Eur. Phys. J. B **48**, 557–565 (2005)
DOI: 10.1140/epjb/e2006-00016-2

THE EUROPEAN PHYSICAL JOURNAL B

The random replicator model at nonzero temperature

F.C. Poderoso and J.F. Fontanari[a]

Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos, São Paulo, Brazil

Received 7 August 2005
Published online 19 January 2006 –

Abstract. The random replicator model with interspecies coupling strengths prescribed by the competitive exclusion principle – the Hebb rule – is studied analytically in the presence of fast noise that describes the flow of migrants between the ecosystem and the outer world. The stochastic dynamics leads to stationary states distributed according to the Gibbs distribution permitting thus an equilibrium statistical mechanics analysis. We find that a discontinuous phase transition separates a regime of strong competition, and consequently of low diversity, from more cooperative regimes. The statistical analysis is carried out for the annealed scheme, for which the evolutionary and ecological timescales coincide, as well as for the quenched scheme, for which the features that identify the species are fixed.

PACS. 75.10.Nr Spin-glass and other random models – 87.10.+e General theory and mathematical aspects – 87.23.Cc Population dynamics and ecological pattern formation

1 Introduction

Ecology was one of the first disciplines of biology to espouse mathematical modeling as a means to attain a comprehensive knowledge of nature (see [1] for a historical account of the mathematical thinking in population ecology). Following this tradition, in 1970 MacArthur introduced a minimum principle in ecology, analogous to the principle of least action of mechanics, by showing that a special kind of competition equation — the noiseless version of equation (2) with *symmetric* interspecies interactions — minimizes a quadratic expression, and then used this finding to interpret species packing and competitive equilibria [2]. More recently, this idea was taken up and considerably refined through the use of techniques borrowed from the statistical mechanics of disordered systems. These modern tools permitted the analytical study of very large ecosystems in which the coupling strengths between species are assigned randomly, the so-called random replicator model [3–5].

For large ecosystems, i.e. for communities composed of many different species $i = 1, \ldots, N$, an analytical statistical approach is more insightful than the direct numerical solution of the dynamical equations, since it is practically impossible to specify fully, not to mention to explore, the space of interaction strengths J_{ij}, which measure how the encounter between ith and jth species affects the growth of species i. Moreover, the uncertainty and complexity of the interspecies interactions in nature suggest that the J_{ij}s be considered as random variables. Nevertheless, to assume simply that these couplings are independent, Gaussian distributed random variables is not good enough [3] since in doing so the model fails to account for the possibility of an underlying, non-random structure of the interspecies interactions, which might explain cooperative behavior as cross-feeding and symbioses as well as the competitive behavior that results from the exploitation of the same niche by different species [6]. An easy solution to this problem is given by assuming that the pairwise species interactions are regulated by the degree of complementarity between species, i.e., by the number of features or characters that distinguish the interacting species [7]. By assigning the defining characters of each species at random, the model retains the randomness ingredient of the original Gaussian model, while accounting for an explicit, biologically motivated structure for the interspecies interactions.

Explicitly, we assume that each species is characterized by a set of p phenotypic characters, $\mu = 1, \ldots, p$ and that the resulting interaction between a pair of species, say J_{ij}, depends on the presence or not of the same character in both species, being given by the rule

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \qquad i \neq j \tag{1}$$

where the ξ_i^μs are independent random variables that take on the values ± 1 with equal probability. If species i exhibits character μ then ξ_i^μ is set to 1; otherwise it is set to -1. Note that equation (1) is the celebrated Hebb rule,

[a] e-mail: fontanari@ifsc.usp.br

extensively studied in the 1980s in the context of attractor neural networks [8]. In line with the competitive exclusion principle [6], which asserts that two species living together cannot occupy the same ecological niche, we assume that the larger the number of features shared by a pair of species the stronger the competition between them, so that $J_{ij} > 0$ corresponds to pairs of competing species whereas $J_{ij} < 0$ to pairs of cooperating species. In addition, according to equation (1), for large p most of the interspecies interaction strengths are very small, corresponding to the situation in which the species live on different resources or occupy different niches. While competitive interactions seem to be the norm in macroscopic ecosystems, cooperative interactions such as cross feeding (i.e., the ability of species to use metabolites excreted by other species) are acknowledged as a crucial mechanism to generate and maintain diversity in microbial, mainly bacterial, ecosystems [9]. As a matter of fact, the mathematical formalism of the random replicator model is very well suited to describe evolution experiments on microbial cultures carried out in chemostatic conditions, in which the total concentration of individuals is regulated to a constant level via some flux control mechanism.

Strictly, taxonomists define species through the identification of a reliable group of characters, i.e., characters that are exhibited by all members of the species to be defined but not by members of other species [10]. Here we define a species by a list of p identifiable morphological characters that a particular group of individuals may or may not possess. In particular, by assuming that the number of characters is extensive, i.e., $p = \alpha N$, we guarantee that each species is assigned a unique set of characters since the probability that two species are assigned the same set of characters vanishes as $2^{-\alpha N} N^2$ in the limit of large N. The use of a complementarity principle in the form of an overlap between binary vectors that identify the species, equation (1), is also a central feature of the Tangled Nature model [11,12]. However, analogously to the models of adaptive walks on complex fitness landscapes [13,14], the Tangled Nature model is an individual-based model that can be studied chiefly through computer simulations.

The study of the zero-temperature statistical properties of the random replicator model with interspecies interactions given by the Hebb rule has revealed a discontinuous phase transition between a cooperative and a competitive regime when the ratio α between the number of features and the number of species reaches a certain value $\alpha_c < 1/2$. As in the original Gaussian model, no similar phenomenon is observed for $\alpha > 1/2$ [15,16]. This threshold phenomenon, however, was shown to be fragile against the presence of static noise of vanishingly small intensity or, equivalently, of any non-linear distortion of the Hebb rule [16]. In this contribution we probe the robustness of the phase transition to the presence of the fast noise that results from considering the random replicator model at nonzero temperature. Rather unexpectedly, we find that this kind of noise actually promotes the phase transition, which can take place even for $\alpha > 1/2$ at nonzero temperature. A significant effect of the fast noise is the spontaneous generation of species, which is interpreted as the flow of migrants to the ecosystem under consideration.

The remainder of the paper is organized as follows. In Section 2 we present the model and discuss in detail the biological interpretation of the control parameters, giving emphasis to the interpretation of the fast noise term. In Section 3 we study the case in which the evolutionary and ecological timescales are the same. This is known as the annealed regime in statistical physics. We note that the species features change in an evolutionary timescale while the species density in an ecological one. The other extreme case, in which the species features are kept fixed, known as the quenched regime is studied in Section 4. The organization of the ecosystem for the quenched case as inferred by the analysis of the probability distribution of the species concentrations is considered in Section 5. Finally, in Section 6 we present our concluding remarks.

2 Model

We assume that the abundance of individuals of species $i = 1, \ldots, N$ in the ecosystem, described by the real-valued quantity $x_i \in [0, \infty)$, is governed by the generalized replicator equation,

$$\frac{dx_i}{dt} = x_i \left(\mathcal{F}_i - \phi\right) + \zeta_i \left(t\right) \tag{2}$$

where $\mathcal{F}_i = -\sum_j J_{ij} x_j$ can be identified with the fitness of species i and ϕ is a Lagrange multiplier that enforces the constraint

$$\sum_{i=1}^{N} x_i = N \tag{3}$$

for all t. We assume in addition that there is a reflecting barrier at $x = 0$. Here we have introduced the Gaussian white noise ζ with zero mean and correlation function

$$\langle \zeta_i \left(t\right) \zeta_j \left(t'\right) \rangle = 2T\delta_{ij} \; \delta \left(t - t'\right). \tag{4}$$

Henceforth we will refer to the parameter T that measures the noise intensity as the temperature. In the case of *symmetric* interactions and for zero temperature, the asymptotic regime of equation (2) can be fully described by looking at the maxima of the fitness functional

$$\mathcal{F}\left(\{x_i\}\right) = -\sum_{i,j} J_{ij} x_i x_j \tag{5}$$

and so it can be shown that the only stationary states are fixed points [17]. In addition, we note that in this case the Lagrange multiplier in equation (2) is interpreted as the mean fitness of the ecosystem, i.e., $\phi = \frac{1}{N}\sum_i x_i \mathcal{F}_i$. In the case of nonzero temperature, the long time regime ($t \to \infty$) of equation (2) leads to the Gibbs probability distribution

$$\mathcal{W}(\{x_i\}) = \frac{1}{Z} \; \delta\left(N - \sum_i x_i\right) \exp\left[\beta \mathcal{F}(\{x_i\})\right] \tag{6}$$

where

$$Z = \int_0^\infty \prod_i \mathrm{d}x_i \; \delta\left(N - \sum_i x_i\right) \exp\left[\beta \mathcal{F}\left(\{x_i\}\right)\right] \qquad (7)$$

is the partition function and $\beta = 1/T$ is the inverse temperature.

The dynamics of ecosystems in general and Lotka-Volterra models in particular have been studied in the presence of white, as well as colored, *multiplicative* noise (see, e.g., [18–21]) revealing a variety of intriguing phenomena such as temporal oscillations, intermittency and stochastic resonance. The noise is said multiplicative because it enters the dynamic equations through the term $x_i\zeta_i$, describing thus random variations in the reproduction factors of the species members. In this contribution we consider *additive* noise instead, the main advantage of which is to produce a stationary regime described by the Gibbs distribution, thus permitting the use of statistical mechanics tools to derive the properties of the ecosystem. In addition, the additive noise term in equation (2) has a meaningful biological interpretation, as the spontaneous generation of species can be interpreted as the random flow of migrants between some source habitat and the ecosystem under analysis. We note also that plant regrowth equations in certain plant-herbivore models do exhibit a rate of spontaneous generation that play a similar role as the additive noise ζ [22]. Strictly, the noise is not simply additive since the Lagrange multiplier ϕ depends on ζ as well. For instance, in the case of independent species, i.e., $J_{ij} = 0$ for all pair i, j we have $\phi = \frac{1}{N}\sum_i \zeta_i$.

The constraint (3) that keeps the total concentration of species in the ecosystem at a constant level does not represent any serious limitation to our model (chemostat cultures of microorganisms are usually grown under that condition), since it can be easily shown that the replicator dynamics for N species is equivalent to the Lotka-Volterra equation for $N - 1$ species [17]. Finally, to prevent the unbounded growth of any single species — a possibility in the limit $N \to \infty$ — we introduce a quadratic damping term that accounts for the self-limitation in the growth of each species. This is achieved by setting $J_{ii} = u > 0$ for all i. With this last ingredient the fitness functional (5) becomes

$$\mathcal{F}\left(\{x_i\}\right) = -\frac{1}{N}\sum_{i\neq j}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu x_i x_j - u\sum_i x_i^2. \qquad (8)$$

To rid our results of the dependence on a specific realization of the couplings between species, we must carry out an average over the probability distribution of the random variables ξ_i^μ. The way this average is performed depends on our assumptions on the timescales of the changes in the species concentrations x_i (ecological scale) and in the species features ξ_i^μ (evolutionary scale). Here we consider the two extreme situations. The first is the annealed case in which we assume that the feature values change as fast as the species concentrations. In this case we must average the partition function directly. The second is the quenched case in which the species features are kept fixed as the species concentrations change according to equation (2). Then we must average the logarithm of the partition function, which is usually done through the replica trick. This involves considering n identical copies of the system and then the limit $n \to 0$ (see Sect. 4 for details). Interestingly, intermediate situations between the annealed, which corresponds to $n = 1$, and the quenched options can also be treated within the replica framework by exploring the physical interpretation of the finite number of replicas n proposed by Coolen et al. [23]. We shall pursue this research line in a future contribution. In what follows we will consider only the annealed and quenched cases.

3 The annealed case

The free-energy in this case is given by

$$-\beta f_a = \lim_{N\to\infty} \ln \langle Z \rangle \qquad (9)$$

with Z given by equation (7). The notation $\langle \ldots \rangle$ stands for the average over the probability distribution of the quenched random variables ξ_i^μ. In addition to being of interest by its own as pointed out before, the calculation of the annealed free energy has the advantage of being free from any mathematical delicacies, such as those used in the replica approach, and resulting in saddle-point equations relatively easy to solve numerically, while capturing some singular features of the quenched model, especially at nonzero temperature. We note, in particular, that the annealed free-energy provides a rigorous lower bound to the quenched free energy. Performing the average in equation (9) yields

$$\langle Z \rangle = \int \frac{dQ d\hat{Q}}{2\pi i/N} \frac{d\hat{R}}{2\pi i/N} \; \mathrm{e}^{N\left[\hat{R} + Q\left(\hat{Q} - \beta u\right) + G_0 + \alpha G_1\right]} \qquad (10)$$

where

$$G_0 = \ln \int_0^\infty dx \exp\left(-\hat{Q}x^2 - \hat{R}x\right) \qquad (11)$$

$$G_1 = \ln \int_{-\infty}^{\infty} Dz \exp\left(-\beta Q z^2\right) + \beta Q \qquad (12)$$

and $Dz = \exp\left(-z^2/2\right) dz/\sqrt{2\pi}$ is the Gaussian measure. In the thermodynamic limit, $N \to \infty$, the integral (10) is dominated by its value at the saddle-point where the derivatives with respect to the three variables Q, $\hat{Q}$, $\hat{R}$ vanish. Only the parameter Q has a direct physical interpretation, namely, it is proportional to the probability that two randomly selected individuals belong to the same species, a measure known as Simpson's index in the ecology literature [24]. Henceforth we will refer to Q as Simpson's index, though, strictly, the correct definition of that index is Q/N. We note that Q is related to the Rényi entropy [25]

$$\mathcal{S}_\gamma = \frac{1}{1-\gamma} \ln\left[\sum_i^N \left(x_i/N\right)^2\right] \qquad (13)$$

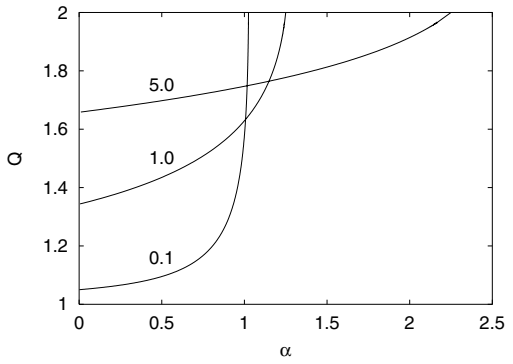

Fig. 1. Simpson's index Q in the annealed case for $u = 1$ and T as indicated in the figure. The lines end at $\alpha = \alpha_c$, at which $Q = 2$. There is no solution to the saddle-point equations beyond α_c.

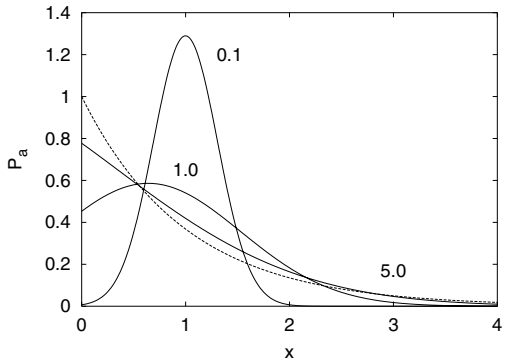

Fig. 2. Probability density that the abundance of a randomly chosen species equals x in the annealed case for $u = 1$, $\alpha = 0.5$ and T as indicated. The dashed line is the exponential distribution e^{-x} valid in the limit $T \to \infty$.

for $\gamma = 2$ (this entropy reduces to the Shannon entropy in the limit $\gamma \to 1$). The other two parameters, $\hat{R}$ and $\hat{Q}$, enter the calculations as Lagrange multipliers. The saddle-point equations are easily derived, yielding

$$Q = \left\langle x^2 \right\rangle_a \tag{14}$$

$$1 = \langle x \rangle_a \tag{15}$$

$$\hat{Q} = \beta \left(u - \alpha \frac{2\beta Q}{1 + 2\beta Q} \right) \tag{16}$$

where

$$\left\langle x^k \right\rangle_a = \mathrm{e}^{-G_0} \int_0^\infty dx\ x^k \exp\left(-\hat{Q}x^2 - \hat{R}x \right). \tag{17}$$

Figure 1 illustrates the nontrivial dependence of Q on α. For nonzero T, the parameter Q increases monotonously with increasing $\alpha \leq \alpha_c$, where

$$\alpha_c = u \left(1 + \frac{T}{4} \right). \tag{18}$$

At $\alpha = \alpha_c$ we have $Q = 2$, but we find that there is no solution to the saddle point equations for $\alpha > \alpha_c$. Of course, the absence of solutions implies that in this regime Q takes on its maximum value, $Q \to \infty$. For $T \to \infty$ we find $Q = 2$, regardless of the values of the other control parameters, while for $T = 0$ we find $Q = 1$ for $\alpha < \alpha_c = u$ and $Q = 2$ at $\alpha = u$, so there is a discontinuity at this point. From the mathematical standpoint it is not difficult to understand this behavior pattern. In fact, there are no (finite) solutions to the saddle-point equations for $\alpha > \alpha_c$ because $\hat{Q} < 0$ in this regime leading to the divergence of the moments defined in equation (17). Moreover, there is no divergence at α_c at which $\hat{Q}$ vanishes because the linear term in the argument of the exponential in that equation guarantees the convergence of the integrals.

More information on the organization of the species in the ecosystem is obtained by looking at the probability that the abundance of a randomly chosen species, say species j, is in the range $(x, x + dx)$. This is given by $P_a(x)dx$ where

$$\begin{aligned} P_a(x) &= \langle \delta(x - x_j) \rangle_a \\ &= \mathrm{e}^{-G_0} \exp\left(-\hat{Q}x^2 - \hat{R}x \right). \end{aligned} \tag{19}$$

For $T = 0$ we find $P_a = \delta(x - 1)$ and in the limit $T \to \infty$, as well as for $\alpha = \alpha_c$ regardless the value of T, we find $P_a = \mathrm{e}^{-x}$, which is consistent with $Q = 2$.

The interpretation of the saddle-point parameter Q as the Simpson's index allows us to distinguish between three distinct regimes. The first is a cooperative or high-diversity regime in which all species coexist and contribute evenly to the ecosystem composition so that $Q \approx 1$. The second is a highly disordered regime of intermediate diversity, characterized by $Q \approx 2$ in which the species interactions play little role. Finally, the third regime is a competitive or low-diversity regime in which the ecosystem is composed predominantly of a few dominant species, corresponding to $Q \gg 1$ (or $Q \to \infty$ in the annealed case). These same three regimes appear in the analysis of the quenched case, as we will see next.

4 The quenched case

The other extreme, but perhaps more realistic, situation is when the species features remain fixed and the only quantities allowed to change are the species concentrations. This is the *quenched* case which requires the calculation of the average of the logarithm of the partition function. The average free-energy density in this case is defined as

$$-\beta f_q = \lim_{N \to \infty} \frac{1}{N} \langle \ln Z \rangle \tag{20}$$

where the evaluation of the quenched average is carried out through the replica method, which consists of calculating $\langle Z^n \rangle$ for *integer* n, i.e., $Z^n = \prod_{a=1}^n Z^a$ and then using the identity

$$\langle \ln Z \rangle = \lim_{n \to 0} \frac{1}{n} \ln\ \langle Z^n \rangle, \tag{21}$$

in which it is implicit the analytical continuation to $n = 0$ [26,27]. After standard algebraic manipulations (see, e.g., [28]) we calculate $\langle Z^n \rangle$, obtaining

$$\langle Z^n \rangle = \int \prod_{a<b} \frac{dq_{ab} d\hat{q}_{ab}}{2\pi i/N} \int \prod_a \frac{dQ_a d\hat{Q}_a}{2\pi i/N} \int \prod_a \frac{d\hat{R}_a}{2\pi i/N} \times \exp\left[N \left(-\sum_{a<b} q_{ab}\hat{q}_{ab} + \sum_a \hat{R}_a + \sum_a Q_a \hat{Q}_a \right) \right] \times \exp\left[N \left(-\beta u \sum_a Q_a + G_0 + \alpha G_1 \right) \right] \quad (22)$$

where

$$G_0 = \ln \int_0^\infty \prod_a dx_a \exp\left[-\sum_a x_a \left(\hat{R}_a + \hat{Q}_a x_a \right) + \sum_{a<b} \hat{q}_{ab} x_a x_b \right] \quad (23)$$

and

$$G_1 = \ln \int_{-\infty}^\infty \prod_a Dz_a \exp\left[-\sum_a Q_a z_a^2 - 2\beta \sum_{a<b} q_{ab} z_a z_b \right] + \beta \sum_a Q_a. \quad (24)$$

The relevant physical order parameters are

$$q_{ab} = \frac{1}{N} \sum_i \langle \langle x_{ia} x_{ib} \rangle_T \rangle \qquad a < b \quad (25)$$

and

$$Q_a = \frac{1}{N} \sum_i \left\langle \langle x_{ia}^2 \rangle_T \right\rangle \quad (26)$$

which measure the overlap between a pair of stationary states labeled by the replica indices a and b, and the overlap between the stationary state labeled by a with itself. Here $\langle \ldots \rangle_T$ stands for a thermal average taken with the Gibbs distribution (6).

To proceed further we must calculate the integrals in (22) by the saddle-point method in the thermodynamic limit $N \to \infty$. The saddle-point equations are obtained by taking derivatives of the exponent of the integrand with respect to all integration variables. In order to carry out these calculations we must make some simplifying assumption about the structure of the saddle-point parameters.

4.1 Replica-symmetric solution

The simplest guess is that the saddle-point parameters are symmetric under permutations of the replica indices, i.e., $q_{ab} = q$, $\hat{q}_{ab} = \hat{q}$, $Q_a = Q$, $\hat{Q}_a = \hat{Q}$, and $\hat{R}_a = \hat{R}$. This prescription turns the evaluation of the integrals in equations (23) and (24) into a trivial task, resulting in the following replica-symmetric free energy density

$$-\beta f_{rs} = \frac{q\hat{q}}{2} - \frac{\alpha\beta q}{1 + 2\beta(Q - q)} + Q\left[\hat{Q} - \beta(u - \alpha) \right] + \hat{R} - \frac{1}{2} \ln\left(\hat{Q} + \hat{q}/2 \right) - \frac{\alpha}{2} \ln\left[1 + 2\beta(Q - q) \right] + \frac{1}{2} \ln(\pi/4) + \int Dz \ln\left[\mathrm{e}^{\Xi_z^2} \mathrm{erfc}(\Xi_z) \right] \quad (27)$$

where

$$\Xi_z = \frac{\hat{R} - \hat{q}^{1/2} z}{2\left(\hat{Q} + \hat{q}/2 \right)^{1/2}}. \quad (28)$$

In this framework the definitions (25) and (26) become

$$q = \frac{1}{N} \sum_i \left\langle \langle x_i \rangle_T^2 \right\rangle \quad (29)$$

$$Q = \frac{1}{N} \sum_i \left\langle \langle x_i^2 \rangle_T \right\rangle \quad (30)$$

where the thermal average is now calculated using the replica-symmetry prescription for the Gibbs distribution (6). The saddle-point parameters q, Q, $\hat{R}$, $\hat{q}$, $\hat{Q}$ can be obtained by solving the five coupled nonlinear equations that result from extremizing the replica symmetric free-energy with respect to each of them. It is a formidable task to solve these equations numerically. We have accomplished that by replacing the parameters $\hat{R}$ and $\hat{Q}$ by $\Delta = \hat{R}/\hat{q}^{1/2}$ and $\eta = \hat{q}^{1/2}/2\left(\hat{Q} + \hat{q}/2 \right)^{1/2}$, respectively, in order to reduce the dependence of Ξ_z from three to two parameters only: $\Xi_z = \eta(\Delta - z)$. The trick to solve the saddle-point equations is to consider η as a fixed, given parameter and α as unknown. By doing so we can easily reduce the five coupled equations to a single equation for Δ which then can be solved using standard techniques. By varying η we can find the solutions for different values of α and hence draw graphs of, say, Q against α. The success of this strategy depends on a rather technical detail, which in fact is crucial to solve the saddle-point equations in the annealed case as well. The difficulty is the numerical evaluation of the expression $\Xi_z - H(\Xi_z)$, where

$$H(\Xi_z) = \frac{1}{\sqrt{\pi}} \frac{\mathrm{e}^{-\Xi_z^2}}{\mathrm{erfc}(\Xi_z)}, \quad (31)$$

which appears when we take derivatives of the free energy (27) with respect to Δ and η. To minimize numerical errors in the case $\Xi_z > 0$ we use a routine based on Chebyshev fitting that returns the complementary error function `erfc` with fractional error everywhere less than 1.2×10^{-7} [29]. For $\Xi_z > 6$, however, we must resort to the well-known asymptotic expansion of the `erfc` [30], since that routine is not sufficient to evaluate $\Xi_z - H(\Xi_z)$ within a reasonable precision requirement.

Figures 3 and 4 show the dependence on α of the reciprocal of Q and of the susceptibility $v \equiv Q - q > 0$

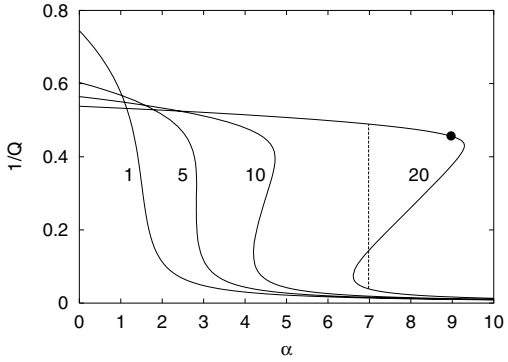

Fig. 3. Reciprocal of Simpson's index as function of α for $u = 1$ and T as indicated in the figure. The dashed vertical line indicates the point at which the thermodynamic phase transition takes place. This phase transition disappears for $T < 4.7$. The symbol • indicates the point where the replica-symmetric solution becomes unstable.

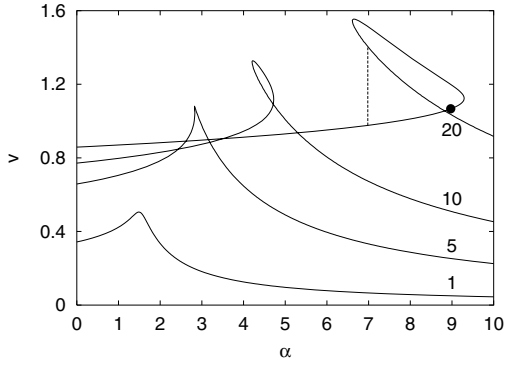

Fig. 4. Susceptibility v as function of α. The parameters and conventions are the same as in Figure 3.

which, according to equations (29) and (30), measures the fluctuations of the species concentrations around the equilibrium value. We note that at zero temperature the three solutions observed for certain ranges of α, that signal the existence of a phase transition, appear for $u < 0.5$ only [15,16] and disappear altogether when a vanishingly small static noise is added to the Hebb interactions [16]. These figures demonstrate that the effect of fast noise is the reverse – increasing the temperature actually enlarges the many-solutions region. To decide which solution must be chosen for a given value of α we must consider the free energy of all them as illustrated in Figure 5 for a single value of T. Note that only by inspection of this figure we can distinguish clearly between the three solutions. The point $\alpha_c\,(T, u)$ at which the free energies of two of the solutions intersects determines the location of the thermodynamic phase transition. The requirement that the free energy changes continuously as one moves from one solution to another results in a discontinuity in the saddle-point parameters, as indicated by the vertical dashed lines shown in Figures 3 and 4. For $u > 0.5$, the size of the many-solution region shrinks as T decreases and vanishes at a critical end point $T_e > 0$, while for $u < 0.5$, as already said, this region persists at $T = 0$. These results are

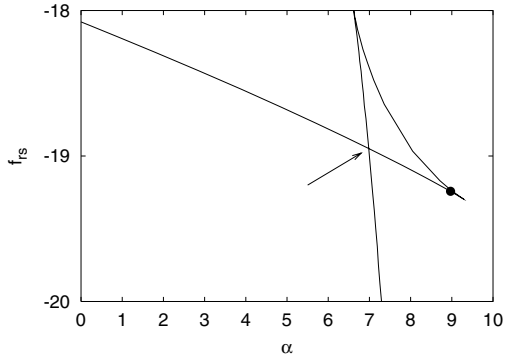

Fig. 5. Replica-symmetric free energy f_{rs} as function of α for $u = 1$ and $T = 20$. The arrow indicates the point at which the thermodynamic transition takes place, i.e., where the free energy of two distinct solutions of the saddle-point equations intersect. This point correspond to the dashed vertical lines of the previous figures. The symbol • points the limit of stability of the replica symmetric solution.

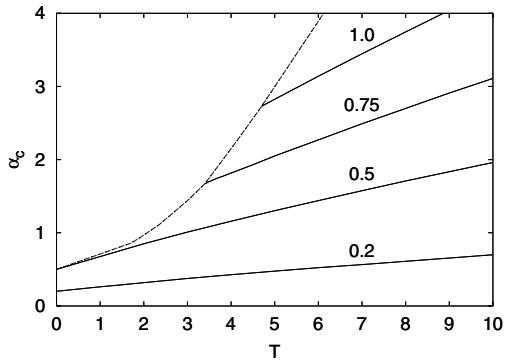

Fig. 6. Critical value α_c at which the discontinuous thermodynamic phase transition takes places as function of T for the values of u indicated in the figure. The transition lines terminate at critical end points T_e, represented by the dashed line, at which the discontinuity disappears.

summarized in Figure 6 where the transition point α_c is shown against the temperature T.

4.2 Stability analysis

In using the replica symmetric prescription to evaluate the saddle-point parameters it is important to check that the solution is in fact locally stable. An instability of the replica-symmetric solution is determined by a sign change in (at least) one of the eigenvalues of the matrix of quadratic fluctuations around the replica-symmetric solution. Following the standard stability analysis [31] it can be shown that the stability is determined by the eigenvalues of the matrix

$$\begin{pmatrix} \partial^2 G_0 & -1 \\ -1 & \alpha \partial^2 G_1 \end{pmatrix} \tag{32}$$

where $\partial^2 G_0$ is the $\frac{1}{2}n\,(n+3)$-dimensional matrix of second derivatives of $G_0(\hat{q}_{ab}, \hat{Q}_a, \hat{R}_a)$ with respect to its arguments. Similarly, $\partial^2 G_1$ is the $\frac{1}{2}n\,(n+1)$-dimensional matrix of second derivatives of G_1 with respect to q_{ab} and Q_a.

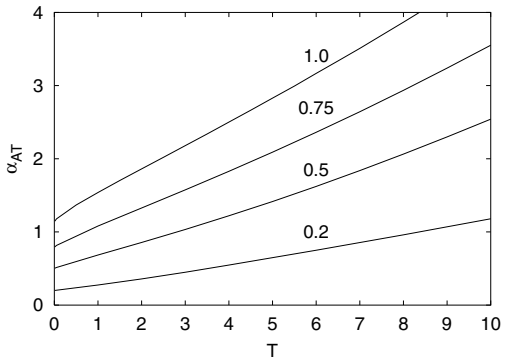

Fig. 7. Almeida-Thouless point α_{AT} above which the replica-symmetric solution becomes locally unstable as function of T for the values of u indicated in the figure.

Requiring all the eigenvalues of this matrix to be positive leads to the following condition (see [28,32] for a similar calculation)

$$\frac{4\eta^4}{q\hat{q}} \int_{-\infty}^{\infty} Dz \left\{1 + 2H\left(\Xi_z\right)\left[\Xi_z - H\left(\Xi_z\right)\right]\right\}^2 < 1 \qquad (33)$$

where $H\left(\Xi_z\right)$ is given by equation (31). Figure 7 shows the Almeida-Thouless line $\alpha_{AT}\left(T, u\right)$ for fixed values of u. These results show that only the solution corresponding to the larger value of $1/Q$ (upper branch in Fig. 3) is stable against replica symmetry breaking and that the instability sets in after the discontinuous transition takes place. The instability points are indicated in Figures 3 and 4 by the black circles. Hence, the phase characterized by small $1/Q$ is everywhere unstable and so a more sophisticated prescription, involving the breaking of the symmetry of permutation in the replica space, is necessary to describe correctly the saddle-point parameters in this phase [26,27]. Nevertheless, about three decades of research on mean-field versions of a variety of disordered systems has shown that the replica-symmetric prescription yields *qualitatively* correct results and in what follows we will use it as an approximation to the correct, replica symmetry broken solution.

5 Ecosystem structure

A common measure used to gain information on the structure of ecosystems is the relative abundance of each species in the community. In particular, field ecologists had long observed that most species of plants in a secluded community were relatively rare, while a few species were fairly common [33], resulting in patterns that can be well described by a geometric distribution [34] (see also [35]). Similarly to the annealed case, see equation (19), the probability density that the concentration of a given species – say x_k – takes on the value x is defined by

$$P_k\left(x\right) = \left\langle \int_0^{\infty} \prod_j dx_j \; \delta\left(x - x_k\right) \mathcal{W}\left(\{x_i\}\right) \right\rangle \qquad (34)$$

where $\mathcal{W}$ is the Gibbs distribution (6). Since all species are equivalent we can drop the species index and write $P_k\left(x\right) = P\left(x\right) \forall k$. Evaluation of this quantity using the replica symmetric prescription is straightforward and yields

$$P_{rs}\left(x\right) = \frac{\hat{q}^{1/2}}{\pi^{1/2}\eta} \int_{-\infty}^{\infty} Dz \frac{\exp\left[-\left(\Xi_z + x\hat{q}^{1/2}/2\eta\right)^2\right]}{\operatorname{erfc}\left(\Xi_z\right)} \qquad (35)$$

with $\Xi_z = \eta\left(\Delta - z\right)$. In the zero temperature limit this distribution exhibits a singularity (a Dirac delta) at the origin $x = 0$, implying thus that a fraction of the species are extinct in the stationary regime [15,16]. Such a singularity does not occur for nonzero temperatures – the probability density (35) is everywhere a well behaved function of the species concentrations. In fact, Figure 8 shows this probability distribution for the same parameters used in the annealed case (see Fig. 2). The results are qualitatively very similar and become practically identical for large T, as expected. This similarity supports our claim that the replica symmetric solution is useful to draw a good qualitative picture of the equilibrium properties of the model, even beyond its stability domain. The main difference between the annealed and quenched results is a greater contribution to the community organization from species with small concentration values in the quenched case. For large α, however, the annealed scheme produces divergent results and in figure 9 we show the results of the distribution of species concentrations as predicted by the replica symmetric prescription in this regime. The remarkable feature of these distributions for large α is the rapid decay at small concentrations (much faster than exponential), the leveling off at intermediate concentration values and, finally, the resume of a rapid decay at very large concentrations. This means that a few species are uncommonly abundant and practically dominate the entire community. Of course, this conclusion accords with the large value of Q in this regime which implies a high probability of finding that two individuals picked at random belong to the same species.

The increase of the number of features of each species results in a reduction of the effective diversity, in the sense that only a few species are statistically relevant to the ecosystem composition. This can be understood by noting that the variance of the interspecies interaction strengths (1) grows linearly with α, which means that the larger α, the greater the odds of producing pairs of strongly cooperating species such that the strength of the interactions among them exceed the self-restraint parameter u as well as the disturbance of the fast noise. The diversity reduction observed in Figure 3 for large α is then a consequence of the unrestricted growth of small groups of cooperating species, the success of which drive the other species to a marginal position in the ecosystem organization since constraint (3) must be fulfilled.

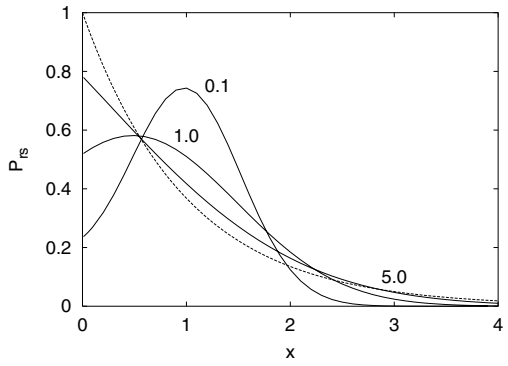

Fig. 8. Probability density that the abundance of a randomly chosen species equals x in the replica-symmetric approximation for $u = 1$, $\alpha = 0.5$ and T as indicated. The dashed line is the exponential distribution.

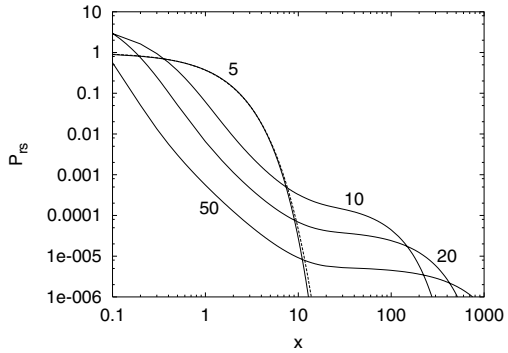

Fig. 9. Probability density that the abundance of a randomly chosen species equals x in the replica-symmetric approximation for $u = 1$, $T = 20$ and α as indicated. The dashed line is the exponential distribution.

6 Conclusion

The equilibrium properties of the model ecosystem addressed in this contribution are regulated by three control parameters only. The ratio α between the number of features p used to specify the species and the total number of species N is the more important since it can be identified with the overall complexity of the species. Large values of α promote the dominance of a few species leading to a regime of low diversity characterized by a very large value of Simpson's index, $Q \gg 1$. The self-restraint parameter $u > 0$ that models the competition between members of a same species comes in second place. It can be viewed as a cooperation pressure since for large u we find $Q \approx 1$ implying thus that all species contribute evenly to the ecosystem composition. A nonzero value of the intraspecies competition parameter is crucial to obtain a well-defined thermodynamic limit – setting $u = 0$ destabilizes the ecosystem yielding $Q \to \infty$. The third control parameter is the temperature T that measures the intensity of the noise that disturbs the otherwise uphill evolution of the species concentrations towards a local maximum of the fitness functional. For large T we find $Q \approx 2$ but the probability distribution of the species concentrations differs markedly from the distribution characteristic of the large u regime: it is an exponential distribution for large T while it approaches a Gaussian of mean 1 for large u (see Figs. 2 and 8). These three regimes are present in both the annealed and quenched cases, except that in the former the competitive regime is characterized by $Q \to \infty$. We stress that in the context of replicator models the annealed case is not a mere approximation to the quenched case as it describes ecosystems for which the evolutionary and ecological timescales coincide.

In the quenched case, we find a discontinuous phase transition separating the low-diversity competitive regime, which is everywhere unstable against replica symmetry breaking, from the other two cooperative regimes, for which the replica symmetric prescription yields the correct description. At zero temperature this threshold phenomenon takes place at $\alpha = u$, provided that $u < 1/2$. In contrast to the effect of static noise, which smoothes out that transition even for vanishingly small intensities [16], the fast noise considered here actually promotes the transition which can take place for arbitrarily large u, provided that T is large too (see Fig. 6). In the annealed case, we find a similar situation with the transition point given by equation (18) for all u.

Similarly to what was done in the neural networks context [36], the ecosystem model can also be studied analytically for a generalized form of the interspecies interactions, namely, $J'_{ij} = F(J_{ij})$ where F is an *odd*, nonlinear functions of the Hebbian terms. In doing so the effects of clipping and deterministic, as well as random, dilution of the interactions can readily be considered (see [16] for the zero-temperature analysis) and, as already pointed out, the overall effect of the nonlinearity is to destroy the thermodynamic phase transitions altogether. However, while such a generalized framework is appropriate to study the case where the interaction strengths with intensity smaller than a given threshold value are set to zero (this is an example of deterministic dilution) it cannot be used to study an asymmetric case in which, for instance, the interactions are set to zero if they are cooperative ($J_{ij} < 0$) since F must be an odd function of its argument. A glance at equation (5), however, is sufficient to realize that the model with competitive interactions only leads to an uninteresting situation in which a single species dominates the entire ecosystem. Hence this result corroborates the notion that cooperation is necessary to produce diversity in ecosystems.

The physical interpretation of the noise term in the generalized replicator equation (2) is of a migration flow that randomly introduces and removes members of the species, modeling thus the exchange of individuals between the ecosystem and its bounding environment. Actually, the effect of noise is more complicated than that since the addition and removal of individuals must be such that the total species concentration is unaltered. The main point, of course, is that this kind of fast noise is the one that permits the statistical mechanics analysis of the equilibrium states of the ecosystem (see [4] for a similar study of the Gaussian model). In the quenched case, this analysis has revealed a rich phase diagram exhibiting discontinuous

transition lines that end at critical end points. In addition, the instability of the low-diversity phase to replica symmetry breaking shown in Figure 7 turns the continuous spin model (5) into an exemplary model to study discontinuous transitions between replica symmetric and replica symmetry broken phases.

The work of J.F.F was supported in part by CNPq and FAPESP, Project No.04/06156-3. F.C.P. was supported by FAPESP.

References

1. S.E. Kingsland, *Modeling Nature* (University of Chicago Press, Chicago, 1985)
2. R. MacArthur, Theor. Pop. Biol. **1**, 1 (1970)
3. S. Diederich, M. Opper, Phys. Rev. A **39**, 4333 (1989)
4. P. Biscari, G. Parisi, J. Phys. A **28**, 4697 (1995)
5. V.M. de Oliveira, J.F. Fontanari, Phys. Rev. Lett. **85**, 4984 (2000); V.M. de Oliveira, J.F. Fontanari, Phys. Rev. E **64**, 051911 (2001)
6. G.F. Gause, *The Struggle for Existence* (Dover, New York, 1971)
7. V.M. de Oliveira, J.F. Fontanari, Phys. Rev. Lett. **89**, 148101 (2002)
8. D.J. Amit, *Modeling brain function: the world of attractor neural networks* (Cambridge University Press, Cambridge, 1989)
9. P.B. Rainey, A. Buckling, R. Kassen, M. Travisano, Trends Ecol. Evol. **15**, 243 (2000).
10. M. Ridley, *Evolution* (Blackwell Science, Cambridge, 1996)
11. K. Christensen, S. A. di Collobiano, M. Hall, H.J. Jensen, J. Theor. Biol. **216**, 73 (2002)
12. M. Hall, K. Christensen, S.A. di Collobiano, H.J. Jensen, Phys. Rev. E **66**, 011904 (2002)
13. S.A. Kauffman, S. Levin, J. Theor. Biol. **128**, 11 (1987)
14. C. Amitrano, L. Peliti, M. Saber, J. Mol. Evol. **29** 513 (1989)
15. V.M. de Oliveira, Eur. Phys. J. B **31**, 259 (2003)
16. D.O.C. Santos, J.F. Fontanari, Phys. Rev. E **70**, 061914 (2004)
17. J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998)
18. J.M.G. Vilar, R.V. Solé, Phys. Rev. Lett. **80**, 4099 (1998)
19. M.F. Dimentberg, Phys. Rev. E **65**, 036204 (2002)
20. B. Spagnolo, A. La Barbera, Physica A **315**, 114 (2002)
21. R. Mankin, A. Sauga, A. Ainsaar, A. Haljas, K. Paunel, Phys. Rev. E **69**, 061106 (2004)
22. P. Turchin, *Complex Population Dynamics* (Princeton University Press, Princeton, 2003)
23. A.C.C. Coolen, R.W. Penney, D. Sherrington, Phys. Rev. B **48**, 16116 (1993)
24. E.H. Simpson, Nature (London) **163**, 688 (1949).
25. A. Rényi, *Proc. Fourth Berkeley Symp. Math. Stat. and Probability*, (University of California Press, Berkeley, 1961), Vol. 1, pp. 547-561
26. M. Mézard, G. Parisi, M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987)
27. K.H. Fischer, J.A. Hertz, *Spin glasses* (Cambridge University Press, Cambridge, 1991)
28. E. Gardner, B. Derrida, J. Phys. A **21**, 271 (1988)
29. W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1988)
30. M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions* (Dover Publication, New York, 1972)
31. J.R.L. de Almeida, D.J. Thouless, J. Phys. A **11**, 983 (1978)
32. D.J. Amit, H. Gutfreund, H. Sompolinsky, Ann. Phys. N.Y. **173**, 30 (1987).
33. N.G. Hairston, Ecology **40**, 404 (1959)
34. G.U. Yule, Proc. R. Soc. London B **213**, 21 (1925)
35. J. Chu, C. Adami, Proc. Natl. Acad. Sci. USA **96**, 15017 (1999)
36. H. Sompolinsky, in *Heidelberg Colloquium on Glassy Dynamics*, edited by J.L. van Hemmen, I. Morgenstern, Lecture Notes in Physics (Springer, Berlin, 1987), Vol. 275, pp. 485–527