Statistical Mechanics of Systems with Heterogeneous Agents: Minority Games

Damien Challet,[1] Matteo Marsili,[2] and Riccardo Zecchina[3]
[1]*Institut de Physique Théorique, Université de Fribourg, CH-1700 Fribourg, Switzerland*
[2]*Istituto Nazionale per la Fisica della Materia (INFM), Trieste-SISSA Unit, Via Beirut 2-4, I-34014 Trieste, Italy*
[3]*The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, P.O. Box 586, I-34100 Trieste, Italy*
(Received 27 April 1999; revised manuscript received 30 June 1999)

We study analytically a simple game theoretical model of heterogeneous interacting agents. We show that the stationary state of the system is described by the ground state of a disordered spin model which is exactly solvable within the simple replica symmetric ansatz. Such a stationary state differs from the Nash equilibrium where each agent maximizes her own utility. The latter turns out to be characterized by a replica symmetry broken structure. Numerical results fully agree with our analytical findings.

PACS numbers: 02.50.Le, 05.20.Dd, 64.60.Ak, 87.23.Ge

Statistical mechanics of disordered systems provides analytical and numerical tools for the description of complex systems, which have found applications in many interdisciplinary areas [1]. When the precise realization of the interactions in a heterogeneous system is expected not to be crucial for the overall macroscopic behavior, then the system itself can be modeled as having random interactions drawn from an appropriate distribution. Such an approach appears to be very promising also for the study of systems with many heterogeneous agents, such as markets, which have recently attracted much interest in the statistical physics community [2,3]. Indeed it provides a workable alternative to the so-called *representative agent* approach of microeconomic theory, where, assuming that agents are identical, one is lead to a theory with one single (representative) agent [4].

In this Letter, we present analytical results for a simple model of heterogeneous interacting agents, the so-called minority game (MG) [3,5], which is a toy model of N agents interacting through a global quantity representing a market mechanism. Agents aim at anticipating market movements by following a simple adaptive dynamics inspired at Arthur's *inductive reasoning* [6]. This is based on simple *speculative* strategies that take advantage of the available public information concerning the recent market history, which can take the form of one of P patterns. Numerical studies [3,7–9] have shown that the model displays a remarkably rich behavior. The relevant control parameter [3,7] turns out to be the ratio $\alpha = P/N$ between the "complexity" of information P and the number N of agents, and the model undergoes a phase transition with symmetry breaking [8] independently of the origin of information [9].

We shall limit the discussion on the interpretation of the model—which is discussed at some length in Refs. [3,7]—to a minimum and rather focus on its mathematical structure and to the analysis of its statistical properties for $N \gg 1$. Our main aim is indeed to show that the model can be analyzed within the framework of statistical mechanics of a disordered system [1].

We find that dynamical steady states can be mapped onto the ground state properties of a model very similar to that proposed in Ref. [10] in the context of optimal dynamics for attractor neural networks. There [10] one shows that the minimization of the interference noise is equivalent to maximizing the dynamical stability of each device composing the system. Conversely, we show that the individual utility maximization in interacting agent systems is equivalent to the minimization of a global function. We also find that different learning models lead to different patterns of replica symmetry breaking.

The model is defined as follows [8]: Agents live in a world which can be in one of P *states.* These are labeled by an integer $\mu = 1, \ldots, P$ which encodes all the information available to agents. For the moment being, we follow Ref. [9] and assume that this information concerns some external system so that μ is drawn from a uniform distribution $\varrho^\mu = 1/P$ in $\{1, \ldots, P\}$. Each agent $i = 1, \ldots, N$ can choose between one of two *strategies*, labeled by a spin variable $s_i \in \{\pm 1\}$, which prescribes an *action* $a^\mu_{s_i,i}$ for each state μ. Strategies may be "look up tables," behavioral rules [3,6], or information processing devices. The actions $a^\mu_{s,i}$ are drawn from a bimodal distribution $P(a^\mu_{s,i} = \pm 1) = 1/2$ for all i, s, and μ, and they will play the role of quenched disorder [1]. Hence, there are only two possible actions, such as "do something" ($a^\mu_{s,i} = 1$) or "do the opposite" ($a^\mu_{s,i} = -1$). It is convenient [8] to make the dependence on s explicit in $a^\mu_{s,i}$, introducing ω^μ_i and ξ^μ_i so that $a^\mu_{s,i} = \omega^\mu_i + s\xi^\mu_i$ [11]. If agent i chooses strategy s_i and her opponents choose strategies $s_{-i} \equiv \{s_j, j \neq i\}$, in state μ, she receives a payoff,

$$u^\mu_i(s_i, s_{-i}) = -a^\mu_{s_i,i} G(A^\mu) , \tag{1}$$

where, defining $\Omega^\mu = \sum_j \omega^\mu_j$,

$$A^\mu = \sum_j a^\mu_{s_j,j} = \Omega^\mu + \sum_j \xi^\mu_j s_j . \tag{2}$$

The function $G(x)$, which describes the market mechanism, is such that $xG(x) > 0$ for all x so that the total payoff to agents is always negative: the majority of agents

 0031-9007/00/84(8)/1824(4)$15.00 © 2000 The American Physical Society

receives a negative payoff whereas only the minority of them gain. Note that the agent-agent interaction, which comes from the aggregate quantity $G(A^\mu)$, is of mean-field character.

The game defined by the payoffs in Eq. (1) can be analyzed along the lines of game theory [12] by looking for its Nash equilibria in the strategies space $\{s_j, j = 1, \ldots, N\}$. Before doing this, we prefer to discuss the dynamics of *inductive agents* following Refs. [3,7,8]: There, the game is repeated many times and agents try to estimate empirically which of the two strategies they have is the best one, using past observations. More precisely, each agent i assigns a *score* $U_{s,i}(t)$ to her sth strategy at time t, and we assume, as in Ref. [13], that she chooses that strategy with probability [14]

$$\pi_{s,i}(t) \equiv \text{Prob}\{s_i(t) = s\} = Ce^{\Gamma U_{s,i}(t)}, \quad (3)$$

with $C^{-1} = \sum_{s'} e^{\Gamma U_{s',i}(t)}$ and $\Gamma > 0$. The scores are initially set to $U_{s,i}(0) = 0$, and they are updated as

$$U_{s,i}(t+1) = U_{s,i}(t) - a_{s,i}^{\mu(t)} G(A^{\mu(t)})/P. \quad (4)$$

The idea is that if a strategy s has predicted the right sign, i.e., if $a_{s,i}^\mu = -\text{sgn}G(A^\mu)$, its score, and, hence, its probability of being used, increases. Note that $a_{s,i}^\mu G(A^\mu)$ in Eq. (4) is *not* the payoff $u_i^\mu(s, s_{-i})$ which agent i would have received if she had actually played strategy $s \neq s_i(t)$. Indeed $G(A^\mu)$ depends on the strategy $s_i(t)$ that agent i has actually played through A^μ. Agents in the MG neglect this effect and behave as if they were facing an external process $G(A^\mu)$ rather than playing against other $N - 1$ agents. This may seem reasonable for $N \gg 1$ since the relative dependence of aggregate quantities on each agent's choice is expected to be small. We shall see below [see Eq. (11)] that this is not true: If agents consider the impact of their actions on A^μ, the collective behavior changes considerably.

We focus on the linear case $G(x) = x$, which allows for a simple treatment. Other choices, such as the original one, $G(x) = \text{sgn}x$, lead to similar conclusions, as will be discussed elsewhere [15]. With this choice, the total losses of agents is $-\sum_i u_i^\mu = (A^\mu)^2$. The time average σ^2 of $(A^\mu)^2$ is shown in Fig. 1, as a function of $\alpha \equiv P/N$. The system shows a complex behavior characterized, among other things, by a phase transition at $\alpha_c \simeq 0.34$ [8], where σ^2 shows a cusp and a small α phase where σ^2 increases with Γ [13].

In order to uncover this behavior, let us focus on the long time behavior of the dynamics. The key observation is that, in the long run, the score of a strategy depends on its performance in all P states. Hence, the behavior of agents will change systematically only on time scales of order P. This suggests introduction of the rescaled time $\tau = t/P$. As $P \to \infty$, any finite interval $d\tau = \Delta t/P$ is made of infinitely many time steps, and we can use the law of large numbers to approximate time averages with statistical averages over the variables $\mu(t)$ and $s_i(t)$ from their respective distributions ϱ^μ and $\pi_{s,i}$. We henceforth

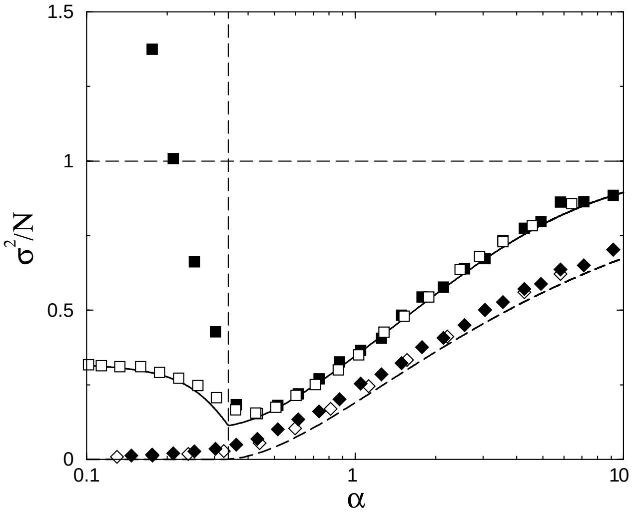

FIG. 1. σ^2/N versus $\alpha = P/N$ for $P = 2^6$ for inductive dynamics (full squares), for the numerical minimization of Eq. (7) (open squares), corrected inductive dynamics (full diamonds), and the ground state of σ^2 (open diamonds). The full and the dashed lines are the corresponding analytic results. Averages are taken over 200 realizations.

use the notation $\overline{o} = \sum_\mu \varrho^\mu o^\mu$ for averages over μ and $\langle\cdot\rangle$ for averages on $s_i(t)$, and we define $m_i(\tau) \equiv \langle s_i(t)\rangle$. With this notation, σ^2 reads

$$\sigma^2 = \overline{\langle A^2\rangle} = \overline{\Omega^2} + \sum_i [\overline{\xi_i^2} + 2\overline{\Omega\,\xi_i} m_i] + \sum_{i\neq j} \overline{\xi_i \xi_j} m_i m_j, \quad (5)$$

where we have used statistical independence of s_i, i.e., $\langle s_i s_j\rangle = m_i m_j + (1 - m_i^2)\delta_{i,j}$. The evolution of scores $U_{s,i}$ in continuum time τ is obtained iterating Eq. (4) for $\Delta t = P d\tau$ time steps. Using Eq. (3) in the form $m_i = \tanh[\Gamma(U_{+1,i} - U_{-1,i})]$, we find

$$\frac{dm_i}{d\tau} = -2\Gamma(1 - m_i^2)\left[\overline{\Omega\,\xi_i} + \sum_j \overline{\xi_i \xi_j} m_j\right]. \quad (6)$$

This can be easily written as a gradient descent dynamics $dm_i/d\tau = -\Gamma(1 - m_i^2)(\partial H/\partial m_i)$ which minimizes the Hamiltonian

$$H = \overline{\langle A\rangle^2} = \sigma^2 - \sum_i \overline{\xi_i^2}(1 - m_i^2). \quad (7)$$

As a function of m_i, H is a positive definite quadratic form, which has a unique minimum. This implies that *the stationary state of the MG is described by the ground state properties of H*. It is easy to see [15] that H is closely related to the order parameter $\theta = \sqrt{\overline{\langle \text{sgn}A\rangle^2}}$ introduced in [8], which is a measure of the system predictability [8]. Indeed $H \propto \theta^2$ when θ is small, suggested that inductive agents actually minimize predictability rather than their collective losses σ^2.

It is possible to study the ground state properties of H in Eq. (7) using the replica method [1]. First, we introduce an inverse temperature β [16] and compute the average over the disorder variables $\Xi = \{a_{s,i}^{\mu}\}$ of the partition function of n replicas of the system, $\langle Z^n \rangle_\Xi$. Next, we perform an analytic continuation for noninteger values of n, thus obtaining $\langle \ln Z \rangle_\Xi = \lim_{n\to 0} \frac{\langle Z^n\rangle_\Xi - 1}{n}$. The "free energy" $F_{ID} = -\langle \ln Z \rangle_\Xi / \beta$ depends on the overlap matrix $Q_{a,b} = \langle m_i^a m_i^b \rangle$ $(a, b = 1, \ldots, n, a \neq b)$ and on the order parameter $Q_a = (1/N)\sum_i (m_i^a)^2$, together with their Lagrange multipliers $r_{a,b}$ and R_a, respectively. F_{ID} can be calculated using a saddle point method that, within the replica symmetric (RS) ansatz $Q_{a,b} = q$, $r_{a,b} = r$ (for all $a < b$), and $Q_a = Q$, $R_a = R$ (for all a), leads to

$$F_{ID} = \frac{\alpha}{2} \frac{1 + q}{\alpha + \beta(Q - q)} + \frac{\alpha}{2\beta} \log\left[1 + \frac{\beta(Q - q)}{\alpha}\right] + \frac{\beta}{2}(RQ - rq) - \frac{1}{\beta} \int d\Phi(\zeta) \log \int_{-1}^{1} ds\, e^{-\beta V(s|\zeta)},$$

where $V(x \mid \zeta) = \beta(r - R)(x^2/2) - \sqrt{r}\,\zeta x$, and Φ is the normal distribution. The ground state properties of H are obtained solving the saddle point equations [1] in the limit $\beta \to \infty$. Figure 1 compares the analytic and numerical findings for σ^2. For $\alpha > \alpha_c = 0.337\,40\ldots$, the solution leads to $Q = q < 1$ and a ground state energy $H_0 > 0$. $H_0 \to 0$ as $\alpha \to \alpha_c^+$ and $H_0 = 0$ for $\alpha \leq \alpha_c$.

This confirms the conclusion $\langle A^\mu \rangle = 0\ \forall \mu$ [8] (or $\theta = 0$) for $\alpha \leq \alpha_c$ and it implies the relation

$$\sigma^2 = \sum_i \overline{\xi_i^2}(1 - m_i^2) \cong \frac{N}{2}(1 - Q), \qquad \alpha \leq \alpha_c\,. \tag{8}$$

The RS solution is stable against replica symmetry breaking (RSB) for any α, as expected from positive definiteness of H. Following Ref. [10], we compute the probability distribution of the strategies, which for $\alpha > \alpha_c$ is bimodal and it assumes the particularly simple form

$$\mathcal{P}(m) = \phi(z)[\delta(m - 1) + \delta(m + 1)] + \frac{z}{\sqrt{2\pi}} e^{-(zm)^2/2}, \tag{9}$$

with $z = \sqrt{\alpha/(1 + Q)}$ (Q taking its saddle point value) and where $\phi(z) = [1 - \mathrm{erf}(z/\sqrt{2})]/2$ is the fraction of frozen agents (those who always play one and the same strategy). Below α_c, $\mathcal{P}(m)$ is continuous, i.e., $\phi = 0$ in agreement with numerical findings [8].

At the transition the spin susceptibility $\chi = \lim_{\beta\to\infty}\beta(Q - q)$ diverges as $\alpha \to \alpha_c^+$, and it remains infinite for all $\alpha \leq \alpha_c$. This is because the ground state is degenerate in many directions (zero modes) and an infinitesimal perturbation can cause a finite shift in the equilibrium values of m_i. This implies that in the long run the dynamics (6) leads to an equilibrium state which depends on the initial conditions $U_{s,i}(t = 0)$. The underconstrained nature of the system is also responsible for the occurrence of antipersistent effects for $\alpha < \alpha_c$ [8]. The periodic motion in the subspace $H = 0$ is probably induced by inertial terms $d^2U_{s,i}/d\tau^2$ which we have neglected, and which require a more careful study of dynamical solutions of Eqs. (3) and (4). It is, however, clear that the amplitude of the excursion of $U_{+1,i}(t) - U_{-1,i}(t)$ decreases with Γ, by the smoothing effect of Eq. (3). When this amplitude becomes of the same order of $1/\Gamma$ antipersistence is destroyed, which explains the sudden drop of σ^2 with Γ found in Ref. [13].

A natural question arises: Is this state individually optimal, i.e., it is a Nash equilibrium of the game where agents maximize the expected utility $\overline{u_i} = -\overline{a_{s,i}A}$? One way to find the Nash equilibria is to consider stationary solutions of the multipopulation replicator dynamics [17]. This takes the form of an equation for the so-called *mixed* strategies, i.e., for the probabilities $\pi_{s,i}$ with which agent i plays strategy s. In terms of $m_i = \pi_{+,i} - \pi_{-,i}$, with a little algebra, these equations [17] read

$$\frac{dm_i}{d\tau} = (1 - m_i^2)\frac{\partial \overline{u_i}}{\partial m_i}\,. \tag{10}$$

Observing that $\partial\overline{u_i}/\partial m_i = -\partial\sigma^2/\partial m_i$, we can rewrite Eq. (10) as a gradient descent dynamics which minimizes a global function which is *exactly* the total loss σ^2 of agents. Nash equilibria then correspond to the local minima of σ^2 in the domain $[-1, 1]^N$. The quadratic form σ^2 is not positive definite, which means that there shall be many local minima and the Nash equilibrium is not unique. It is easy to see [15] that Nash equilibria are in *pure* strategies, i.e., $m_i^2 = 1\ \forall i$, which implies $\sigma^2 = H$, by Eq. (7). A detailed characterization of the Nash equilibria shall be given elsewhere [15]. The best Nash equilibrium can be studied applying the replica method to σ^2 for $\beta \to \infty$. The multiplicity of Nash equilibria (metastable states) manifests itself in the occurrence of replica symmetry breaking for any $\alpha > 0$ with a nonvanishing σ^2/N [15]. The simple RS solution, though incorrect, provides a close lower bound $F_{\rm NE}^{\rm (RS)} = F_{ID} + \frac{1}{2}(1 - Q)$ to σ^2/N for $\beta \to \infty$ (see Fig. 1). For $\alpha > 1/\pi$, we have $Q = q = 1$ and $F_{\rm NE}^{\rm (RS)}$ $(\beta = \infty) = [1 - 1/\sqrt{\pi\alpha}\,]^2$ positive, whereas $1 = Q < q$ and $F_{\rm NE}^{\rm (RS)} = 0$ for $\alpha < 1/\pi$.

Figure 1 shows that in a Nash equilibrium agents perform much better than in the MG. This is the consequence of the fact that agents do not take into account their impact on the market (i.e., on A^μ) when they update the scores of their strategies by Eq. (4). It is indeed known [18] that reinforcement-learning dynamics based on Eq. (3) is closely related to the replicator dynamics and, hence, it converges to rational expectation outcomes, i.e., to Nash equilibria. More precisely, Ref. [18] suggests that this occurs if Eq. (4) is replaced with

$$U_{i,s}(t + 1) = U_{i,s}(t) + u_i^{\mu(t)}[s, s_{-i}(t)]/P\,. \tag{11}$$

Now $U_{s,i}(t)$ is proportional to the cumulated payoff that agent i would have received had she always played strategy s (with other agents playing what they actually played) until time t. As Fig. 1 again shows, this leads to results which coincide with those of the Nash equilibrium. It is remarkable that the (relative) difference between Eqs. (4) and (11) is small, i.e., of order $1/A^\mu \sim 1/\sqrt{N}$. Yet, it is *not negligible* because, when averaged over all states μ, it produces a finite effect, especially for $\alpha < \alpha_c$, and it effects considerably the nature of the stationary state. This term has the same origin of the cavity reaction term in spin glasses [1]. In order to follow Eq. (11), agents need to know the payoff they would have received for any strategy s they could have played. That may not be realistic in complex situations where agents know only the payoffs they receive and are unable to disentangle their contribution from $G(A^\mu)$. However, agents can account approximately for their impact on the market by adding a cavity term $+\eta\delta_{s,s_i(t)}$ to Eq. (4) which "rewards" the strategy $s_i(t)$ used with respect to those $s \neq s_i(t)$ not used. The most striking effect of this new term, as discussed elsewhere [15] in detail, is that for $\alpha < \alpha_c$ an *infinitesimal* $\eta > 0$ is sufficient to cause RSB and to reduce σ^2/N by a *finite* amount.

Thus far, the information $\mu(t)$ was randomly and independently drawn at each time t from the distribution $\varrho^\mu = 1/P$. In the original version of the MG [3], μ is instead endogenously determined by the collective dynamics of agents: $\mu(t)$ indeed labels the sequence of the last $M = \log_2 P$ "minority" signs, i.e., $\mu(t+1) = [2\mu(t) + 1]_{\mathrm{mod}P}$ if $A^{\mu(t)} > 0$, and $\mu(t+1) = [2\mu(t)]_{\mathrm{mod}P}$ otherwise. The idea [3] is that the information refers to the recent past history of the market, and agents try to guess trends and patterns in the time evolution of the process $G(A^{\mu(t)})$. We may say that $\mu(t)$ is *endogenous* information, since it refers to the market itself, as opposed to the *exogenous* information case discussed above.

Numerical simulations [9] show that the collective behavior of the MG, based on Eq. (4), under endogenous information is the same as that under exogenous information. Within our approach, the relevant feature of the dynamics of $\mu(t)$ is its stationary state distribution ϱ^μ. The key point is that a finite fraction $1 - \phi$ of agents behave stochastically ($m_i^2 < 1$) because $Q < 1$. As a consequence, A^μ has stochastic fluctuations of order $\sqrt{N(1-Q)}$ which are of the same order of its average $\langle A^\mu \rangle \sim \sqrt{H}$. With endogenous information, these fluctuations of A^μ induce a dynamics of $\mu(t)$ which is ergodic in the sense that typically each μ is visited with a frequency $\varrho^\mu \simeq 1/P$ in the stationary state [15]. The situation changes completely when agents follow Eq. (11). Indeed the system converges to a Nash equilibrium where agents play in a deterministic way, i.e., $m_i^2 = 1$ (or $Q = \phi = 1$). The noise due to the stochastic choice of s_i by Eq. (3) is totally suppressed. The system becomes deterministic and the dynamics of $\mu(t)$ locks into some periodic orbit. The ergodicity assumption then breaks down: Only a small number $\tilde{P} \ll P$ of patterns μ are visited in the stationary state of the system, whereas the others never occur ($\varrho^\mu = 0$). This leads to an effective reduction of the parameter $\alpha \to \tilde{\alpha} = \tilde{P}/N$, which further diminishes σ^2. Numerical simulations show that $\tilde{P} \propto \sqrt{P}$ which imply that $\tilde{\alpha} \to 0$ in the limit $P = \alpha N \to \infty$, i.e., $\sigma^2/N \to 0$.

In summary, we have shown how methods of statistical physics of disordered systems can successfully be applied to study models of interacting heterogeneous agents. Our results extend easily to more general models [15] and, more importantly, the key ideas can be applied to more realistic models of financial markets, where heterogeneities arise, e.g., from asymmetric information.

We acknowledge J. Berg, A. De Martino, S. Franz, F. Ricci-Tersenghi, S. Solla, M. Virasoro, and Y.-C. Zhang for discussions and useful suggestions. This work was partially supported by Swiss National Science Foundation Grant No. 20-46918.98.

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

[2] P. W. Anderson, K. Arrow, and D. Pines, *The Economy as an Evolving Complex System* (Addison-Wesley, Reading, MA, 1988); *Econophysics: An Emerging Science*, edited by J. Kertesz and I. Kondor (Kluwer, Dordrecht, 1998).

[3] D. Challet and Y.-C. Zhang, Physica (Amsterdam) **246A**, 407 (1997); Y.-C. Zhang, Europhys. News **29**, 51 (1998).

[4] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory* (Oxford University, New York, 1995).

[5] See www.unifr.ch/econophysics for a complete collection of references on the minority game.

[6] W. B. Arthur, Am. Econ. Assoc. Papers Proc. **84**, 406 (1994).

[7] R. Savit, R. Manuca, and R. Riolo, Phys. Rev. Lett. **82**, 2203 (1999).

[8] D. Challet and M. Marsili, Phys. Rev. E **60**, R6271 (1999).

[9] A. Cavagna, Phys. Rev. E **59**, R3783 (1999).

[10] N. Brunel and R. Zecchina, Phys. Rev. E **49**, R1823 (1994).

[11] Both ω_i^μ and ξ_i^μ take values in $\{0, \pm 1\}$ but they are not independent: $\omega_i^\mu \xi_i^\mu = 0$ and $\omega_i^\mu + \xi_i^\mu \neq 0$ for all i, μ.

[12] D. M. Kreps, *Game Theory and Economic Modelling* (Oxford University, New York, 1990).

[13] A. Cavagna, J. P. Garrahan, I. Giardina, and D. Sherrington, Phys. Rev. Lett. **83**, 4429 (1999).

[14] See, e.g., M. Marsili, Physica (Amsterdam) **269A**, 9 (1999).

[15] M. Marsili, D. Challet, and R. Zecchina, cond-mat/9908480; D. Challet, M. Marsili, and Y.-C. Zhang (to be published).

[16] β is introduced as a device to study the minima of H and it should not be confused with Γ. Equation (6) suggests that Γ is not an inverse temperature, in this context, but rather the learning rate. However, a system with a finite memory—i.e., replacing Eq. (4) by $U_{s,i}(t+1) = (1 - \epsilon)U_{s,i}(t) - a_{s,i}^{\mu(t)} G(A^{\mu(t)})$—is described by the properties of H at a finite temperature $\beta = \Gamma/\epsilon$.

[17] J. W. Weibull, *Evolutionary Game Theory* (MIT, Cambridge, MA, 1995).

[18] A. Rustichini (to be published); D. Fudenberg and D. K. Levine, *The Theory of Learning in Games* (MIT, Cambridge, MA, 1998).