Exact solution of a modified El Farol's bar problem

Matteo Marsili

Istituto Nazionale per la Fisica della Materia (INFM), Trieste-SISSA Unit, V. Beirut 2-4, Trieste I-34014

Damien Challet

Institut de Physique Théorique, Université de Fribourg, CH-1700

and Riccardo Zecchina

The Abdus Salam International Centre for Theoretical Physics Strada Costiera 11, P.O. Box 586, I-34014 Trieste

Abstract

We discuss the Minority Game, a model of heterogeneous, indictive rational agents inspired by the El Farol Bar problem. We derive the exact solution of the model in the limit of infinitely many agents using tools of statistical physics of disordered systems. Our results show that the impact of agents on the collective behavior plays a key role: even though for each agent this effect is very small, the collective behavior crucially depends on whether agents account for it or not.

1 Introduction

The El Farol bar problem [1] has become a popular paradigm of complex systems. It describes the situation where N persons have to choose whether to go or not to a bar which is enjoyable only if it is not too crowded. In order to choose, each person forms mental schemes, hypotheses or behavioral rules based on her beliefs and she adopts the most successful one on the basis of

⋆ We acknowledge J. Berg, S. Franz, F. Vega-Redondo, M. Virasoro and Y.-C. Zhang for discussions and useful suggestions. This work was partially supported by Swiss National Science Foundation Grant Nr 20-46918.98.

past performance. This approach to human behavior, based on inductive rationality, is quite different from the one of mathematical economy and game theory[2], which assumes deductive rationality. Inductive [1], low [3] or generally bounded rationality based on learning theory [4] is regarded as a more realistic model of how real agents behave in complex strategic situations. This is specially true in contexts involving many heterogeneous agents with limited information, such as the El Farol bar problem. Theoretical advances, beyond computer experiments, is technically very hard for these problems and it has been regarded as a major step forward in the understanding of complex systems[5].

The minority game [6,7] represents a first step in this direction. It indeed describes a system of interacting agents with inductive rationality which face the problem of finding which of two alternatives shall be chosen by the minority. This problem is quite similar in nature to the El Farol bar problem as the result for each agent depends on what all other agents will do and there is no best alternative. These same kind of situations arise generally in systems of many interacting adaptive agents, such as markets[7].

Computer experiments by several authors [6,8–12] have shown that the minority game (MG) displays a remarkably rich emergent collective behavior, which has been qualitatively understood to some extent by approximate schemes [7,13,14]. In this paper, which follows refs. [11,15], we show that a full statistical characterization of the MG in the long run can be derived analytically in the "thermodynamic" limit of infinitely many agents. Our approach is based on tools and ideas of statistical physics of disordered systems[16].

The minority game, as the El Farol bar problem, allows for a relatively easy definition in words. This may be enough for setting up a computer code to run numerical simulations, but it is clearly insufficient for an analytical approach. Therefore we shall, in the next section, define carefully its mathematical formulation in levels of increasing complexity. The minority game, as it has been introduced[6,7], shall only emerge little by little as we go through the stages of our construction. This may disorient readers familiar with the literature on the MG but we hope they shall benefit, as we did, from a deeper understanding of its structure and it strategic implications. Clarifying the structure of the game also allows to understand which observed feature depends on which ingredient of the model, what are fundamental and secondary elements and, finally, what situations the structure of the model describes in general.

It is worth to stress that we take the minority game as given and derive an analytical understanding of it. We shall only mention some briefly its motivation, for which we refer the reader to refs. [6–8]. At the same time we shall not try to justify the (crude) assumptions it makes on agents behavior but just discuss their consequences.

After defining the stage game, we shall briefly discuss its Nash equilibria: these are the reference equilibria of deductive rational agents. Finally we shall pass to the repeated game with adaptive agents. We show that the key difference between agents playing a Nash equilibrium and agents in the usual minority game is *not* that the first are deductive whereas the latter are inductive. Indeed with inductive rationality of the same form as that assumed in the MG agents converge indeed to a Nash equilibrium. The key issue is instead related to the "market impact". If agents knew their "market impact" and consider it when evaluating the performance of their strategies, the system converges towards a Nash equilibrium. In the minority game agents are *naive* since they behave as "price takers", i.e. as if their choices did not affect the aggregate. As a result they converge to a sub-optimal state which is not a Nash equilibrium. In other words, the neglect of market impact, which seems an innocent approximation in the limit $N \to \infty$ [1], plays a very important role in this complex system. We find it more appropriate to discuss these results in the final section of this paper.

2 The stage game

2.1 Level 0: actions and payoffs

The minority game describes a situation where a large odd number N of agents have to make one of two opposite actions – such as e.g. "buy" or "sell" – and only those agents who choose the minority action are rewarded. This is similar to the El Farol bar problem, where each one of N agents may either choose to go or not to a bar which is enjoyable only when it is not too crowded. In order to model this situation, let $\mathcal{N} = (1, \ldots, N)$ be the set of agents and let $\mathcal{A} = (-1, +1)$ be the set of the two possible actions. If $a_i \in \mathcal{A}$ is the action of agent $i \in \mathcal{N}$, the payoffs to agent i is given by

$$u_i(a_i, a_{-i}) = -a_i\, A \quad \text{where} \quad A = \sum_{i\in\mathcal{N}} a_i, \tag{1}$$

where $a_{-i} = \{a_j,\, j \neq i\}$ stands for opponents actions. The game rewards the minority group. To see this, note that the total payoff to agents $\sum_i u_i = -A^2$ is always negative. Then the majority of agents, who have $a_i = \operatorname{sign} A$, receives a negative payoff $-|A|$, whereas the minority "wins" a payoff of $|A|$. Eq. (1) can be generalized to $u_i(a_i, a_{-i}) = -a_i\, U(A)$: if the function $U(x)$ is such that $x\, U(x) = -xU(-x) \geq 0$ for all $x \in \mathbb{R}$, the game again rewards the minority.

[1] The impact of each agent on the aggregate is of relative order $1/N$ and it vanishes as $N \to \infty$.

The original model[6,7] takes $U(x) = \operatorname{sign} x$, but the collective behavior is qualitatively the same [11] as that of the linear case $U(x) = x$ on which we focus. Note that the "inversion" symmetry $u(-a_i, -a_{-i}) = u(a_i, a_{-i})$ implies that the two actions are *a priori* equivalent: there cannot be any best actions, because otherwise everybody would do that and loose.

The game's structure is similar to that of coordination games [2, chapt. 6]. Agents cannot communicate. If communication were possible, agents would have incentives to stipulate contracts – such as "We toss a coin, if the outcome is head I do $a_{\rm me} = +1$ and you do $a_{\rm you} = -1$, and if it is tail we do the other way round". Both players would benefit from this contract because it transforms the negative sum game into a zero sum game for the two players. The contract would then be self-enforcing.

Agents interact only through a *global* or *aggregate* quantity A^μ which is produced by all of them. This type of interaction is typical of market systems [7] ant it is similar to the long-range interaction assumed in mean-field models of statistical physics[16]. Finally note that the El Farol bar problem has a similar structure but with A replaced by $(A - A_0)$ in Eq. (1) where A_0 is related to the bar's comfort level [1,17].

2.2 Level 1: information

Next we introduce *public information* in the form of an integer variable μ which takes integer values $\mu \in \mathcal{P} \equiv \{1, 2, \ldots, P\}$. We assume that P is large and of the same order of N and we define $\alpha \equiv P/N$, which we eventually keep finite in the limit $N \to \infty$. We shall discuss later the origin of information. For the moment being, let it suffice to say that μ labels the state of the environment where agents live and that it is drawn from a distribution ϱ^μ on $\mathcal{P}$. In what follows, we shall denote with an over-line $\overline{O} \equiv \sum_{\mu\in\mathcal{P}} \varrho^\mu O^\mu$ the average of a quantity O^μ over μ. For any $\mu \in \mathcal{P}$, payoffs are still given by Eq. (1). Strictly speaking μ is a so called *sun-spot*[18] because the payoffs only depend on the actions of agents. Now, however, the *pure* strategies of each agent may depend on μ. We call $\mathcal{A}^\mathcal{P}$ the set of all such strategies: An element of $\mathcal{A}^\mathcal{P}$ is a function $a : \mu \in \mathcal{P} \to a^\mu \in \mathcal{A}$ or a P dimensional vector with coordinates a^μ, $\forall \mu \in \mathcal{P}$ [2]. There are $|\mathcal{A}^\mathcal{P}| = 2^P$ possible such functions. We call $a_i \in \mathcal{A}^\mathcal{P}$ a possible pure strategy for agent $i \in \mathcal{N}$, with elements $a_i^\mu \in \mathcal{A}$ for all $\mu \in \mathcal{P}$. With this notations, the payoff "matrix" reads $\overline{u_i}(a_i, a_{-i}) = -\overline{a\,A}$ where $A^\mu = \sum_i a_i^\mu$. At this level we have just replicated the game of the previous level P times. Again there cannot be a best strategy $a \in \mathcal{A}^\mathcal{P}$ for the same reasons as before.

[2] We use the simple letter a without the index $^\mu$ to denote the function.

2.3 Level 2: Heterogeneous beliefs and strategies

Now we assume that each agent only restricts her choice on a small subset of S elements of $\mathcal{A}^{\mathcal{P}}$. We use the vector notation $\vec{a}_i = (a_{1,i}, \ldots, a_{S,i})$ to denote the subset of strategies available to agent i, with elements $a_{s,i} \in \mathcal{A}^{\mathcal{P}}$ (the action of agent i, when the information is μ and she chooses her s^{th} strategy shall then be $a^\mu_{s,i} \in \mathcal{A}$. The strategies $a_{s,i}$ are randomly and independently drawn (with replacement) from $\mathcal{A}^{\mathcal{P}}$. More precisely

$$P(a^\mu_{s,i} = +1) = P(a^\mu_{s,i} = -1) = \frac{1}{2}, \quad \forall i \in \mathcal{N},\ s \in (1, \ldots, S),\ \mu \in \mathcal{P}. \quad (2)$$

The utility of agent i, given the value of μ, his choice s_i and the choice of other agents $s_{-i} = \{s_j,\ j \neq i\}$, now becomes

$$u^\mu_i(s_i, s_{-i}) = -a^\mu_{s_i,i}\, A^\mu \quad \text{with} \quad A^\mu \equiv \sum_{j \in \mathcal{N}} a^\mu_{s_j,j}. \quad (3)$$

The goal of each agent is to maximize his expected payoff over all possible values of μ, which, for agent i, reads

$$\overline{u_i}(s_i, s_{-i}) = -\overline{a_{s_i,i}\, A} \equiv \quad (4)$$

This structure was introduced in refs. [6,7] in order to model *inductive rational* behavior of agents [1]: If agents have not a completely detailed model of the game they are engaged in, they may think that the value μ has some effect on the game's outcome A, eventually because she believes that other agents will believe the same. This is a self-reinforcing belief because if agents behave differently for different values of μ, the aggregate outcome A will indeed depend on μ, thus confirming agents' beliefs. The game's structure, however, is such that there can be no "commonality of expectations" [1]. This means that, since there is no rational best strategy a_{best} – because otherwise everybody would use that and loose – agents' expectations (or beliefs) are forced to differ.

On the basis of her mental schemes or hypotheses on the situation she faces, an agent may consider a particular strategy more "likely" to predict the "correct" action than another one. More precisely, she may consider that only the S forecasting rules $\vec{a}_i = (a_{1,i} \ldots, a_{S,i})$, out of all the 2^P such rules in $\mathcal{A}^{\mathcal{P}}$, are "reasonable" or compatible with her beliefs and then restrict her choice to just those [3]. Heterogeneous beliefs are represented by the fact that each agent

[3] This rises the following question: Do agents have incentives to increase the number S of strategies? As we shall see later, the answer is generally yes. To fix ideas we may think that agents consider only S strategies because of computational costs we

draws her strategies at random, independently from others using Eq. (2). We refer the reader to refs. [1,7] for a deeper discussion of this behavioral model.

It is worth to point out that construction may seem a bit arbitrary: why should agents restrict to only S strategies? If the information μ is known before they take their decision, why don't they take a decision s_i^μ which depends on μ – or even an action a_i^μ which depends on μ? On the other hand, one can assume that the information μ is not known and that agents resort to S "devices" which take actions $a_{s,i}^\mu$ for them. They do not observe μ but only the global outcome $A(t)$ and the result $a_{s_i,i}^\mu$ of their choice s_i and that $a_{s,i}^\mu$ of all other possible choices $s \neq s_i$. Note that, if $\vec{a}_i^\mu$ where known to agent i before taking her decision, it would be reasonable for her to consider independent decisions conditional on her *private* information $\vec{a}_i^\mu$. Again that would be quite different from the present model.

We shall not try to defend the behavioral assumptions in refs. [1,6–8], but rather observe that there are other situations for which this structure can be relevant. For instance, the game also describes a situation where agents have only limited control on their actions which are also determined by the variable μ. In this case we should assume that μ is revealed only after agents take their decision or that μ is not observable altogether. Alternatively we can think that agents are engaged in P simultaneous games and that they weight the outcome of game μ by ϱ^μ. Anyway, the restriction on the strategic choice of agents (from $\mathcal{A}^{\mathcal{P}}$ to $\vec{a}_i$) is due to implicit individual constraints we are not modeling explicitly.

Before coming to the analysis of the game, let us introduce *mixed strategies*. The mixed strategy $\pi_{s,i}$, $s = 1, \ldots, S$ of each agent i is a distribution over her available strategies: $\pi_{s,i}$ is, in other words, the probability with which agent i plays her s^{th} strategy. Again we use a vector notation $\vec{\pi}_i = (\pi_{1,i}, \ldots, \pi_{S,i}) \in \Delta_i$, where Δ_i is the S dimensional simplex of i^{th} agent. We introduce the scalar product $\vec{u} \cdot \vec{v} \equiv \sum_{s=1}^S u_s, v_s$ for vectors $\vec{u}, \vec{v} \in I\!R^S$. We also define the norm $|v|^2 \equiv \vec{v} \cdot \vec{v}$ of vectors in $I\!R^S$. Expectations on the mixed strategy of player i reads $E_{\vec{\pi}_i}(\vec{v}) = \vec{\pi}_i \cdot \vec{v}$. We also define the direct product $\Delta^{\mathcal{N}} = \prod_{i \in \mathcal{N}} \Delta_i$ which we shall also call the *phase space*. A point $(\vec{\pi}_1, \ldots, \vec{\pi}_N) \in \Delta^{\mathcal{N}}$ is indeed a possible state of the system. Finally we use the shorthand notation

$$\langle O \rangle = \sum_{s_1=1}^S \cdots \sum_{s_N=1}^S \pi_{s_1,1} \cdots \pi_{s_N,N} \, O_{s_1,\ldots,s_N} \tag{5}$$

do not model explicitly

for the expectation on the product measure of mixed strategies over the phase space $\Delta^{\mathcal{N}}$. For example we shall frequently refer to the quantity

$$\langle A^\mu \rangle = \sum_{i\in\mathcal{N}} \vec{\pi}_i \cdot \vec{a}_i^\mu . \tag{6}$$

3 Global quantities of interest

As a preliminary to a more detailed discussion, we find it useful to introduce the key quantities which describe the collective behavior. First, as a measure of global efficiency, we take

$$\sigma^2 \equiv \overline{\langle A^2 \rangle} = \sum_{\mu\in\mathcal{P}} \varrho^\mu \left\langle \left(\sum_{i\in\mathcal{N}} a^\mu_{s_i,i} \right)^2 \right\rangle \tag{7}$$

where we remind that $\langle \ldots \rangle$ here means expectation over the variables s_i with the corresponding mixed strategy distribution $\pi_{s_i,i}$. Note that $\sigma^2 = -\sum_i \langle \overline{u_i} \rangle$ is just the total loss of agents. A small value of σ^2 implies an efficient coordination among agents.

By construction, the model is symmetric in the sense that, for any μ, no particular sign of A^μ is *a priori* preferred. We shall see, however, that this symmetry can be "broken" resulting in a state where A^μ may take more probably positive than negative values for some μ and *vice-versa* for other values of μ. As a measure of this asymmetry, it is useful to introduce the quantity

$$H \equiv \overline{\langle A \rangle^2} = \sum_{\mu=1}^{P} \varrho^\mu \langle A^\mu \rangle^2 = \sum_{\mu=1}^{P} \varrho^\mu \left(\sum_{i\in\mathcal{N}} \vec{\pi}_i \vec{a}_i^\mu \right)^2 . \tag{8}$$

Note that $H > 0$ implies that there is a *best* strategy $a^\mu_{\mathrm{best}} = -\mathrm{sign}\,\langle A^\mu \rangle$ that could ensures a positive payoff to a new-comer agent. Ideally because if the new-comer really starts playing the game, she will also affect the outcome A^μ. This suggests that H can be regarded as a measure of the exploitable *information content* of the system by an external agent[11].

Both of these quantities are *extensive*, i.e. are proportional to N for $N \to \infty$, and we shall mainly be interested in the finite quantities σ^2/N and H/N. As a statistical characterization of the equilibrium, it is useful to introduce the *self overlap*

$$G = \frac{1}{N} \sum_{i\in\mathcal{N}} |\pi_i|^2 = \frac{1}{N} \sum_{i\in\mathcal{N}} \sum_{s=1}^{S} \pi_{s,i}^2 \tag{9}$$

which gives a measure of the average spread of mixed strategies played by agents. If all agents play pure strategies $G = 1$ whereas $G = 1/S$ if $\pi_{s,i} = 1/S$ $\forall s, i$. Therefore $1/G$ is a measure of the "effective" number of strategies that agents play on average. Note that σ^2 can be written, for large N, as

$$\frac{\sigma^2}{N} \cong \frac{H}{N} + 1 - G \tag{10}$$

where we neglected terms which vanish in the limit $N \to \infty$.

4 Nash Equilibria

Given the specification of payoffs and available choices given above, at each level of the game, we now briefly discuss Nash equilibria of the stage game[4]. These equilibria shall provide a reference framework for the following discussion.

4.0.0.1 Level 0 : Given the symmetry of the game, let us first look for symmetric Nash equilibria. These cannot be in pure actions[5] so let us look for Nash equilibria in mixed actions: Each player either plays $a_i = +1$ with probability π_i or she plays $a_i = -1$ otherwise. It is easy to see that $\pi_i = 1/2$ is the only symmetric Nash equilibrium: No agent has incentives to deviate from the choice $\pi_i = 1/2$ if others stick to it. This state, which we shall call the *random agent state*, is usually taken as a reference state[8,13] and it is characterized by $\sigma^2 = N$ and $H = 0$.

The game has many more Nash equilibria than the symmetric one. With N odd, indeed, any state where $|A| = 1$ is a Nash equilibrium. Indeed no agent in the minority (playing $a_i = -A$) would decrease her payoff, switching to the majority side. On the other hand agents in the majority cannot increase their payoff changing from $a_i = A$ to $a_i = -A$ because then, also the majority would change $A \to -A$. The number $\Omega^{(0)}_{\text{Nash}}$ of these Nash equilibria is exponentially large in N and it is given by

$$\Omega^{(0)}_{\text{Nash}} = \binom{N}{\frac{N-1}{2}} + \binom{N}{\frac{N+1}{2}}, \qquad \binom{N}{k} \equiv \frac{N!}{k!(n-k)!} \tag{11}$$

[4] For our purposes, Nash equilibria are those states which are stable under payoff incentives, given the choices available to agents. We shall not discuss refinements.
[5] Strategies at this level are actions. We use both terms interchangeably

Each of these states is globally optimal, since it has no fluctuation: $\sigma^2 = H = 1$ as compared to $\sigma^2 = N$, $H = 0$ in the symmetric Nash equilibrium.

4.0.0.2 Level 1:

Again we have the symmetric Nash equilibrium – the random agents state – where, for any μ, each agent plays $a = +1$ with probability 1/2. The number of Nash equilibria in pure actions is huge. Indeed any combination of P Nash equilibria at level 0 (one for each value of μ), is a Nash equilibrium. There are therefore $\Omega^{(1)}_{\text{Nash}} = \left(\Omega^{(0)}_{\text{Nash}}\right)^P$ such equilibria each of them with minimal fluctuations $\sigma^2 = 1$.

4.0.0.3 Level 2:

With heterogeneous strategies the problem of identifying Nash equilibria becomes more complex. One way to tackle the problem is to write down replicator dynamics [19] and then identify (evolutionarily stable) Nash equilibria by its stationary points. The replicator dynamics [19] (RD) reads

$$\frac{d\pi_{s,i}}{dt} = -\pi_{s,i} \sum_{j\in\mathcal{N}, j\neq i} \left[\overline{a_{s,i}(\vec{\pi}_j \cdot \vec{a}_j)} - \overline{(\vec{\pi}_i \cdot \vec{a}_i)(\vec{\pi}_j \cdot \vec{a}_j)}\right]. \tag{12}$$

We observe that $\sigma^2 = \sum_{i,j\neq i} \overline{(\vec{\pi}_i \cdot \vec{a}_i)(\vec{\pi}_j \cdot \vec{a}_j)} + N$ is a Lyapunov function under this dynamics, i.e.

$$\frac{d\sigma^2}{dt} = -2 \sum_{i\in\mathcal{N}} \sum_{s=1}^{S} \pi_{s,i} \left[\overline{(a_{s,i} - \vec{\pi}_i \cdot \vec{a}_i) \sum_{j\neq i} \vec{\pi}_j \cdot \vec{a}_j}\right]^2 \leq 0. \tag{13}$$

Therefore Nash equilibria are local minima of σ^2 in $\Delta^{\mathcal{N}}$. Furthermore, since σ^2 is a linear function of $\pi_{i,s}$ for any i, s, it is also an harmonic function in $\Delta^{\mathcal{N}}$ which therefore implies that the minima are on its boundary. This holds for any subset of variables $\pi_{i,s}$ which therefore implies that minima are located in the corners of the simplex, i.e. Nash equilibria are in pure strategies ($G = 1$).

A detailed characterization of Nash equilibria shall be given elsewhere [20]. Here we briefly mention that also in this case Nash equilibria are exponentially many in N. This makes the analytic calculation a step more difficult than the one we shall present later for inductive rational agents. A simplified

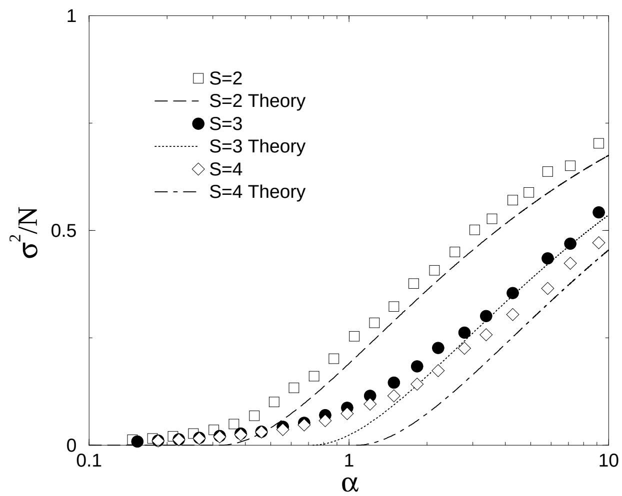

Fig. 1. Global efficiency σ^2/N as a function of α for $S = 2, 3$, and 4 from computer experiments with $P = 128$ averaged over 100 realizations of $\vec{a}_i$ (symbols) and from the theoretical lower bound (lines).

approximate calculation gives the following lower bound[6] :

$$\frac{\sigma^2}{N} \geq \begin{cases} \left[1 - \frac{z(S)}{\sqrt{\alpha}}\right]^2 & \text{for } \alpha > z(S)^2 \\ 0 & \text{for } \alpha \leq z(S)^2 \end{cases} \tag{14}$$

where $\alpha = P/N$ and $z(S) = \sqrt{2/\pi} S \int_{-\infty}^{\infty} dz\, e^{-z^2} z[1 - \mathrm{erfc}(z)/2]^{S-1}$ is the expected value of the maximum among S standard random variable (for $S \gg 1$, $z(S) \simeq \sqrt{2 \ln S}$).

Figure 1 shows that the lower bound is already a good approximation to the typical value of σ^2 in the Nash equilibrium, specially for small values of S. Eq. (14) implies that, for fixed S, σ^2 increases with α, which is reasonable because the complexity of information increases and the resources of agents is limited by S. For fixed α, Eq. (14) suggests that σ^2 decreases with S. So if agents are given more resources (larger S), they attain a better equilibrium. Both of these features are confirmed by computer experiments (see fig. 1).

It is worth to point out that the game specified by the payoffs of Eq. (4), for N and P very large, implies a fantastic computational complexity. Deductive rational agents should be able to master a chain of logical deductions of formidable complexity in order to derive their best response. The efforts required by this strategic situation may well exceed the bounds of memory and computational capabilities of any realistic agent or more simply the re-

[6] Strictly speaking, the meaning of this lower bound is that the probability to observe σ^2 smaller than the lower bound decreases exponentially with N

sources she is likely to devote to the problem. Furthermore her assumption that everybody else behaves as a rational deductive player becomes more and more unrealistic as N grows large. Finally, even with deductive rational agents there would still be the problem of equilibrium selection which, in this case, involves a huge number of possible equilibria.

5 Repeated game: Learning and inductive rationality

Deductive rationality, as suggested in refs. [1,7], is unrealistic in such complex strategic situations [7] and it has to be replaced by inductive rationality. This amounts to assume that agents try to learn what their best choice is from their past performance. We henceforth focus on the repeated game in which agents meet once and again to play the stage game of the previous sections. Different stage games are distinguished by the time label by $t \in I\!N$. For example, $s_i(t)$ denotes the strategy chosen by agent i at time t and $\mu(t)$ the information available at that time.

5.1 Exponential learning:

It is generally accepted that agents follow more likely strategies which have been more successful in the past (the "law of effect" see e.g. [21]). Here we assume that agents follow an *exponential learning* behavior [8]: Each agent i assigns *scores* $U_{s,i}(t)$ to each of their strategies $s = 1, \ldots, S$ and she plays strategy s with a probability which depends exponentially on its score:

$$\pi_{s,i}(t) = \frac{e^{\Gamma_i U_{s,i}(t)}}{\sum_{s'=1}^{S} e^{\Gamma_i U_{s',i}(t)}} \tag{15}$$

where $\Gamma_i > 0$ is a numerical constant, which may differ for each agent $i \in \mathcal{N}$. This particular model for discrete choice – called the *Logit model* – has a long tradition in economics [22] and some experimental support (see e.g. [23]). The MG has been originally introduced with $\Gamma_i = \infty$ [6–8] and only recently it has been generalized to $\Gamma_i < \infty$ [24]. Note that $\pi_{s,i}$ here is no more a mixed strategies – which is the object of agents' strategic choice – but rather it encodes a particular behavioral model.

[7] We do not discuss learning and inductive rationality at the simple levels 0 and 1. For a discussion of reinforcement learning El Farol problem at this level, i.e. on actions rather than on *schemata*, see R. Franke (1999).

[8] There are several behavioral models implementing the law of effect, some of which converge to rational expectation whereas others do not (see e.g. [21]).

At time $t = 0$ scores are set to some $U_{s,i}(0)$, which encodes *prior beliefs*: e.g. $U_{s,i}(0) > U_{s',i}(0)$ means that agent i considers strategy s *a priori* more likely to be successful than s'.

5.2 Full and partial information

Each agent updates the scores of her strategies $U_{s,i}(t) \to U_{s,i}(t+1)$ from time t to time $t+1$ depending on her perception about the success of that strategy at that time. The success of a strategy is quantified by the payoff it delivers to the agent when it is played. A key question is "what is the perceived success of a strategy which has not been played at that time?". One answer is to assume that agents know for each strategy what payoff she would have got if she had actually played that strategy. In the second – which is the one assumed in the MG [9] agents neglect their impact on the aggregate A^μ and update scores *as if* A^μ had not changed if they had used a different strategy. While in the first case agents have *full information* on the effects (payoffs) of all of their strategies, in the second they have only *partial information* and they *naively extrapolate* the effect of the strategies they do not play, neglecting their "market" impact.

We shall consider these two cases – of *full information* and of *partial information* with *naive* agents – separately below, and see that the collective behavior is remarkably different. Before doing that, we shall first discuss the dynamics of information.

5.3 Exogenous vs endogenous information

When the game is repeated, the information variable becomes time dependent – which we denote by $\mu(t)$ – and we need to specify its dynamics. In the El Farol problem and in the MG information refers to the outcome of past games. We call $\mu(t)$ *endogenous* information in this case since it encodes information on the game itself. More precisely, in the MG $\mu(t)$ encodes (in its binary representation) the information on which has been the winning action in the last $M = \log_2 P$ games. This means that μ is updated at each time as:

$$\mu(t+1) = \mathrm{mod}\left(2\mu(t) + \frac{1 + \mathrm{sign}\, A(t)}{2}, P\right), \qquad A(t) = \sum_{i\in\mathcal{N}} a^{\mu(t)}_{s_i(t),i}, \quad (16)$$

Note that $A(t)$ depends on time both through $\mu(t)$ and through the choice $s_i(t)$ of each agent $i \in \mathcal{N}$ at time t. Therefore Eq. (16) implies that the dynamics

[9] and probably also in the El Farol problem, ref. [1] is not very clear on this point.

of $\mu(t)$ is defined by the collective behavior of the game itself. Still payoffs do not depend on $\mu(t)$ which is a sun-spot. However agents can coordinate in such a way that some information $\mu(t)$ – i.e. some pattern in the time series of $A(t)$ – can occur more frequently than some other and eventually some can even never occur. As we shall see below, the only relevant information of this dynamics is the stationary state distribution ϱ^μ of the process in Eq. (16). This equation shows that, in the long run, the distribution ϱ^μ is determined by the collective behavior of agents through $A(t)$.

In the context of the MG, Cavagna [10] has studied the case where $\mu(t)$ is just randomly drawn from an uniform distribution $\varrho^\mu = 1/P$. Remarkably he found by computer experiments that the collective behavior differs only weakly from that under endogenous information. We call μ *exogenous information* in this case since it can be considered to encode information about an external system, eventually the environment where agents live.

6 Results

6.1 Full exogenous information

Agents update their scores $U_{s,i}(t)$ by the following rule:

$$U_{s,i}(t+1) = U_{s,i}(t) + u_i^{\mu(t)}[s, s_{-i}(t)] = -a_{s,i}^{\mu(t)} \left[\sum_{j \neq i} a_{s_j(t),j}^{\mu(t)} + a_{s,i}^{\mu(t)} \right], \quad (17)$$

where $s_{-i}(t)$ stands for the strategies played by other agents at time t. It is known [21] that exponential learning with full information, for a single agent playing against a stationary stochastic process, converges to rational expectations. If we can regard the opponents of i as a stationary process, this implies that i^{th} strategy converges to the best response. If this happens for all players the system converges to a Nash equilibrium. This is indeed what numerical experiments show (see figure 1). In order to get a theoretical understanding of why this is so, let us iterate Eq. (17) for Δt time steps and divide by Δt. This gives

$$\begin{aligned} \frac{dU_{s,i}}{dt} \cong \frac{U_{s,i}(t+\Delta t) - U_{s,i}(t)}{\Delta t} &= \frac{1}{\Delta t} \sum_{\tau=t}^{t+\Delta t-1} u_i^{\mu(\tau)}[s, s_{-i}(\tau)] \\ &\cong \langle \overline{u_i}[s, s_{-i}(t)] \rangle \, . \end{aligned} \quad (18)$$

The $\cong$ sign means we are making two assumptions: *i)* we are taking a continuum time limit where the "infinitesimal" time step Δt is not small. We expect that $U_{s,i}(t)$ changes only significantly and systematically over time intervals of order P. Indeed the score of each strategy depends on its performance for all of the P values of μ. Therefore this peculiar time limit, where $\Delta t \propto P = \alpha N$ is very large, captures the systematic changes in the play of agents. *ii)* we neglect the fluctuations around the average of the right hand side. This is legitimate since, by the central limit theorem, these fluctuations are of order $1/\sqrt{\Delta t} \propto 1/\sqrt{N}$ which in the limit of large N is very small. Note that the time average in Eq. (18) is replaced, in this approximation, by the expectation over the stationary distribution ϱ^μ of $\mu(t)$ and over the distributions $\vec{\pi}_i\ \forall i \in \mathcal{N}$. In the present perspective, $\pi_{s,i}$ becomes the frequency with which agent i plays her s^{th} strategy, rather than her mixed strategy. We can now use Eq. (15) to get an equation for the dynamics of $\pi_{s,i}(t)$. It is easy to check that this leads immediately to

$$\frac{d\pi_{s,i}}{dt} = -\Gamma_i \pi_{s,i} \sum_{j\in\mathcal{N}, j\neq i} \left[\overline{a_{s,i}(\vec{\pi}_j \cdot \vec{a}_j)} - \overline{(\vec{\pi}_i \cdot \vec{a}_i)(\vec{\pi}_j \cdot \vec{a}_j)}\right]. \tag{19}$$

Apart from the factor Γ_i, this coincides with the RD of Eq. (12). Again σ^2 is minimized along the trajectories of Eq. (19): it is easy to check that the time derivative of σ^2 is given by Eq. (13) with an extra factor Γ_i inside the sum on $i \in \mathcal{N}$. We therefore conclude that the *with exponential learning and full information agents coordinate on a Nash equilibrium*. Each agent plays, in the long run a pure strategy, i.e. $G = 1$. The Nash equilibrium to which agents converge depends on the initial conditions $U_{s,i}(0)$, i.e. on prior beliefs: Different initial conditions select different Nash equilibria.

6.2 Full endogenous information

If agents play pure strategies, then $A(t)$ takes the same value A^μ each time $\mu(t) = \mu$. This implies that the dynamics of Eq. (16) for $\mu(t)$ becomes deterministic. More precisely it locks into a periodic orbit $\mu(t+\tau) = \mu(t)$ with some period τ. Only the values of μ into this orbit shall occur in the long run, whereas all other values of μ shall never occur. This means that agents strongly influence the time series of the aggregate $A(t)$. Most remarkably, in doing so, they achieve a much better coordination with respect to the exogenous information case because they reduce the parameter $\alpha = P/N$ to $\tilde{\alpha} = \tau/N$ [10]. Numerical studies show that $\tau \propto \sqrt{P}$ so that, in the limit $N \to \infty$ with $P/N = \alpha$ finite, $\tilde{\alpha} \propto 1/\sqrt{N}$ and also σ^2/N takes a very small value.

[10] Note that the values of μ which occur in the long run are sampled uniformly.

6.3 *Partial information and naive agents*

If agents cannot know what payoff they would have obtained if they had played a strategy $s \neq s_i(t)$ different from the one they actually played, they cannot implement the above dynamics. There may be other learning procedures under which agents may converge to Nash equilibria with only partial information (see [21] for example). The MG assumes that agents still follow exponential learning dynamics (with $\Gamma_i = \infty$ eventually) but they behave as *naive* agents. The MG tries to model the behavior of real agents (speculators) in a toy market[7]: they have several (S) speculative strategies which they test against market moves and eventually use the most successful of them. In evaluating the performance of their forecasting rules, agents think as "price takers". In other words they neglect the impact of their behavior on the market: They approximate the payoff the would have received playing a strategy $s \neq s_i(t)$ different from the one they actually played, assuming that the aggregate $A^{\mu(t)}(t)$ did not depend on it. This seems, at first sight an innocent approximation. After all the impact of each agent on $A(t)$ is of relative order $1/N$ and it should become "negligible" as $N \to \infty$. We shall see that this is not so. Naive agents update scores by

$$U_{s,i}(t+1) = U_{s,i}(t) - a_{s,i}^{\mu(t)} A(t) = U_{s,i}(t) - a_{s,i}^{\mu(t)} \sum_{j=1}^{N} a_{s_j(t),j}^{\mu(t)}. \tag{20}$$

Note that, at odds with Eq. (17), this depends on the action $a_{s_i(t),i}^{\mu(t)}$ which player i actually played through $A(t)$. Given the scores $U_{s,i}(t)$, Eq. (15) gives the probability $\pi_{s,i}(t)$ with which agent i shall choose her s^{th} strategy.

We can derive a continuum time limit equation for $\pi_{s,i}$ following the same approach as before. It is not difficult to check that this leads to the equation

$$\frac{d\pi_{s,i}}{dt} = -\Gamma_i \pi_{s,i} \sum_{j\in\mathcal{N}} \left[\overline{a_{s,i}(\vec{\pi}_j \cdot \vec{a}_j)} - \overline{(\vec{\pi}_i \cdot \vec{a}_i)(\vec{\pi}_j \cdot \vec{a}_j)}\right]. \tag{21}$$

This is different from RD (Eq. 12) because of the $j = i$ term in the sum. So $\vec{\pi}_i$ does not converge to a Nash equilibrium. Rather we can easily show that H is a Lyapunov function of this dynamics. Indeed

$$\frac{dH}{dt} = -2 \sum_{i\in\mathcal{N}} \Gamma_i \sum_{s=1}^{S} \pi_{s,i} \overline{\langle A\rangle (a_{s,i} - \vec{\pi}_i \cdot \vec{a}_i)}^2 < 0. \tag{22}$$

The dynamics converges therefore to the minima of H. It is easy to see that H is a non-negative quadratic form of the variables $\pi_{s,i}$ and therefore it attains its minimum on a connected subset $\mathcal{M} \in \Delta^{\mathcal{N}}$ [11]. We therefore conclude that *the long run dynamics of this system is described by the minimum of H.*

A complete statistical characterization of the minima of H in the limit $N \to \infty$ with $P/N = \alpha$ finite, can be obtained from the *replica method* a tool of statistical mechanics devised to deal with disordered systems. An account of this method is given in the appendix together with technical details on the calculation. Here we only discuss the results and their interpretation.

6.3.0.4 Exogenous information We first discuss exogenous information where $\varrho^\mu = 1/P$. We distinguish two regimes separated by a *phase transition* which occurs as $\alpha \to \alpha_c(S) \cong S/2 - 0.6626\ldots$.

For $\alpha > \alpha_c$ we find an *asymmetric phase*. Indeed $H > 0$ which means that $\langle A^\mu \rangle \neq 0$ at least for some $\mu \in \mathcal{P}$. The symmetry between the two actions in $\mathcal{A}$ is broken and a *best* strategy $a^\mu_{\mathrm{best}} = -\mathrm{sign}\,\langle A^\mu \rangle$ arises in $\mathcal{A}^{\mathcal{P}}$. An $N+1^{\mathrm{st}}$ agent who joined the game with this strategy would receive a payoff $\overline{|A|} - \overline{a^2_{\mathrm{best}}} = \overline{|A|} - 1$ [12].

The set $\mathcal{M}$ where H attains its minimum consists of a single point, so that, for any initial conditions, the dynamics converges to the same final state in the long run. In other words prior beliefs of agents about their strategies are irrelevant in the long run.

The asymmetry H decreases with decreasing α, which means increasing the number N of agents at fixed P. Naively speaking the asymmetry in $\langle A^\mu \rangle$ is exploited by the adaptive behavior of agents who then reduce it. Indeed agents are more and more selective in their choice, as shown by the fact that G increases as α decreases and the effective number of strategies used $1/G$ decreases. At the same time, as α decreases, the equilibrium becomes more and more “fragile” in the sense that its *susceptibility* to a generic perturbation increases (see the Appendix and ref. [16] for more precise statements).

As $\alpha \to \alpha_c$ the asymmetry vanishes $H \to 0$ and the response of the system to a generic perturbation diverges. This signals a phase transition to the symmetric phase $\alpha < \alpha_c$ where $H = 0$ and any perturbation can change dramatically the equilibrium. The set $\mathcal{M}$ where H attains its minimum is no more a single point, but rather an hyper-plane of a dimension which increases as α decreases. Any point in $\mathcal{M}$ is an equilibrium of Eqs. (21) and any displacement along

[11] Note that, on the contrary, σ^2 is not positive definite and it attains its minima, the Nash equilibria, on a non-connected subset of $\Delta^{\mathcal{N}}$.

[12] The term $-\overline{a^2_{\mathrm{best}}}$ is the “market” impact caused by the new agents. It arises because if the strategy a^μ_{best} where actually played by the $N+1^{\mathrm{th}}$ agent, that would also modify $A^\mu \to A^\mu + a^\mu_{\mathrm{best}}$. Ref. [25] discusses in greater detail these issues.

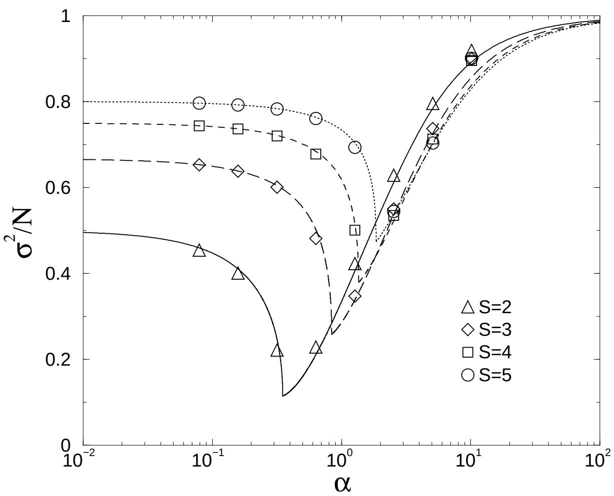

Fig. 2. Global efficiency σ^2/N as a function of α for $S = 2, 3, 4$ and 5 from computer experiments with $N = 101$ averaged over 100 realizations of $\vec{a}_i$ (symbols) and from the theoretical calculation (lines).

this set can occur freely. In particular, with different initial conditions, the system reaches different points of $\mathcal{M}$. The dynamics (21) indeed converges to the "closest" point on $\mathcal{M}$ which is on its trajectory. In other words, prior beliefs $U_{s,i}(0)$ are relevant for $\alpha < \alpha_c$: With different $U_{s,i}(0)$ the system reaches different equilibria.

These conclusions hold under the approximations who led us to Eq. (21). In the symmetric phase, these are only valid for $\Gamma_i \ll 1/P$ [15]. The reason is that Eq. (21) neglects time derivatives of order higher than the first one in

$$\frac{U_{s,i}(t+\Delta t) - U_{s,i}(t)}{\Delta t} = \frac{dU_{s,i}}{dt} + \frac{\Delta t}{2}\frac{d^2U_{s,i}}{dt^2} + \dots \tag{23}$$

These higher order terms become relevant in view of the dynamical degeneracy of the system below α_c and are responsible for the "crowd effect" [8,13,14] observed for $\Gamma_i \gg 1/P$ [24]: This effect [13] manifests in an increases of σ^2/N as α decreases, which is much faster than what predicted by our theory. This effect can be traced back to a dynamical anti-persistence [11] resulting from the fact that agents, neglecting their impact on $A(t)$, over-estimate the performance of the strategies they do not play and they keep switching from one to the other. Each time a particular information μ shows up, agents tend to do the opposite action of what they did the last time they saw the same information μ. Therefore the period of this dynamics is $2P$ [14]. In the present

[13] For the sake of simplicity, we assume, in the following discussion that Γ_i are all nearly equal for all$i \in \mathcal{N}$.

[14] The fact that agents do not realize that they have an impact on the aggregate in this case is probably unrealistic. From their point of view, the same strategy

context, this effect arises because the second order term in Eq. (23) induces a periodic orbits on the dynamically degenerate set $\mathcal{M}$. This behavior persists as long as $1/\Gamma_i$ is much smaller than the amplitude of the periodic oscillations of $U_{s,i}(t)$. As Γ_i decreases, Eq. (15) finally smoothes the oscillations in agents choices and the anti-persistent behavior disappears, as observed in [24].

As far as global efficiency is concerned, we find that σ^2/N increases with α towards the random agent limit as $1-\sigma^2/N \sim 1/\alpha$. This is shown in figure 2, which also shows that computer experiments for finite populations fully confirm our theoretical results. For fixed α, as S increases σ^2/N first decreases moderately as long as $\alpha > \alpha_c(S)$. Then the system enters the symmetric phase ($\alpha < \alpha_c$) because $\alpha_c(S)$ increases, and σ^2/N increases with S towards the random agent limit as $1-\sigma^2/N \sim 1/S$. We then conclude that allowing agents to have more strategies does not increase global efficiency as in the Nash equilibrium. Rather it pushes the system in the symmetric phase where σ^2 converges to the random agent limit as $S \to \infty$.

As long as the system is in the asymmetric phase, agents have incentives to consider more than S strategies, because that allows them to detect better the asymmetry. However if every agent enlarges the set of strategies she considers, i.e. if S increases, the system enters in the symmetric phase. Then global efficiency starts to decline. This behavior is reminiscent of the *Tragedy of the Commons* [27]: a situation where individual utility maximization by all agents leads to over-exploitation of common resources and poor payoffs.

6.3.0.5 Endogenous information With endogenous information the system behaves qualitatively in the same way, as first observed in ref. [10] by computer experiments. This is because the stationary state distribution ϱ^μ of the process $\mu(t)$ – which is induced by the dynamics of agents through Eq. (16) – is almost uniform on $\mathcal{P}$. Actually $\varrho^\mu = 1/P$ for $\alpha \le \alpha_c$ because of the symmetry of A^μ.

In order to measure the deviation of ϱ^μ from the uniform distribution $\varrho^\mu_{\mathrm{unif}} = 1/P$, we compute the entropy $\Sigma(P) = -\sum_{\mu\in\mathcal{P}} \varrho^\mu \log_P \varrho^\mu$. With the choice of base P for the logarithm $\Sigma(P) = 1$ for $\varrho^\mu = 1/P$ so that $1-\Sigma(P)$ is a reasonable measure of the deviation of ϱ^μ from a uniform distribution. In figure 3, $1-\Sigma(P)$ is plotted for several values of P as a function of α. While for $\alpha < \alpha_c$ we find $\Sigma(P) = 1$ to a great accuracy, for $\alpha > \alpha_c$ numerical results suggest that $\Sigma(P) \to 1$ as $P = \alpha N \to \infty$. On this basis, we conclude as in ref. [10], that the MG with endogenous information gives the same results as

which had a good score of performance when they were not using it, starts performing badly as soon as they use it. This could either be considered a manifestation of *Murphy's law* or the fact that agents rationality is bounded below the level of common sense.

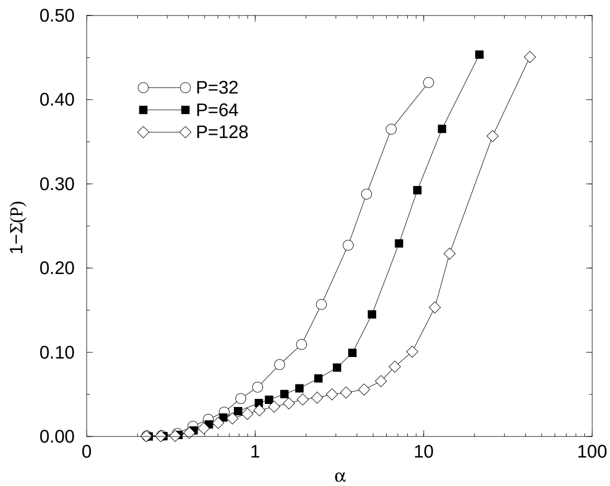

Fig. 3. Deviation of the distribution ϱ^μ from the uniform one from numerical simulations.

the MG with exogenous information. A detailed study of the dynamics of the process $\mu(t)$, to give a deeper theoretical foundation to this conclusion, shall be pursued elsewhere.

7 Discussion

There is a growing literature on learning which addresses the issue of which learning procedure (modeling inductive rationality) may eventually lead to deductive rational outcomes[4,21]. The choice made in the El Farol bar problem and in the MG – which is exponential learning (Eq. 15) eventually with $\Gamma_i \to \infty$ – is one of these as we have shown. What leads to equilibria different from Nash equilibria is the fact that agents *i)* have not full information on the effects of their strategies and *ii)* that they neglect their impact on the aggregate in the evaluation of their strategies.

This new equilibrium, which we call *naive agents equilibrium* (NAE) differs substantially from a Nash equilibrium (NE) because:

(1) In a NE global efficiency always increases as the number N of agents increases (with P fixed), whereas in the NAE it only decreases as far as $N < P/\alpha_c$ and then it increases in a way which depends on initial conditions (prior beliefs) and on the parameters Γ_i.
(2) There are (exponentially) many *disconnected* Nash equilibria which are selected by initial conditions. For $\alpha > \alpha_c$ there is a unique NAE and, for all initial conditions, the system converges to it. For $\alpha < \alpha_c$ there is a continuum of NAE, but they are all connected.
(3) Global efficiency (σ^2), for fixed α, always decreases as agents resources (S)

increase, and it eventually converges to perfect optimization for $S \to \infty$. In the NAE efficiency only mildly improves with S in the asymmetric phase. But increasing S also increases $\alpha_c(S)$ and when $\alpha_c(S) > \alpha$ the system enters into the symmetric phase where σ^2 increases with S towards the random agent limit (this occurs for $\Gamma_i \sim 1/P$ [24]).

(4) In the NE agents play *pure* strategies, i.e. $\vec{\pi}_i$ is a singleton $\forall i \in \mathcal{N}$. Indeed, for fixed opponent strategies s_{-i}, each agent typically [15] has a pure strategy s_i – the best response – which is superior to others. In the NAE, agents mix this strategy with others because, neglecting their impact on $A(t)$, they over-estimate the performance of the strategies they do not play [16]. Playing a strategy reduces its perceived performance and this is why agents mix strategies in the NAE (see ref. [25] for a deeper discussion). The probability $\pi_{s,i}$ in the NAE is such that the perceived performance of all strategies which are played ($\pi_{s,i} > 0$) is the same.

(5) A consequence of the previous point is that, the origin of information is quite important for inductive agents with full information, while it is irrelevant in the NAE [10]. Inductive agents with full information lead, under *endogenous* information, to a deterministic dynamics of $\mu(t)$ and only a small subset of informations μ is ever visited. This in turn leads to a much more efficient coordination. Naive agents with endogenous information, on the other hand, induce a dynamics on $\mu(t)$ which is "ergodic", i.e. which visits each information μ with nearly the same frequency ϱ^μ. Therefore the collective behavior is the same as that of the NAE with exogenous information.

Our results clearly allow for several extensions, as those of ref. [25]. It also suggests a theoretical approach to the El Farol problem [1]. The key problem lies in the parametrization of forecasting rules: agents in the El Farol problem consider the record of past attendance to the bar – i.e. $A(t')$ for $t' < t$ – whereas in the minority game agents only consider the record of the sign of $A(t')$. Focusing on the last M games, there can be N^M possible records in the El Farol problem, instead of 2^M. This causes no problem in principle since one can take $P = N^M$. In practice however forecasting rules have to be "reasonable": For example a randomly drawn rule can easily predict a different outcome if the attendance of one of the past weeks just changes by one unit. Some sort of continuity in strategies should be introduced so that "similar" histories (=information μ) lead to similar forecasts. Even though it is not clear how to translate this requirement mathematically (one way could

[15] Typically here means that we disregard unlikely realizations where two pure strategies happen to yield the same payoffs.

[16] In order to see this, take the expectation of Eq. (21) over μ and consider the cross term involving i^{th} strategies $-\overline{a_{s,i}a_{s_i,i}}$, where s_i is the strategy actually played by i. When $s \neq s_i$ this term is small ($\overline{a_{s,i}a_{s_i,i}} \sim 1/\sqrt{P}$), whereas for $s = s_i$ it is not: $\overline{a_{s_i,i}a_{s_i,i}} = 1$.

be that followed in ref. [28]) it is clear that it suggests that the number of relevant informations P is much less than N^M.

We expect that the distinction between inductive agents with full information and naive agents to be of key importance also in the El Farol problem and we believe that, eventually, an analytical solution in the limit $N \to \infty$ is possible along the same lines followed here.

A Brief review of statistical mechanics of optimization problems

As we have seen the typical behavior of agents in the Minority Game can be recast as the problem of finding the probability distribution of mixed strategies which optimizes (minimizes) a global Lyapunov function. This is not the general case in games where agents optimize their own utility function. However for inductive learning behavior of agents, like in the MG, it may happen that there exist a global quantity, an *energy* of the system, which is bounded below and which decreases as the system evolves.

Once such an energy function is defined we may hope to be able to apply all the methods and ideas developed in the context of statistical physics just by thinking to the optimization problem as to the problem of studying its zero temperature properties.

Notice that for games in which such a global energy function does not exists (indeed the most general case!) one could still hope to be able to apply the techniques developed in the context of off-equilibrium statistical mechanics. This issue will be discusses elsewhere.

There is another advantage in resorting to the statistical mechanics approach. The theory of random systems and the methods developed for their study, allow to model highly complex (deterministic) systems as properly generated random ones and to study their most probable outcome. This has been done with success for many complex physical systems and is what we have been doing with the Minority Game.

In order to give a brief and intuitive description of the connections between statistical mechanics and optimization, let us recall some basic definitions and concepts (that will be heavily used in the calculations for the MG described in Appendix B).

(1) Consider a finite set $\mathcal{C}$ of cardinality Ω including elements $C_1, C_2, \ldots, C_\Omega$ hereafter referred to as *configurations* of the system.
(2) Consider a finite set $\mathcal{E}$ of cardinality $L \le \Omega$ including real-valued elements

$E_1 < E_2 < \ldots < E_L$ hereafter referred to as *energy levels*.

(3) Consider a function $E \ : \mathcal{C} \to \mathcal{E}$. $E(C_i)$, $i = 1, \ldots, \Omega$ is called *energy* of the i-th configuration.

(4) Call g_l the *multiplicity* of the l-th energy level, that is the number of configurations C_i having energy $E(C_i) = E_l$.

(5) Let β be a real parameter called *inverse temperature*.

(6) The *partition function* associated to function E at inverse temperature β is *defined* as

$$Z = \sum_{C \in \mathcal{C}} e^{-\beta E(C)} \qquad . \tag{A.1}$$

By noticing that expression (A.1) may be rewritten as

$$Z = \sum_{l=1}^{L} g_l \, e^{-\beta E_l} \qquad , \tag{A.2}$$

it follows that Z may also be seen as the generating function of the multiplicities g_l.

Such multiplicities have a straightforward combinatorial interpretation. Indeed g_l is the number of solutions to the equation

$$E(C) = E_l \qquad \text{for} \;\; C \in \mathcal{C}. \tag{A.3}$$

Statistical mechanics deals with the study and calculation of the partition function Z, or of the so called *free energy*

$$\mathcal{F} = -\frac{1}{\beta} \ln Z \quad . \tag{A.4}$$

Once Z or $\mathcal{F}$ have been computed, one can extract information about the multiplicities of the energy levels by varying the parameter β. Of particular relevance is the infinite β limit, in that it gives access to the minimal energy level, the so called ground state. One can easily check that

$$\begin{aligned} E_1 &= \lim_{\beta \to \infty} -\frac{\partial \ln Z}{\partial \beta} \\ \ln g_1 &= \lim_{\beta \to \infty} -\beta^2 \frac{\partial (\beta \ln Z)}{\partial \beta} \qquad . \end{aligned} \tag{A.5}$$

Therefore, from the knowledge of $\ln Z$ one directly derives the value of the minimal cost and of the number of solution to the optimization problem of minimizing the function E over $\mathcal{C}$. The logarithm of the number of such optimal solutions ($\ln g_1$), is called the *entropy* of the ground state.

The statistical mechanics formalism can be directly generalized to continuous variable, the sum over configurations becoming a multiple integral over some domain.

In order to extract the typical behavior of a system which depends on some random parameters we need to average the free energy. For sufficiently simple models, this can be done resorting to the so called replica method. We discuss this issue directly by describing the details of the calculations for the MG.

B Replica calculation for the MG

Our goal is to compute and characterize the minimum of $H = \overline{\langle A\rangle^2}$ on the simplex $\Delta^{\mathcal{N}} = \{\vec{\pi}_i,\ i \in \mathcal{N}\}$. Considering H as an Hamiltonian of a statistical mechanic's system, this can be done analyzing the zero temperature limit. First we build the partition function

$$Z(\beta) = \mathrm{Tr}_\pi e^{-\beta H\{\pi\}}, \tag{B.1}$$

where β is the inverse temperature and Tr_π stands for an integral on $\Delta^{\mathcal{N}}$ (we call simply π an element of $\Delta^{\mathcal{N}}$). The quantity of interest is then

$$\min_{\pi\in\Delta^{\mathcal{N}}} H\{\pi\} = -\lim_{\beta\to\infty} \beta^{-1} \ln Z(\beta). \tag{B.2}$$

This in principle depends on the specific realization $a^\mu_{s,i}$ of rules chosen by agents. In practice however, to leading order in N, all realizations of $a^\mu_{s,i}$ yield the same limit, which then coincides with the average of $\min_{\pi\in\Delta^{\mathcal{N}}} H\{\pi\}$ over $a^\mu_{s,i}$. The average of $\ln Z$ over the a's, which we denote by $\langle\ldots\rangle_a$, is reduced to that of moments of Z using the replica trick[16]:

$$\langle \ln Z\rangle_a = \lim_{n\to 0} \frac{1}{n} \ln \langle Z^n\rangle_a \tag{B.3}$$

With integer n the calculation of $\langle Z^n\rangle_a$ amounts to study n *replicas* of the the same system with the same realization of $a^\mu_{s,i}$. To do this we introduce a set of dynamical variables $\pi_a \equiv \{\pi_{s,i,a}\}$ for each replica, which are labeled by the additional index $a = 1, \ldots, n$. Each replica has its corresponding Hamiltonian, which we write as $H_a\{\pi_a\} = \overline{A_a^2}$ where $A^\mu_a = \sum_{i\in\mathcal{N}} \vec{\pi}_{i,a} \cdot \vec{a}_i$. The set of all dynamical variables for all replicas is the direct product $\Delta^{\mathcal{N}^n}$ of n phase spaces $\Delta^{\mathcal{N}}$. In order to compute the limit $n \to 0$ in Eq. (B.3) one appeals to analytic continuation of $\langle Z^n \rangle_a$ for real n. We give here the details of the calculation in our specific case. More details on the nature of the method can be found in ref. [16]. We write

$$\langle Z^n \rangle = \prod_{a=1}^{n} \prod_{\mu \in \mathcal{P}} \mathrm{Tr}_\pi \left\langle e^{-\beta \varrho^\mu (A_a^\mu)^2} \right\rangle_a = \prod_{\mu \in \mathcal{P}} \prod_{a=1}^{n} E_{z_a^\mu} \left[\mathrm{Tr}_\pi \left\langle e^{-i\sqrt{2\beta\varrho^\mu} z_a^\mu A_a^\mu} \right\rangle_a \right] \quad \text{(B.4)}$$

where $E_z[\dots]$ stands for the expectation over the Gaussian variable (unit variance an zero mean) z and we have introduced one such variable z_a^μ for each a and μ and used the identity $E_z[e^{-ixz}] = e^{-x^2/2}$. In addition we used the shorthand Tr_π for the integral over $\Delta^{\mathcal{N}n}$. The average over $a_{s,i}^\mu$ now factorizes

$$\prod_{a=1}^{n} \left\langle e^{-i\sqrt{2\beta\varrho^\mu} z_a^\mu A_a^\mu} \right\rangle_a = \prod_{i \in \mathcal{N}} \prod_{s=1}^{S} \left\langle e^{-i\sqrt{2\beta\varrho^\mu} (\sum_a z_a^\mu \pi_{s,i}^a) a_{s,i}^\mu} \right\rangle_a =$$

and we can explicitly compute it using the distribution (2). This gives

$$= \prod_{i \in \mathcal{N}} \prod_{s=1}^{S} \cos \left[\sqrt{2\beta\varrho^\mu} \sum_{a=1}^{n} z_a^\mu \pi_{s,i}^a \right] \simeq \prod_{i \in \mathcal{N}} \exp \left[-\beta\varrho^\mu \sum_{a,b=1}^{n} z_a^\mu z_b^\mu \sum_{i \in \mathcal{N}} \vec{\pi}_i^a \cdot \vec{\pi}_i^b \right].$$

In the last passage we used the relation $\cos x \simeq e^{-x^2/2}$ which is correct to order x^2 in a power expansion. This is justified as long as $\varrho^\mu \to 0$ as $P = \alpha N \to \infty$ for each $\mu \in \mathcal{P}$. Before going back to Eq. (B.4), we introduce the matrices $\hat{G} \equiv \{G_{a,b},\ a,b = 1,\dots,n\}$ and $\hat{r} \equiv \{r_{a,b},\ a,b = 1,\dots,n\}$ through the identities

$$1 = \int dG_{a,b} \delta \left(G_{a,b} - \frac{1}{N} \sum_{i \in \mathcal{N}} \vec{\pi}_i^a \cdot \vec{\pi}_i^b \right) \propto \int dr_{a,b} dG_{a,b} e^{\frac{\alpha\beta^2 r_{a,b}}{2} (\sum_i \vec{\pi}_i^a \cdot \vec{\pi}_i^b - N G_{a,b})}$$

for all $a \geq b$, where $\delta(x)$ is Dirac's delta function and we used its integral representation. The only part depending on the $\pi_{s,i}^a$ in $\langle Z^n \rangle$ is $e^{\alpha\beta^2 \sum_{a \geq b} r_{a,b} \sum_i \vec{\pi}_i^a \cdot \vec{\pi}_i^b /2}$. This can be factorized in the agent's index i and so the integral Tr_π on $\Delta^{\mathcal{N}n}$ can be factorized into N integrals over Δ^n (=the direct product of the simplexes of the n replicas of the same agent's mixed strategies). With this we can write

$$\langle Z^n \rangle = \int dr_{a,b} dG_{a,b} e^{-\beta n N F(\hat{G}, \hat{r})} \quad \text{(B.5)}$$

where, specializing to the case $\varrho^\mu = 1/P$ [17],

[17] A generic distribution ϱ^μ can also be handled, though with heavier notations.

$$F(\hat{G},\hat{r}) = \frac{\alpha}{2n\beta}\ln\det\left[\hat{I}+\frac{2\beta}{\alpha}\hat{G}\right] + \frac{\alpha\beta}{2n}\sum_{a,b} r_{a,b}G_{a,b}$$
$$-\frac{1}{n\beta}\ln \mathrm{Tr}_{\pi\in\Delta^n}\exp\left[\frac{\alpha\beta^2}{2}\sum_{a,b} r_{a,b}\vec{\pi}^a\vec{\pi}^b\right], \tag{B.6}$$

where $\hat{I}$ is the identity matrix. The first term arises from the expectation over z_a^μ. This factorizes for each μ and one is left with a Gaussian integral over $\vec{z}\in I\!R^n$. The second and the third terms arise from the integral representation of the delta functions [18].

The key point is that, in the limit $N\to\infty$ the integral over the matrices $\hat{r}$ and $\hat{G}$ in Eq. (B.5) are dominated by their saddle point value, i.e. by the values of $r_{a,b}$ and $G_{a,b}$ for which F attains its minimum value [19]. One should then study the first order conditions $\partial F/\partial r_{a,b}=0$ and $\partial F/\partial G_{a,b}=0$. In this particular case, F attains its minimum for *replica symmetric* matrices of the form

$$G_{a,b}=g+(G-g)\delta_{a,b}, \qquad r_{a,b}=r+(R-r)\delta_{a,b}. \tag{B.7}$$

The reason for this is related to the fact that the *energy landscape* is very simple in this case because H is a non-negative definite quadratic form in $\Delta^{\mathcal{N}}$ (see ref. [16] for a deeper discussion). Taking the limit $n\to 0$, Eq. (B.3) then gives

$$\frac{\langle \ln Z(\beta)\rangle_a}{N} = \frac{\alpha g}{\alpha+2\beta(G-g)} + \frac{\alpha}{2\beta}\ln\left[1+\frac{2\beta(G-g)}{\alpha}\right] + \frac{\alpha\beta}{2}(RG-rg)$$
$$-\frac{1}{\beta}E_{\vec{z}}\{\ln \mathrm{Tr}_\pi \exp\left[-\beta V_{\vec{z}}(\vec{\pi})\right]\} + \epsilon G \tag{B.8}$$

where Tr_π is now the integral over the simplex Δ of a single agent's mixed strategies and we defined, for convenience, the potential $V_{\vec{z}}(\vec{\pi}) = \sqrt{\alpha r}\,\vec{z}\cdot\vec{\pi} - \frac{\alpha}{2}\beta(R-r)|\pi|^2$. We have also added a new term ϵG to the *free energy* for future use. This is obtained if we start, right from the beginning, by the Hamiltonian $H+\epsilon N G$ with G given by Eq. (9). Eventually we shall restore the original problem in the limit $\epsilon\to 0$. The parameters g, G, r and R are fixed by the first order conditions $\partial F/\partial g=0$, $\partial F/\partial G=0$, $\partial F/\partial r=0$ and $\partial F/\partial R=0$. These equations, finally, have to be studied in the limit $\beta\to\infty$, where one recovers the minimum of H by Eq. (B.2).

[18] For simplicity we have also done the transformation $r_{a,b}\to r_{a,b}/2$ for $a\neq b$ so that $\sum_{a\geq b}\to\sum_{a,b}$.
[19] Note that, by Eq. (B.2), we shall also be interested in the limit of $\beta\to\infty$ in the end!

It is convenient to define the parameters

$$\chi = \frac{2\beta(G-g)}{\alpha}, \qquad y = \frac{\sqrt{g/\alpha}}{1-\epsilon(1+\chi)} \tag{B.9}$$

In the limit $\beta \to \infty$, we are going to find solutions where $g \to G$ and χ, which we call *susceptibility*, remains finite. This implies that two replicas of the same system converge in the long run to the same stationary state. Using the saddle point equations, and $g = G$, we can rewrite

$$V_{\vec{z}}(\vec{\pi}) = \frac{2y\,\vec{z}\cdot\vec{\pi} + \pi^2}{1+\chi}, \qquad \beta = \infty \tag{B.10}$$

The last term in Eq. (B.8) is dominated by the mixed strategy $\vec{\pi}^*(\vec{z})$ which is the solution of

$$\vec{\pi}^*(\vec{z}) = \arg\min_{\pi\in\Delta} V_{\vec{z}}(\vec{\pi}). \tag{B.11}$$

We find that $G = g = E_{\vec{z}}[\vec{\pi}^*(\vec{z})]$, which is then a function of y only $G \equiv G(y)$. Upon defining $\zeta(y) = E_{\vec{z}}[\vec{z}\cdot\vec{\pi}^*(\vec{z})]$, we find

$$\chi(y) = -\frac{\zeta(y)}{\sqrt{\alpha G(y)} + \zeta(y)}$$

The second of Eqs. (B.9) becomes an equation for y as a function of α which has two implicit solutions

$$(1-\epsilon)^2\alpha y^2 = G(y), \qquad (1-\epsilon)^2\alpha G(y) = \zeta^2(y), \tag{B.12}$$

where now we can set $\epsilon = 0$. The *free energy* per agent is

$$\lim_{N\to\infty}\frac{H}{N} = \lim_{N\to\infty}\frac{\langle \ln Z(\beta)\rangle_a}{N} = \frac{G}{(1+\chi)^2} \tag{B.13}$$

These equations are transcendental and we could not find an explicit solution for generic S. Nevertheless, they represent a great simplification with respect to the original problem. The main technical difficulty lies in the evaluation of the functions $G(y) = E_{\vec{z}}[|\pi^*(\vec{z})|^2]$ and $\zeta(y) = E_{\vec{z}}[\vec{z}\cdot\vec{\pi}^*(\vec{z})]$, which can be computed numerically to any desired accuracy $\forall S$.

The first of Eqs. (B.12) gives the $\alpha > \alpha_c$ phase. One recovers this solution also if one starts from Eq. (B.8) with $\epsilon = 0$. This solution has $\chi > 0$ finite and

$H > 0$ non-zero. As α decreases χ increases and it diverges as $|\alpha - \alpha_c|^{-1}$ when $\alpha \to \alpha_c^+$. In this limit Eq. (B.13) implies that $H \sim |\alpha - \alpha_c|^2$ vanishes. The critical point $\alpha_c = \alpha(y_c)$ is obtained imposing $\chi = \infty$, which gives $G(y_c) = -y_c\zeta(y_c)$. By the numerical evaluation of the functions $G(y)$ and $\zeta(y)$, we find

$$\alpha_c(S) \cong \alpha_c(2) + \frac{S-2}{2} \tag{B.14}$$

to a high degree of accuracy. It might be that this equation is exact but we could not prove it. An interesting relation for $\alpha_c(S)$ can be derived by algebraic considerations: Note that for each $\pi_{s,i} > 0$ the equation

$$\frac{\partial H}{\partial \pi_{s,i}} = 2\sum_{j,s'} \overline{a_{s,i}a_{s',j}}\pi_{s',j} = 0 \tag{B.15}$$

must hold. This is a set of linear equations in the variables $\pi_{s,i} > 0$. The $NS \times NS$ matrix $\overline{a_{s,i}a_{s',j}}$ is built with P dimensional vectors $a_{s,i}^{\mu}$ and therefore has at most rank P. In other words there are only P independent equations (B.15). In addition there are N normalization conditions on $\pi_{s,i}$. The system becomes dynamically degenerate when the number of free variables $\pi_{s,i}$ becomes bigger than the number $P+N$ of independent equations and, exactly at α_c the two are equal. Dividing this condition by N gives the desired equation

$$\sum_{s=1}^{S} E_{\vec{z}}\{\theta[\pi_s^*(\vec{z})]\} = \alpha_c(S) + 1. \tag{B.16}$$

The left hand side is the average number of strategies used by agents. Note that this equation implies that $\alpha_c(S)$ cannot grow faster than linear in S. Also $\alpha_c(S) \propto S/2$ imply that agents use on average $1/2$ of their strategies at α_c.

The second of Eqs. (B.12) gives the $\alpha < \alpha_c$ phase. Note indeed that with this choice $\chi \simeq 1/\epsilon \to \infty$ and $H \sim \epsilon^2 \to 0$ as $\epsilon \to 0$. At odds with the solution for $\alpha > \alpha_c$, this equation only arises if $\epsilon > 0$ is sent to zero only at the end of the calculation, in Eq. (B.12). With $\epsilon = 0$ the saddle point equations have only a solution with $G > g$ in the limit $\beta \to \infty$. This is because for $\alpha < \alpha_c$ the set $\mathcal{M}$ where $H = 0$ is not a single point, which means that the equilibrium is not unique. The replica method with $\epsilon = 0$ takes an average on all the set $\mathcal{M}$ and so it gives results which are not representative of a particular system [20]. The term $\epsilon\, N\, G$ in H, which results in the last term $\epsilon\, G$ in Eq. (B.8), is introduced to break this "degeneracy" of equilibria and to select only one of them. With this particular choice we indeed took the equilibrium which is closest to the

[20] Note indeed that g has the interpretation of the overlap between two replicas of the same system, so that $g < G$ means that the two replicas are not identical.

random initial condition $\pi_{s,i}(0) = 1/S$ for all $i \in \mathcal{N}$ and $s = 1, \ldots, S$, which assumes no prior beliefs of agents ($U_{s,i}(t = 0) = 0$, $\forall s, i$). A different state could be singled out by a different choice of the term proportional to ϵ. Once the degeneracy is broken and an equilibrium is singled out, we can safely set $\epsilon = 0$.

In both phases, once the saddle point equations are solved, one can derive the full statistical characterization of the system. For instance the fraction of agents playing a strategies in a neighborhood $d\vec{\pi}$ of $\vec{\pi}$ is given by $p(\vec{\pi})d\vec{\pi} = E_{\vec{z}}[\delta(\vec{\pi}^*(\vec{z}) - \vec{\pi})]d\vec{\pi}$.

References

[1] Arthur W. B., *Inductive reasoning and bounded rationality*, Am. Econ. Assoc. Papers and Proc **84**, 406, 1994.

[2] Myerson, R.B., *Game Theory: analysis of conflict*, Harvard Univ. Press (1997).

[3] Erev I. and A.E. Roth, *Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria*, Am. Ec. Rev. **88**, 848 (1998).

[4] D. Fudenberg and D. K. Levine, *The theory of learning in games* (MIT Press, 1998).

[5] Casti J. L. *Seeing the light at El Farol*, Complexity (**1**)5, 7 (1996).

[6] Challet D. and Zhang Y.-C., *Emergence of cooperation and organization in an evolutionary game*, Physica A **246**, 407 (1997).

[7] Y.-C. Zhang, *Toward a Theory of Marginally Efficient Markets*, Physica A **269**, 30 (1999).

[8] Savit R., Manuca R., and Riolo R., *Adaptative competition, market efficiency and phase transition*, Phys. Rev. Lett., **82**, 2203 (1990).

[9] Johnson N.F. et al.,*Volatility and Agent Adaptability in a Self-Organizing Market*, Physica A (**258**), 230 (1998).

[10] A. Cavagna, Irrelevance of Memory in the Minority Game, Phys. Rev. E **59**, R3783 (1999), preprint cond-mat/9812215

[11] Challet D. and M. Marsili, *Phase Transition and Symmetry Breaking in the Minority Game*, preprint cond-mat/9904071 (1999).

[12] For a collection of papers and preprints on the minority game see the web site http://www.unifr.ch/econophysics/.

[13] Challet D. and Zhang Y.-C., *On the Minority Game : analytical and numerical studies*, Physica A **256**, 514 (1998).

[14] Johnson N.F. et al., *Crowds Effects and Volatility in a Competitive Market*, Physica A **269**, (1999).

[15] Challet D., M. Marsili and R. Zecchina, *Statistical Mechanics of Systems with Heterogeneous Agents*, preprint cond-mat/9904392 (1999).

[16] M. Mezard, G. Parisi, M. A. Virasoro, *Spin glass theory and beyond* (World Scientific, 1987).

[17] Johnson N.F. et al., *Minority Game with Arbitrary Cutoffs*, preprint cond-mat/9903228 (1999).

[18] M. Woodford, *Learning to believe in sun-spots*, Econometrica **58**, 277 (1990).

[19] Weibull, J.W., (1995), *Evolutionary Game Theory*, The MIT Press, (Cambridge, Massachusetts; London, UK.

[20] A. De Martino, M. Marsili and R. Zecchina forthcoming.

[21] A. Rustichini, *Optimal properties of Stimulus-Response Learning Models*, Games and Economic Behavior (forthcoming, 1999).

[22] Luce R.D., *Individual Choice Behavior: A Theoretical Analysis*, New York, Wiley (1959).

[23] McKelvey, R.D. and T.R. Palfrey, *Quantal response equilibria for normal form games*, Games Econ. Behav. **10**, 6 (1995).

[24] A. Cavagna, J.P. Garrahan, I. Giardina, D. Sherrington, *A thermal model for adaptive competition in a market*, preprint cond-mat/9903415 (1999).

[25] Challet D., M. Marsili and Y.-C. Zhang *Modeling Market Mechanisms*, preprint (1999).

[26] Franke R., *Reinforcement learning in the El Farol problem*, preprint (March 1999).

[27] Hardin, G., *Science* **162**, 1243 (1968).

[28] G. Caldarelli, M. Marsili and Y.-C. Zhang, *A Prototype Model for Stock Exchange Markets*, Europhys. Lett. **40**, 479 (1997).