory was the fraction of the elements which evolve to a constant state, introduced by Flyvbjerg [12] as an order parameter. We concentrated our attention on the relevant elements, defined as the elements that do not become constant and that control some other relevant elements, and studied numerically their distribution and the connections with the dynamical properties of the network [15].

The relevant elements are divided in separate subnetworks which become dynamically independent after a transient time; a clue to the fact that this modular organization is indeed relevant in finite K Kauffman networks lies in the fact that cycles with even length are much more likely than odd ones. We investigated numerically the modular organization of Kauffman networks, finding interesting results useful for a deeper understanding of critical systems [16].

9 Acknowledgments

We are grateful to Bernard Derrida, Henrik Flyvbjerg, David Lancaster and Luca Peliti for interesting discussions and for reading the manuscript.

References

[1] S.A. Kauffman (1969), *J. Theor. Biol.* **22**, 437 Homoeostasis and Differentiation in Random Genetic Control Networks, Nature **244**, 177-178

[2] S.A. Kauffman (1984), *Physica* D **10** 145

[3] B. Derrida and H. Flyvbjerg (1986), Multivalley Structure in Kauffman's Model: Analogy with Spin Glasses, *J.Phys.A: Math.Gen.* **19**, L1003-L1008

[4] M. Mezard, G. Parisi, M.A. Virasoro (1987), *Spin Glass Theory and Beyond*, Singapore: World Scientific

[5] B. Derrida, H. Flyvbjerg (1986), The random map model: a disordered model with deterministic dynamics, *Journal de Physique* **48**, 971-978

[6] H. Flyvbjerg, N.J. Kjaer (1988), Exact Solution of Kauffman Model with Connectivity One, *J.Phys. A:Math.Gen.* **21**(**7**), 1695-1718

[7] B. Harris (1960), *Ann. Math. Stat.* **31**, 1045

[8] S.A. Kauffman (1993), *Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press

[9] Derrida B. and Y. Pomeau (1986), Random Networks of Automata: a Simple Annealed Approximation, *Biophys. Lett.* **1**(**2**), 45-49

[10] B. Derrida and G. Weisbuch (1986), *J. Physique* **47** 1297.

[11] H. J. Hilhorst and M. Nijmajer (1987), On the Approach of the Stationary State in Kauffman's Random Boolean Network, *J. Physique* **48**, 185-191

[12] H. Flyvbjerg (1988), An Order Parameter for Networks of Automata, *J.Phys. A:Math.Gen.* **21** L955-L960

[13] B. Derrida, D. Stauffer (1986): Phase Transitions in Two-dimensional Kauffman Cellular Automata, *Europhys. Lett.* 2(10), 739-745

[14] U. Bastolla and G. Parisi (1996): A Numerical Study of the Critical Line of Kauffman Networks (in preparation).

[15] U. Bastolla and G. Parisi (1996): Kauffman Model: Relevant Elements and Dynamical Properties (in preparation).

[16] U. Bastolla and G. Parisi (1996): The Modular Structure of Kauffman Networks (in preparation).

[17] A. Bhattacharijya and S. Liang (1995), Power-Law Distribution in the Kauffman Net, archived at *http://xxx.lanl.gov/archive/cond-mat/9510154*.