Coevolution in a rugged fitness landscape

Per Bak
Department of Physics, Brookhaven National Laboratory, Upton, New York 11973

Henrik Flyvbjerg and Benny Lautrup
Computational Neural Network Center, The Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark
(Received 3 June 1992)

A variant of Kauffman's *NKC* model for the coevolution of haploid organisms is shown to have two phases: a *frozen* phase in which all species eventually reach local fitness maxima and stop evolving, and a *chaotic* phase in which a fraction of all species is at local maxima, while another fraction evolves towards maxima. In doing so, they set other species back in evolution, thereby maintaining a steady fraction of evolving species. The evolutionary activity of the steady state is a natural order parameter for the ecosystem. Closed expressions are given for this order parameter and for the system's relaxation time. The latter quantity diverges at the phase boundary, showing the system is critical there. All results were obtained analytically for the maximally rugged case of $K+1=N$, and to leading order in N, the number of genes in a species.

PACS number(s): 87.10+e, 05.40+j, 05.50+q

I. INTRODUCTION

In this second of two papers, we discuss a variant of Kauffman's *NKC* model for the coevolution of haploid organisms [1–4]. It has been suggested [3] that this model self-organizes dynamically to criticality [5]. It thereby provides a model for the intermittency of extinction events observed in biological evolution by Raup [6].

We study the *NKC* model with maximally rugged fitness landscapes, and demonstrate analytically that it possesses two phases, one phase with dynamics governed by attractive fixed points, and another phase with chaotic dynamics. The phases are separated by a critical line in the *(N,C)* plane at $C \simeq N/\ln N$. Here N, K, and C parametrize the number of genes in the evolving organisms, the roughness of the fitness landscapes they evolve in, and their mutual dependence, respectively. To obtain the maximally rugged landscapes we study, one must choose $K = N-1$, so K does not occur as an independent parameter in the present paper.

The ultimate goal of our analytical investigation of the *NKC* model is to demonstrate its capacity for self-organized criticality, if it is there. The present paper reports on some progress in this direction, inasmuch as we show that the first prerequisite, critical behavior, is there in the model. We may hope then that a more realistic version of the model, having C and N as dynamical variables, may self-organize by driving these parameters to critical values. Whether this is the case is not addressed in the present paper. Here we keep N, K, and C fixed, and common to all species, during the dynamics. Also, all such necessary partial results, which pertain to the evolution of a single species in isolation, have been derived and collected in the preceding paper [7], hereafter referred to as paper I.

In Sec. II we briefly describe and motivate the version of the *NKC* model studied here. In Sec. III we write down a master equation for its time evolution, correct to leading order in an expansion in $1/N$. In Sec. IV we develop a qualitative understanding of the model's dynamics, before we describe it in quantitative detail in Sec. V. The different phases are characterized by the asymptotic dynamics of the model. The relaxation to this asymptotic behavior is described in Sec. VI. Section VII contains a summary of our findings, a discussion of the conditions under which they were found, and some ideas for future work.

II. THE SYSTEM

We already described and motivated the *NK* model in paper I. So, since the *NK* model is the *NKC* model with $C = 0$, we can be brief here.

We consider an ensemble of mutually dependent and evolving species, an *ecosystem*, so to speak. At any time, the state of any species is given by the state of its genome. This genome contains N genes. We shall assume the genes are binary variables, i.e., there are only two alleles. We do not expect our results to change in any significant way if the number of alleles is changed, as long as it is small compared with N in results based on expansion in $1/N$.

As in paper I we do not distinguish between phenotypes and genotypes, and also neglect variations in type within a species. In real life variation is responsible for the very existence of evolution. In the *NKC* model, however, only this consequence of variation is modeled: evolution takes place and is driven by a constant rate of mutations of individual, randomly chosen genes. If a mutation increases the fitness of a species, it is accepted, and the entire species is changed. If a mutation does not increase the fitness, it is rejected, and the species remains unchanged. If the time scale that selection works

on is much faster than the time scale for mutations, this lends some justification to our "all or nothing" dynamics neglecting variations [8]. Proliferation and extinction of species are both neglected in the present paper.

The fitness of any of the evolving species is a random function of its N genes *and* of C other genes belonging to other species [9]. These C other genes are chosen at random among the genes of other species. For a given sample of the kind of ecosystem described here, the particular choice for these C genes and the random fitness function define the sample, and remain fixed during evolution — the randomness is *quenched*. The choice of a random fitness function is motivated as in paper I.

(i) We conjecture that any fitness landscape with a finite correlation length will be indistinguishable from a random function on a sufficiently coarse-grained scale. So by choosing a random landscape we avoid the particular and treat a whole class, in principle.

(ii) It is technically convenient because the absence of correlations allows us to derive a number of analytical results, as shown below.

As explained in paper I, we may assume that the values of the random fitness function are equidistributed in the interval $[0, 1]$ with no loss of generality.

III. MASTER EQUATION

Because each species evolves by mutation of randomly chosen genes in a random fitness landscape, its path of evolution through configuration space can be replaced by a random walk, to leading order in N. This point was discussed in detail for isolated species in paper I. It remains valid for interacting species, because each species essentially is given a new random fitness landscape to evolve in each time a change occurs in one of the C genes in other species it depends on. This observation causes vast simplifications in the description of the system's dynamics, which, on the other hand, is exact then only to leading order in N. But that is a small price to pay, as we imagine N is large anyway.

We include two additional simplifications in the description: instead of keeping fixed the C randomly chosen foreign genes that any species depends on, we rechoose them at random any time we need them, i.e., we exchange "quenched" randomness for "annealed." If the total number of species in the ecosystem is effectively infinite — and this assumption is the second simplification we add to the description — then there is no difference between results based on quenched (annealed) randomness. This is because the set of species that any species depends on, directly or via other species, forms a C-branched tree, each node of the tree representing a species, each oriented branch a dependency. So while our exchange of quenched for annealed randomness amounts to a mean-field approximation, we nevertheless expect the mean-field theory to be exact, because the system effectively is infinite dimensional through its random connections.

The second assumption, an effectively infinite number of species in the ecosystem, makes a description in terms of density functions possible: let $\rho_M(F;t)$ denote the relative number of species which have fitness F and M less fit one-mutant neighbors at time t. A change in a random gene will then lead to higher fitness — and therefore be accepted — with probability

$$A(t) = \sum_{M=0}^{N} (1 - M/N) \int_0^1 dF \rho_M(F;t) \tag{1}$$

because $1-M/N$ is the probability that the change of one random gene leads to higher fitness in a species which has M less fit one-mutant neighbors. We note that $A(t)$ also is the rate at which mutations are accepted by the ecology from the constant rate of mutations offered. So $A(t)$ is a measure of the evolutionary activity in the ecology. We shall find it a useful quantity below, and refer to it as the *activity*.

The probability that such a mutation is accepted *and* results in fitness F for the changed species is

$$\Phi(F;t) = \int_0^F dF' \phi(F';t), \tag{2}$$

where

$$\phi(F';t) = \frac{1}{1-F'} \sum_{M=0}^{N} (1 - M/N) \rho_M(F';t) \tag{3}$$

is the contribution to this probability from species with fitness F'. This contribution does not depend on F as long as $F \geq F'$. This is so because we have assumed the fitness landscape is uncorrelated. The factor $1/(1-F')$ in this expression is the normalization factor for the constant distribution for F with $F \geq F'$.

With this notation we can easily write down the master equation for $\rho_M(F;t)$:

$$\frac{\partial}{\partial t}\rho_M(F;t) = -\left(1 - \frac{M}{N}\right)\rho_M(F;t) + B_{M,N}(F)\Phi(F;t)$$

$$-\frac{C}{N}A(t)\rho_M(F;t) + \frac{C}{N}A(t)B_{M,N}(F). \tag{4}$$

This nonlinear integro-differential equation expresses that the relative number of species with fitness F, and M less fit one-mutant neighbors, changes for four different reasons, corresponding to the four terms on the right-hand side of Eq. (4). The time scale in Eq. (4) has been chosen such that in one unit of time one mutation is offered per species — to be accepted or rejected.

The first term on the right-hand side of Eq. (4) is the rate at which species with fitness F, and M less-fit neighbors, mutate to higher fitness.

The second term on the right-hand side is a rate of change of less-fit species into species with fitness F and number of less-fit neighbors M. The function $B_{M,N}(F)$ is the binomial distribution with mean F:

$$B_{M,N}(F) = \frac{N!}{M!\,(N-M)!} F^M (1-F)^{N-M}. \tag{5}$$

It represents the probability that M out of N one-mutant neighbors to a genome with fitness F are less fit than F. This probability is binomially distributed because the fit-

ness landscape is random, with fitness F equidistributed in the interval $[0,1]$ [10].

The third term is a rate of loss of species with fitness F, M. This loss is not caused by a change in the genes of the species lost but by a change in its fitness due to genetic changes in other species. Since the C genes in other species that any species depends on are randomly chosen, this change is the product of the probability $A(t)$ that a mutation in a random species is accepted and the probability $C/N\rho_M(F;t)$ that the gene it occurs in is a gene on which a species with fitness F, M depends.

The fourth term on the right-hand side of Eq. (4) is, like the second term, a rate of change of species into species with fitness F, M. It complements the third term: species that change fitness due to genetic changes in other species can change their fitness to F with equidistributed F. When they have done that, they have M less-fit neighbors with probability $B_{M,N}(F)$.

We note that Eq. (4) conserves the total probability, as it should:

$$\frac{\partial}{\partial t}\int_0^1 dF \sum_{M=0}^{N} \rho_M(F;t) = 0. \tag{6}$$

IV. ESTIMATING THE PHASE STRUCTURE

Clearly, a static solution to Eq. (4) is provided by

$$\rho_M(F;t) = \delta_{M,N}\rho(F), \tag{7}$$

for any distribution $\rho(F)$. This solution corresponds to all species being at local fitness maxima. In the language of [1–3], borrowed from economics, the system is at a Nash equilibrium. Whether this fixed point for the dynamics is attractive or repulsive with respect to perturbations of $\rho_M(F)$ depends on the value of C. For $C = 0$ it is attractive, since in this case each species evolves in a fixed landscape, and consequently arrives at a local maximum, as described in paper I. At the other extreme, $C/N \gg 1$,

$$\rho_M(F;t) = B_{M,N}(F) \tag{8}$$

is a static solution to leading order in N/C. It corresponds to totally random fitness F, and maximum activity $A = \frac{1}{2}$.

At intermediate values of C, we can easily imagine the existence of a static solution with a finite activity A corresponding to a certain fraction of all species being in states that evolve. The activity is maintained by a balance between the rate at which species evolve towards fitness maxima and the rate at which species are set back in evolution by their dependence on other species. We expect the activity A to increase with C.

On the other hand, we can also imagine that C can be too small to sustain a finite activity. In paper I we found that isolated species on the average change

$$\mu_1 = \ln N + 0.099\,13... + O(N^{-1}) \tag{9}$$

genes in their evolution to a local maximum. So do species in the *NKC* model studied here, if they are not set back in evolution by their dependence on other species. Thus μ_1 is the minimal number of genetic changes per species by which the *NKC* model can evolve to the fixed point Eq. (7). If, in doing so, each species on the average sets back less (or more) than one other species in evolution, the fixed point Eq. (7) will (or will not) be attractive.

We can make the argument more precise by making it perturbative: suppose for a given value of C the system has been arranged to be at the fixed point solution Eq. (7), and we change the fitness of one species to a random value. Since the other species do not evolve, the one singled out evolves as an isolated species, and arrives at a fitness maximum after having changed typically μ_1 of its genes. But the fitness of other species depend on the state of genes in the species that evolved; typically C other species will each depend on one gene. If any of these C genes were among the μ_1 genes that changed, the species depending on them were set back in evolution and are now evolving, possibly setting back yet other species in their evolution. The question then is, if the chain reaction set off this way is sub- or supercritical. Will it die out or run away? The value for C which separates these two situations we call critical, and write it C_{crit}. It is the value for which, on the average, one out of C randomly chosen genes is among the μ_1 changed genes. Thus $1 = C_{\text{crit}}\mu_1/N$, or

$$C_{\text{crit}} = N/\mu_1. \tag{10}$$

We conclude that the species collectively evolve each to their own local fitness maximum and remain there with vanishing activity A for $C < C_{\text{crit}}$, while they evolve to a state with finite activity $A < \frac{1}{2}$ for $C > C_{\text{crit}}$. The asymptotic value of the activity A for $t \to \infty$ can consequently be used as an order parameter distinguishing the two phases.

The arguments used in this section were based on average values. While we would not expect fluctuations to change the qualitative picture, they might change the coefficient in a scaling law like Eq. (10). Actually they do not. The perturbative result is exact, as we see in the next section, where we also find the activity as a function of C.

V. CALCULATING THE PHASE STRUCTURE

Let us denote a stationary, or fixed point, solution to Eq. (4) by $\rho^*_M(F)$. With the notation $A^* = A[\rho^*]$, $\phi^* = \phi[\rho^*]$, $\Phi^* = \Phi[\rho^*]$, and $c = C/N$, the time-independent version of Eq. (4) can be rewritten

$$\rho^*_M(F) = \frac{N}{N - M + CA^*} B_{M,N}(F)[cA^* + \Phi^*(F)]. \tag{11}$$

Since A^* and Φ^* both depend on ρ^*, Eq. (11) is a nonlinear integral equation for $\rho^*_M(F)$. We can solve it, nevertheless, by temporarily treating A^* as a constant, to be determined by self-consistency in the end. This is done in the following way: By multiplying both sides in Eq. (11) with $(1 - M/N)/(1 - F)$ and summing over M, one finds

$$\phi^*(F) = g(F; cA^*)[cA^* + \Phi^*(F)], \tag{12}$$

where we have introduced the function [11]

$$g(F;x) = \frac{1}{1-F}\sum_{M=0}^{N-1}\frac{N-M}{N-M+Nx}B_{M,N}(F)$$
$$= N\sum_{M=0}^{N-1}\frac{B_{M,N-1}(F)}{N-M+Nx}. \tag{13}$$

For later use we also introduce

$$G(F;x) = \int_0^F dF'\, g(F';x) \tag{14}$$

and

$$\mathcal{G}(x) = \int_0^1 dF e^{G(F;x)}. \tag{15}$$

Since g and G have simple poles at $x = -1, -1+1/N, -1+2/N, \ldots, -1/N$, the function $\mathcal{G}$ has essential singularities at these points. The graph for $\mathcal{G}(x)$ is shown in Fig. 1 for the case of $N = 10$. For $x \gg O(1/N)$ or $x < -1$, $\mathcal{G}$ simplifies to

$$\mathcal{G}(x) = (1+x)\ln(1+x^{-1}) \tag{16}$$

to leading order in $1/N$. The graph for this approximation is shown as the dotted curve in Fig. 1. The approximation has a cut in the interval $[-1,0]$ where $\mathcal{G}(x)$ has N essential singularities.

Now, remembering $\phi^*(F) = \frac{d}{dF}\Phi^*(F)$, we see Eq. (12) is solved by

$$\Phi^*(F) = cA^*(e^{G(F;cA^*)} - 1). \tag{17}$$

Inserting this solution in the definition Eq. (1) of the activity, we finally arrive at a self-consistency equation for A^*, given c:

$$A^* = cA^*[-1 + \mathcal{G}(cA^*)]. \tag{18}$$

This equation is solved by $A^* = 0$ and by A^* satisfying

$$c^{-1} = -1 + \mathcal{G}(cA^*). \tag{19}$$

The last equation gives A^* as an implicit function of c. It has a real, positive solution A^* only for

$$c > c_{\text{crit}} = [-1 + \mathcal{G}(0)]^{-1} = \mu_1^{-1}, \tag{20}$$

where μ_1 is given in Eq. (9). For $cA^* \gg O(1/N)$, Eq. (19) simplifies to leading order in $1/N$ to another implicit expression for $A^*(c)$,

$$c^{-1} = -1 + (1+cA^*)\ln[1 + (cA^*)^{-1}]. \tag{21}$$

According to Eq. (19), $A^* \propto c - c_{\text{crit}}$ for $c \sim c_{\text{crit}}$, i.e., the critical exponent for the order parameter A^* is 1. At the other extreme, for $c \to \infty$, Eq. (19) gives $A^* = \frac{1}{2}$, as we expect from Sec. IV. Figure 2 shows $A^*(c)$ for $N = 10$ and 100 as fully drawn curves. The approximate expression in Eq. (21) is shown as the dotted curve.

For $c > c_{\text{crit}}$, Eq. (11) then gives

$$\rho_M^*(F) = \frac{CA^*}{N-M+CA^*}B_{M,N}(F)\exp[G(F;cA^*)], \tag{22}$$

while for $c < c_{\text{crit}}$ we have

$$\rho_M^*(F) = \delta_{M,N}F^N \exp[G(F;0)]. \tag{23}$$

So, as already seen in Sec. IV, the long-term dynamics of the coevolving species can be of two qualitatively different kinds, depending on whether the parameters C

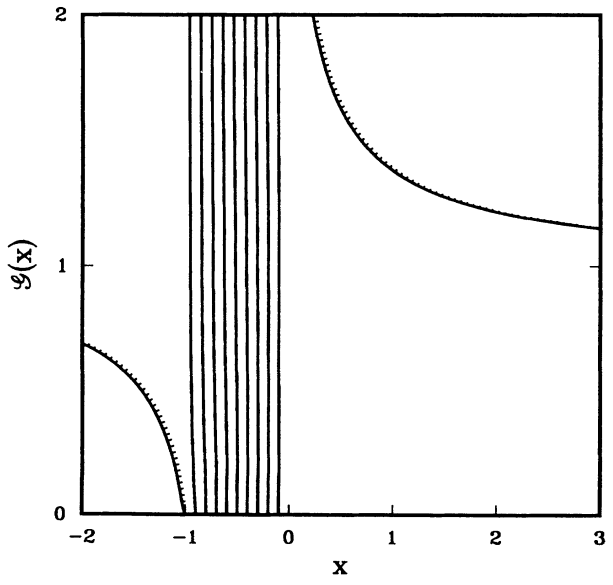

FIG. 1. Graph of the function $\mathcal{G}(x)$ defined in Eq. (15) in the case of $N = 10$ (full curve) and its approximation given in Eq. (16) (dotted curve).

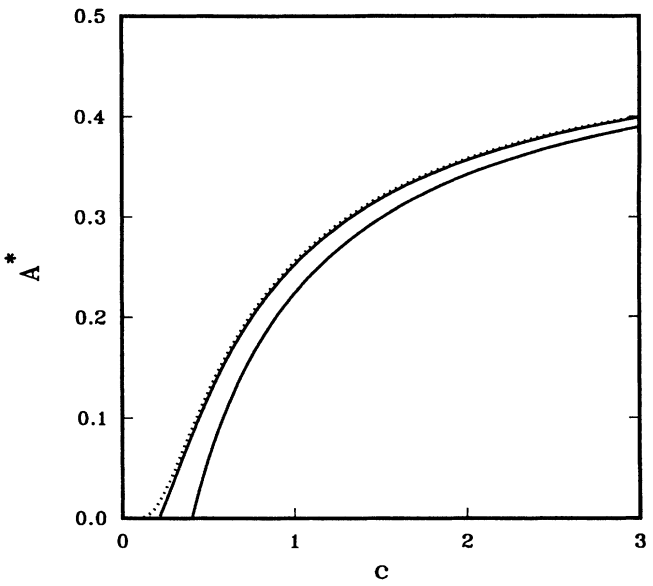

FIG. 2. The asymptotic activity A^* vs the connectivity c for $N = 10$ and 100 according to Eq. (19) (full curves) and Eq. (21) (dotted curve).

and N have values making $c = C/N$ smaller or larger than c_{crit} given above. In the first case, the activity $A(t)$ dies out because all species stop evolving as they reach local fitness maxima. This is *frozen* dynamics, characterizing the *frozen phase*. In the second case the activity converges to a nonzero value A^*, signaling *chaotic* dynamics, characterizing the chaotic phase. In this phase species also evolve towards local maxima in fitness, but in the process of doing so, they change the fitness of other species, typically setting them back in evolution. After a transient time, a balance is reached where a certain fraction of species evolve, while another fraction remains at local fitness maxima, with individual species passing from one fraction to the other every so often.

The line $C/N = c_{\text{crit}}$ dividing the (C, N) plane into two phases is critical in the sense that the relaxation time to asymptotic behavior diverges on this line, as shown in the next section.

VI. RELAXATION TIMES

In order to find the relaxation time to asymptotic values, we linearize Eq. (4) at its fixed point solution. We write

$$\rho_M(F;t) = \rho_M^*(F) + \Delta\rho_M(F;t), \tag{24}$$

$$A(t) = A^* + \Delta A(t), \tag{25}$$

$$\Phi(F;t) = \Phi^*(F) + \Delta\Phi(F;t), \tag{26}$$

$$\phi(F;t) = \phi^*(F) + \Delta\phi(F;t), \tag{27}$$

and insert these expressions in Eq. (4). By using Eq. (11) and keeping only terms linear in $\Delta\cdots$, we arrive at the linearized master equation

$$\begin{aligned}\frac{\partial}{\partial t}\Delta\rho_M(F;t) = &-\left(1 - \frac{M}{N} + cA^*\right)\Delta\rho_M(F;t)\\ &+c[B_{M,N}(F) - \rho_M^*(F)]\Delta A(t)\\ &+B_{M,N}(F)\Delta\Phi(F;t).\end{aligned} \tag{28}$$

This equation is more easily solved by writing $\Delta\rho_M(F;t)$ as a Laplace transform:

$$\Delta\rho_M(F;t) = \int_0^\infty d\lambda\, e^{-t\lambda}\Delta\tilde{\rho}_M(F;\lambda). \tag{29}$$

$\Delta A(t)$ and $\Delta\Phi(F;t)$ are linear functionals of $\Delta\rho_M(F;t)$ and therefore commute with Laplace transformation. So with a self-explanatory notation, the inverse Laplace transform of Eq. (28) reads, slightly rewritten,

$$\Delta\tilde{\rho}_M(F;\lambda) = \frac{c\Delta\tilde{A}(\lambda)[B_{M,N}(F) - \rho_M^*(F)] + B_{M,N}(F)\Delta\tilde{\Phi}(F;\lambda)}{1 - M/N + cA^* - \lambda}. \tag{30}$$

By multiplying both sides of this equation with $(1 - M/N)/(1 - F)$, and summing over M, one finds

$$\begin{aligned}\Delta\tilde{\phi}(F;\lambda) = &\, c\Delta\tilde{A}(\lambda)[g(F;cA^* - \lambda) - g_1(F;cA^* - \lambda)]\\ &+g(F;cA^* - \lambda)\Delta\tilde{\Phi}(F;\lambda),\end{aligned} \tag{31}$$

where the function $g(F;x)$ was introduced in the preceding section, and the function g_1 has a similar definition:

$$\begin{aligned}g_1(F;x) &= \frac{1}{1-F}\sum_{M=0}^{N-1}\frac{N-M}{N-M+Nx}\rho_M^*(F)\\ &= \frac{cA^*}{x - cA^*}[g(F,cA^*) - g(F,x)]\exp[G(F,cA^*)].\end{aligned} \tag{32}$$

Equation (31) is solved by

$$\begin{aligned}\Delta\tilde{\Phi}(F;\lambda) &= c\Delta\tilde{A}(\lambda)e^{G(F;cA^*-\lambda)}\\ &\quad\times\int_0^F dF'\, e^{-G(F';cA^*-\lambda)}\\ &\quad\times\ [g(F';cA^* - \lambda) - g_1(F';cA^* - \lambda)]\\ &= c\Delta\tilde{A}(\lambda)\left(-1 + \frac{cA^*}{\lambda}e^{G(F;cA^*)}\right),\end{aligned} \tag{33}$$

where we have used the definition, Eq. (32), for g_1 to obtain the last equality. Using

$$\Delta\tilde{A}(\lambda) = \int_0^1 dF\,\Delta\tilde{\Phi}(F;\lambda), \tag{34}$$

integration over F on both sides of Eq. (33) gives an equation for $\Delta\tilde{A}(\lambda)$ which is solved by $\Delta\tilde{A}(\lambda) = 0$, as we might expect, and by

$$\frac{cA^* - \lambda}{\lambda}[\mathcal{G}(cA^* - \lambda) - \mathcal{G}(cA^*)] = 0. \tag{35}$$

The smallest value for λ solving this equation contributes with the longest relaxation time

$$t_{\text{relax}}^{\text{chaotic}} = \lambda^{-1} \tag{36}$$

to $\Delta\rho_M(F;t)$ in Eq. (29). An obvious solution is

$$\lambda = cA^*. \tag{37}$$

A survey of $\mathcal{G}(x)$'s graph shows there are $N-1$ other solutions to Eq. (35), one in each interval $]cA^* + M/N, cA^* + (M+1)/N[$, where $M = 1, 2, \ldots, N-1$. So all these solutions correspond to contributions to $\Delta\rho_M(F;t)$ which decay faster in time than the mode corresponding to $\lambda = cA^*$. We conclude that the relaxation time in the chaotic phase is

$$t_{\text{relax}}^{\text{chaotic}} = \frac{1}{cA^*}, \tag{38}$$

where A^* is a function of c given implicitly by Eq. (19).

Since $A^* \sim c - c_{\text{crit}}$ for $c - c_{\text{crit}} \sim 0^+$, we see from Eq. (38) that the relaxation time diverges with exponent -1 at the critical connectivity. This typical mean-field value for the exponent comes as no surprise; it is after all a mean-field description we are developing. The value for this exponent is exact, however, in the limit $S = \infty$ of infinitely many species, which we are considering. The only requirement is that each species depends on a vanishing fraction of other species, i.e., $C/S = 0$, and that the species which a given species depends on were chosen at random. Whether this randomness is quenched or annealed does not matter. This point has been explained in

detail in [12, 13] for an in this respect identical problem.

In the frozen phase, where the order parameter $A^* = 0$, Eq. (35) shows

$$c^{-1} + 1 = \mathcal{G}(-\lambda), \tag{39}$$

which for a given value of $c < c_{\text{crit}}$ has N positive solutions for λ, one in each interval $]cA^* + M/N, cA^* + (M + 1)/N[$, where $M = 0, 1, \ldots, N-1$. The smallest solution, which determines the relaxation time, grows from $\lambda = 0$ to $1/N$ for c decreasing from c_{crit} to 0. So the relaxation time grows from N to infinity when c grows from 0 to c_{crit}. This result agrees with the average relaxation time for isolated species found in the preceding paper, and the expected increase in relaxation time with increasing coupling.

We can summarize our results for the relaxation time in the following implicit expressions for it:

$$c^{-1} + 1 = \mathcal{G}(-t_{\text{relax}}^{-1}) \quad \text{for} \quad c < c_{\text{crit}}, \tag{40}$$

$$c^{-1} + 1 = \mathcal{G}(t_{\text{relax}}^{-1}) \quad \text{for} \quad c > c_{\text{crit}}, \tag{41}$$

where the solution for t_{relax} is obtained by using the branch of $\mathcal{G}^{-1}$ characterized by $-1/N < x < \infty$.

VII. SUMMARY; DISCUSSION; PERSPECTIVES

We have shown analytically that Kauffman's *NKC* model has two phases: a *frozen* phase in which all species eventually stop evolving, because they all reach local fitness maxima, and a *chaotic* phase characterized by a balance between the number of species at local fitness maxima, and the number evolving towards such maxima, and changing the fitness landscape of other species in the process. As order parameter we used the asymptotic *activity*, the fraction of species changing genetically per unit of time. We gave a closed expression determining the asymptotic activity as an implicit function of the connectivity between species. We also gave expressions for the system's relaxation time to the asymptotic activity. On the line separating the two phases in the system's parameter space, the relaxation time diverges with mean-field exponent -1.

We obtained these results in a mean-field description of the model, keeping only leading terms in an expansion in $1/N$, N being the number of genes per species. Since N typically is very large, however, our leading-order approximation in N is very good. We do not expect any qualitative differences between our leading order $1/N$-expansion results and exact results as concerns the existence of the two phases, the location of the phase boundary, and the relaxation time. As for the exponent -1 for the divergence of the relaxation time, we have argued that it is an exact result. These results all depend on the number of species S being effectively infinite, and certainly much larger than both the number of genes N and the connectivity C.

It may well be possible to obtain other analytical results for the *NKC* model using the methods of the present paper. For example, one may try to find the Lyapunov exponents of the chaotic phase.

As for the ultimate goal of our investigation — the demonstration of self-organized criticality in the *NKC* model — we see no way that the maximally rugged variant studied here can be driven with perturbations from its frozen phase into a "poised," critical state, as was done in [14] with Conway's *Game of Life*. The maximally rugged variant cannot be "pumped up" to a "poised" state — at least not in the mean-field description — because after the model has responded to a perturbation it is back in the same state as it was before the perturbation was applied. This is not necessarily a shortcoming of the mean-field description. It willingly describes the buildup of the self-organized critical state of conservative sandpile models, for example [15]. Rather it is due to the maximal ruggedness of the fitness landscape. Its total absence of correlations makes any perturbation of a species eliminate all memory of the fitness the species had acquired before the perturbation was applied. There is, so to speak, no such thing as a perturbation *of fitness* in the maximally rugged case. Genetic configurations may be perturbed by having just one or a few genes changed. But that typically results in a finite change of fitness in a maximally rugged landscape.

On the other hand, maximal ruggedness of the model's fitness landscape is crucial for our ability to derive analytical results, and these results are important in view of the difficulty of a numerical simulation of the model. So we are reluctant to abandon it. That leaves us with another, biologically appealing possibility: we can make the model more realistic (and computationally even more difficult) by treating N and C as dynamical parameters of the individual species, add criteria for their evolutionary change, and ask if evolution drives their averages onto the critical line found in the present paper. That study has yet to be done. Methods and results that appear to make such an undertaking feasible, were presented above.

ACKNOWLEDGMENTS

We thank Bernard Derrida for discussions. The research of P.B. was supported by the Division of Materials Science, U.S. DOE, under Contract No. DE-AC02-76CH00016. H.F. and B.L. were supported by the Danish Natural Science Research Council and the Danish Technical Research Council under Contracts Nos. 11-9450-1 and 5.26-1818, respectively.

[1] S. A. Kauffman and S. Levin, J. Theor. Biol. **128**, 11 (1987).

[2] S. A. Kauffman, *Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, Oxford, England, 1990).

[3] S. A. Kauffman and S. Johnsen, J. Theor. Biol. **149**, 467 (1991).

[4] S. A. Kauffman, Sci. Am. **264** (8), 78 (1991).

[5] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1987).

[6] D. M. Raup, Science **231**, 1528 (1986).
[7] H. Flyvbjerg and B. Lautrup, preceding paper, Phys. Rev. A **46**, 6714 (1992).
[8] J. Gillespie, Evolution **38**, 1116 (1984).
[9] At this point we differ slightly from Kauffman's own definition of the *NKC*-model. He defines the fitness function of a species as the sum of N random functions, one for each gene, depending on the gene and on K other genes in the species *plus* on C genes in other species. For $K = N - 1$, a species fitness function is therefore an entirely random function of the N genes in the species, but a rather correlated function of the foreign genes it depends on. There is no good reason that the fitness function should be this way; it is just an accidental consequence of its parametrization. So for convenience we have simply assumed that the fitness function is a random function of *all* its variables.
[10] Strictly speaking, this probability for M less fit neighbors is $B_{M',N'}(F)$, with $N' < N$ and $M' = M - (N - N')$, because we already know that one or more one-mutant neighbor configurations are less fit. But we can neglect this difference in calculations to leading order in $1/N$ for reasons similar to those given in the Appendix in paper I.
[11] Contact is made with paper I by observing that $g(F;0)$ essentially is equal to the function $h_{N-1}(F;1)$ defined there. The two functions differ only because in paper I we chose to account explicitly for the impossibility of backtracking in adaptive walks, though this is an effect of subleading order in $1/N$. In the present article, formulas are simpler when we neglect subleading terms from the start. Thus the difference between $g(F;0)$ and $h_{N-1}(F;1)$.
[12] B. Derrida and G. Weisbuch, J. Phys. (Paris) **47**, 1297 (1987).
[13] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).
[14] P. Bak, K. Chen, and M. Creutz, Nature **342**, 780 (1989).
[15] P. Bak and H. Flyvbjerg (unpublished).