

January 17, 2001

Adaptive Learning and Emergent Coordination in Minority Games¹

G. Bottazzi**, G. Devetag***, G. Dosi**²

**S. Anna School of Advanced Studies
via G. Carducci, 40 Pisa, Italy

***Department of Computer and Management Sciences
University of Trento
Via Inama, 5 - 38100 Trento, Italy.

Abstract

The work studies the properties of a coordination game in which agents repeatedly compete to be in the population minority. The game reflects some essential features of those economic situations in which positive rewards are assigned to individuals who behave in opposition to the modal behavior in a population. Here we model a group of heterogeneous agents who adaptively learn and we investigate the transient and long-run aggregate properties of the system in terms of both allocative and informational efficiency. Our results show that, first, the system long-run properties strongly depend on the behavioral learning rules adopted, and, second, adding noise at the individual decision level and hence increasing heterogeneity in the population substantially improve aggregate welfare, although at the expense of a longer adjustment phase. In fact, the system achieves in that way a higher level of efficiency compared to that attainable by perfectly rational and completely informed agents.

JEL codes: G1,D8,C6

Keywords: minority game, coordination, speculation, adaptive learning, market efficiency, emergent properties.

¹Comments on a previous draft by Cars Hommes and Nicolas Jonard are gratefully acknowledged.

²Corresponding author.

1 Introduction

This work investigates the dynamic properties of *minority games* under different forms of heterogeneity amongst agents and different learning procedures.

In the game, first introduced by Challet and Zhang (1997, 1998), a population of N agents (where N is an odd number) must each simultaneously and independently choose between two sides, say 0 and 1. The side chosen by the minority of the agents is the winning one - each gaining a fixed positive reward -, while all members on the majority side obtain a null payoff.

Formally, the model is an N -person coordination game with multiple asymmetric Nash equilibria in pure strategies, and a unique symmetric mixed strategy equilibrium. In the following, we assume a large population, and we investigate the conditions under which repeated interaction amongst players causes some form of aggregate self-organization to emerge.

The game, notwithstanding its extreme simplicity, does capture some basic features of quite a few processes of social and economic interactions whereby orderly coordination rests upon the existence or the emergence of some behavioral heterogeneity in the relevant population.

Some coordination problems of this kind have long been discussed in Schelling's seminal work on those patterns of "sorting and mixing" where order at an aggregate level might find its roots in persistent micro adjustments within heterogeneous populations (Schelling, 1978). They also bear a close resemblance to the "El Farol" Problem devised by Brian Arthur (1994). In this game, a number of agents must independently decide each night whether or not to attend a bar called "El Farol". Each agent receives a positive utility from attending the bar as long as this is not too crowded. Otherwise he prefers to stay home. Somewhat similar coordination problems are involved in the experimental market entry games [e.g., Ochs (1990), Meyer et al. (1992); Ochs (1995) and references therein] where a group of N players must decide at each stage whether or not to enter one or more markets each having a fixed capacity k , with $k < N$.

Finally, the model captures some basic features of speculation on financial markets - which has been indeed the primary concern of minority-game modeling so far. A classic reference in this respect is to Keynes' "Beauty Contest" metaphor, where the payoff to an individual player does not stem from the accuracy of the appreciation of the intrinsic beauty of various contest candidates but rather from guessing the guesses of the ensemble of the other evaluators.

Strictly speaking, the beauty contest metaphor crucially involves *positive feedback* investment strategies - in the language of technical trading in finance - since the payoff is based on a majority rule, while minority games (MG, hereafter) fundamentally address those aspects of speculation dynamics involving activities of arbitrage against average market behaviors. These *negative feedback* strategies characterize agents trying to infer some - actual or imagined - structure in the history of collective interactions and, through that, trying to "beat the market" - that is, "beat the majority view" - by arbitraging against it. In the literature on chartist rules, this roughly corresponds to *contrarian strategies* [e.g., in the behavioral finance literature, De Bondt and Thaler (1995)].

In this work we shall precisely address the collective outcomes of such interactions in the MG setup. Are orderly aggregate coordination and absence of arbitrage opportunities generic properties of minority-type market processes, independently of any further specification of microeconomic behaviors? Or, conversely, do finer details of the ecology of agents populations (such as their sheer size) and their decision processes (e.g., the number and degree of rational sophistication of their decision

rules, their amount of memory, etc.) affect collective adjustment processes and long term outcomes?

Previous results on the MG based on adaptive agents with deterministic decision rules [e.g., Challet and Zhang (1997, 1998); Savit et al. (1998)] do suggest that diverse degrees of market efficiency fundamentally depend upon some underlying heterogeneity in the population. Here we explore the more general case whereby heterogeneous agents, endowed with memories of variable length, are allowed to adaptively learn. Interestingly, first, efficient coordination turns out robustly to be an emergent property resting on an ecology of diverse agents who do not play Nash equilibrium strategies. Second, we show that collective efficiency is not monotonic either in the rational sophistication of the agents nor in the information they are able to access. Rather, again, it crucially depends on the ecology of behaviors over the population.

In Section 2 we briefly illustrate the basic features of the game, study the properties of its equilibria and introduce a notion of *adaptive strategy*. Section 3 discusses the role of heterogeneity for collective dynamics under the assumption of strategies over the population and adaptive deterministic behaviors. In Section 4 we introduce a probabilistic learning model and study both the transient and limit properties of the dynamics. In particular, we analyze the *allocative* and *informational* efficiencies - which we shall define below - of the system under different degrees of computational complexity of the purported agents and different parameterizations of the learning rule. In general, our results suggest that it is heterogeneity in the population rather than individual computational abilities which mainly accounts for the observed aggregate properties.

2 The Minority Game

2.1 The Baseline Game-theoretic Framework

The minority game is played by a group N of players, where N must be an odd number. On each period of the stage game, each player must choose privately and independently between two actions or sides, say 0 and 1. The payoff π_i for $i \in \{0, 1\}$ is the same for all N players and is equal to

$$\pi_i = \begin{cases} 1 & \text{if } n_i \leq (N - 1)/2 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where n_i is the number of players choosing side i . Each player is rewarded with a unitary payoff whenever the side he chooses happens to be chosen by the minority of the players, while players on the majority side get nil.

It is easy to see that the game has $\binom{N}{(N-1)/2}$ asymmetric Nash equilibria in pure strategies, in which exactly $(N - 1)/2$ players choose either one of the two sides. Clearly, under pure-strategy equilibrium play, the payoffs for the two “parties” are quite different. Players belonging to the minority side are rewarded a fixed positive payoff, while those on the majority side earn nothing. The pure strategy equilibria, hence, are not strict, because players on the majority side are just indifferent between sticking to equilibrium play and deviating. The game also presents a unique symmetric mixed-strategy Nash equilibrium, in which each player selects the two sides with equal probability¹.

¹Of course, there are infinite asymmetric mixed-strategy equilibria.

The Minority Game is a simple game of coordination, in which, however, the payoff asymmetry implied by the pure strategy Nash equilibria is likely to rule out simple precedent-based solutions of the underlying coordination problem. The goodness of the achieved coordination can be easily measured by the average size of the winning minority. The further this size is from $(N - 1)/2$, the higher is the amount of money which is, so to speak, left on the table, hence the lower the resulting aggregate welfare. Note that under mixed strategy equilibrium play, individual expected payoff is equal to .5 on each period of the stage game, and the group payoff follows a binomial distribution with mean equal to $N/2$ and a variance of $N/4$. The measure of variance is indeed an equivalent measure of the degree of efficiency achieved in a population. The higher the variance, the higher the magnitude of fluctuations around the mixed strategy Nash equilibrium and the corresponding aggregate welfare loss.

2.2 The Behavioral Foundations of the Minority Game: Adding Beliefs

Let us start by discussing the nature of the game equilibria when players choose their actions on the basis of idiosyncratic (i.e. agent-specific) beliefs about other players' aggregate behavior, and try to exploit the latter to their advantage. This perspective, which amounts to consider agents endowed with inductive rather than deductive rationality, has been adopted in the minority literature including the El Farol model (Arthur, 1994) and it is common to multi-agent accounts of financial markets [e.g., Arthur et al. (1997); Brock and Hommes (1998); LeBaron (2000) and references therein].

One way to model beliefs in this setting, which has traditionally been used in the previous simulation studies on the MG, is to endow players with sets of adaptive strategies, where a strategy may be defined as a function mapping a particular sequence of observed past outcomes up to a certain period with an action to choose in the next period. In the following we show indeed that an explicit account of this form of strategic behavior does not change the set of the mixed strategy Nash equilibria of the game.

Consider a group of N agents who play the MG with a payoff function² as from Eq. 1. The only information made available after each round is the winning side (0 or 1). We define the market information as the (history) H of play, i.e., a binary string specifying which side has won in each period of the stage game. We also define a parameter m as the portion of the past history H that players retain in memory. So, if $m = 3$, players' strategies will be based only on the outcomes observed in the last 3 rounds of the game.

An *adaptive strategy* may be defined as a prescription on the action to take on the next round of play (i.e., to choose 0 or 1) provided that a particular history (i.e., a particular sequence of m bits) has been observed up to that point. For example, in the case in which $m = 3$, an example of strategy is shown in Table 1.

The history columns specify all the possible sequences of aggregate outcomes (i.e., of winning sides) in the last m periods; the action columns specify which action to choose on the next round in correspondence to each particular sequence observed³. Given a certain value of the parameter m ,

²Various modifications of this payoff function have been proposed by Challet and Zhang and by De Cara et al. (1999)). We stick to the original one for its simplicity and because the essential features of the model are highly insensitive to the proposed variations.

³Note again that this definition of strategy differs considerably from the notion of strategy as complete plan of action used in standard game theory. Rather it represents "mental models" or "subjective beliefs" as in Arthur (1994) and

the total number of strategies that can be generated is equal to 2^{2^m} , corresponding to the 2^{2^m} ways of assigning either action to all the 2^m possible binary strings of length m .

Given the foregoing definition of strategies, we study, as a first important benchmark, the equilibria associated with the “homogenous” game where all the players have access to the full set of strategies. Moreover, rather than considering the dynamics of the history string h , generated by the successive players’ choices, we begin by focusing on the static case requiring that all the strings h can appear with equal probability⁴. Under this hypothesis, despite the high number of strategies, we show that a unique mixed strategy Nash equilibrium exists by which players assign equal probability to choosing either side, 1 and 0, regardless of past history. However, given the above definition of strategy, there are actually infinite ways in which players may end up choosing sides with equal probability.

In order to clarify the analysis that follows, let us introduce some notation. A pure strategy $s \in S_l$ is a mapping $H_l \rightarrow \{0, 1\}$ from the set of binary strings of length l to the set of actions. Consequently, a mixed strategy $m = \{x_\alpha, \alpha \in S_l \mid \sum_{\alpha \in S_l} x_\alpha = 1\}$ can be thought as a map $H_l \rightarrow [0, 1]$ which associates to each binary string $h \in H_l$ the probability $m(h) = \sum_{\alpha \in S_l} x_\alpha s_\alpha(h)$ that the action following it be 1. Let Δ_l be the 2^{2^l} dimensional mixed strategies simplex. For what follows, it is important to notice that the actual mapping between the mixed strategies space and the space $\times^{2^l} [0, 1]$ of maps from H_l to $[0, 1]$ is many to one and surjective, i.e. there are many ways of building up the same mixed strategy starting from the pure strategies set and each mixed strategy can be realized in at least one way.

Consider a population of N (odd) agents. If $\bar{m} \in \times^N \Delta_m$ is a mixed strategy profile (the vector composed by the players mixed strategies) we denote the mixed strategy of player i as m_i and the vector of mixed strategies of all players except i as \bar{m}_{-i} . The mixed strategy profile defines, for each binary string, the population probability of playing 1 (the expected frequency of players playing action 1) associated to a given binary string:

$$\bar{m}(h) = \frac{1}{N} \sum_{n=1}^N \sum_{j \in S_l} m_j(h) = \sum_{j \in S_l} P_j(\bar{m}) s_j(h) \quad (2)$$

where $P_j(\bar{m})$ is the expected probability of playing strategy j (the frequency with which this strategy will be played).

The payoff function in Eq.1 defines the expected payoff of player i given the past history h and the complete mixed strategy profile \bar{m} :

$$U_{m_i, \bar{m}}(h) = m_i(h)(1 - \bar{m}(h)) + \bar{m}(h)(1 - m_i(h)) \quad (3)$$

and we say that a mixed strategy m_1 is superior to m_2 if, given the mixed strategy profile \bar{m} , it is $U_{m_1, \bar{m}}(h) \geq U_{m_2, \bar{m}}(h)$, $\forall h \in H_l$. This definition allows one to discuss equilibria in a way similar to what is done in more canonical population games.

Theorem The strategy profile \bar{m} is a Nash equilibrium iff $m_i(h) = 1/2$, $\forall m_i \in \bar{m}$, $\forall h \in H_l$

Marengo and Tordjman (1996).

⁴This prescription is introduced for purposes of analytical tractability and stems from a symmetry consideration: there is no reason for a player to assign ex-ante a greater probability to a particular history string.

The meaning is rather straightforward: a mixed strategy profile is a Nash equilibrium only when the players' mixed strategies assign probability one half to pick either side, 1 and 0 irrespectively of the past history, i.e. when all the players are perfectly symmetric random players. Notice, however, that the ways of realizing such a mixed strategy are infinite. Hence, the actual number of mixed strategy Nash equilibria is also infinite.

Proof

In order to demonstrate the assertion we proceed by identifying the i -th player's best replies to a given mixed strategy profile \bar{m}_{-i} for the rest of the population. For this purpose, it is convenient to express the payoff in Eq. 3 isolating the contribution of the i -th player's strategy from that of the rest of the population:

$$U_{m_i, \bar{m}_{-i}}(h) = \frac{N+1}{N}m_i(h) + \frac{N-1}{N}\bar{m}_{-i}(h) - \frac{2}{N}m_i(h)^2 - \frac{2(N-1)}{N}m_i(h)\bar{m}_{-i}(h) \quad (4)$$

and $m_i(h)$ has to be chosen in order to maximize the foregoing expression. The solution depends on the value of $\bar{m}_{-i}(h)$ and the quadratic form of Eq. 4 allows one to immediately find it. Defining

$$p^*(h) = \frac{N+1-2(N-1)\bar{m}_{-i}(h)}{4} \quad (5)$$

the solution reads

$$\hat{m}_i(h) = \begin{cases} 0 & p^*(h) \leq 0 \\ 1 & p^*(h) \geq 1 \\ p^*(h) & \text{otherwise} \end{cases} \quad (6)$$

The same procedure can be repeated for each string h to obtain the best reply mixed strategy to $\bar{m}_{-i}(h)$. In order to obtain the set of Nash equilibria, we have to find strategies that belong to the set of mutual best replies, i.e. $\hat{m}_i(h) = \bar{m}_{-i}(h), \forall h \in H_l$. Inspecting Eq. 6, it turns out that this can happen only when $0 < p^*(h) < 1$ and the following condition is satisfied

$$\bar{m}_{-i}(h) = \frac{N+1-2(N-1)\bar{m}_{-i}(h)}{4} \quad (7)$$

The solution is then $\bar{m}_{-i}(h) = 1/2$, implying $\hat{m}_i(h) = 1/2$.

Q.E.D.

3 The Minority Game with Heterogeneous Agents

One of the interesting questions one may explore in the MG framework is whether and how the presence of heterogeneity among players' beliefs may impact on the system degrees of self-organization around the Nash equilibria of the game.

In line with previous simulation studies by Challet and Zhang, one may introduce heterogeneity in the population via a diversity of strategies. Given a certain value of m , each player is initially endowed with a number s of strategies which are randomly drawn from a common pool. Hence, although both m and s are identical over the population, heterogeneity arises from the random initial strategy assignments.

In the course of play, each active strategy i is characterized by a value $q_i(t)$, which indicates the total number of points accumulated by that strategy at time t . Indeed, after each period of the game,

all the strategies that have made a correct prediction in that period (i.e., all strategies prescribing the choice of the side which ex post resulted the winning side) are assigned one point each.

Given the initial strategy assignment and the updating rule just described, players' behavior at each round of the stage game is completely deterministic, as each agent picks, among the strategies he possesses, the one with the highest number of accumulated points.

Note that in such a framework, however, agents' heterogeneity is *only* introduced via the initial assignment of strategies to players and does not stem from different "personal histories". If, for example, the same strategy i is initially assigned to two different players, the value of $q_i(t)$, which determines the strength of that strategy at time t will necessarily be the same for both players, as it will only depend on the number of times that strategy has made a correct prediction whether it was actually played or not.

In order to evaluate the population's performance, it is necessary to introduce a measure of *allocative efficiency*. A natural candidate is provided by the average number of players belonging to the winning party, i.e. the average number of points earned by the whole population⁵. In accordance with the previous literature, as a measure of allocative efficiency we compute a quantity associated to the foregoing one, namely the mean squared deviation from the half population, σ . Let N be the number of agents, and $N_0(t)$ the number of agents choosing side 0 at time t , then in a repeated game of duration T the mean squared deviation is computed as

$$\sigma = \frac{1}{T} \sum_{\tau=0}^T (N_0(\tau) - \frac{N}{2})^2. \quad (8)$$

Clearly, the lower the value of σ , the higher the system efficiency.

The dynamic process can be expected to depend to some degree on the initial strategy distribution and on the initial game history, which are both generated randomly. Therefore, in order to eliminate any dependence on initial conditions from our results and to focus only on asymptotically stable states, in all the simulations presented here we applied an averaging procedure over 50 independent sample paths with randomly generated initial histories and strategy distributions.

In addition, at the beginning of each simulation the system was left to evolve for an initial "adjustment phase" of length T_0 in order to wash away any possible transient effects on the subsequent averaging procedure. The quantities so obtained can thus be considered asymptotic properties of the system as long as T_0 and T are chosen high enough as to provide a good approximation of the limit $T \rightarrow \infty$.

The dependence of the volatility measure σ on N , m and s for the original minority game has been thoroughly investigated in the previous simulation studies and is summarized in Fig. 1 for the case $s = 2$.

As noticed by Savit et al. (1998), the type of market regime is determined, at least in first approximation, by the ratio $z = 2^m/N$: hence the curves for various N collapse if plotted in this variable. In this respect, notice that even if the actual number of possible strategies is 2^{2^m} , their relative strengths are completely defined in term of the frequency $P(0|h_m)$ with which, over a history, a 0 follows a given m -length string h_m . And there are 2^m of such variables. So, z can be interpreted as the density of agents in the strategy space degrees-of-freedom.

⁵Recall from the previous Section that this quantity measures the degree to which the population is close to a Nash equilibrium, whether pure or mixed strategy.

As shown in Fig. 1 three different regimes of the system can be identified. First, a “random regime” occurs when z is large (the agents are sparse in the strategy space). Here the system can hardly organize and its behavior can be described as a collection of random agents choosing their side with a coin toss. In fact suppose the past history to be a given h_m and suppose there are $N_d(h_m)$ agents whose strategies prescribe different actions based on that history while there are $N_0(h_m)$ and $N_1(h_m)$ agents whose strategies prescribe the same party (we restrict ourselves to the $s = 2$ case), respectively 0 and 1. If the agent in N_d choose randomly, the variance is $\sigma(h_m) = N_d(h_m)/4 + (N_0(h_m) - N_1(h_m))^2/4$. The average over the possible h_m will then give $\sigma = N/4$. Notice that σ is shaped by two factors, namely a fluctuation in the choices of agents able to choose and a fluctuation in the initial distribution of strategies. Note also that the allocative efficiency in this case (as measured by σ) equals that of a population of players playing the mixed strategy Nash equilibrium defined over the two actions 0 and 1 (cf. Section 2.1).

The second regime could be called the “inefficient regime” for $z \ll 1$. Here the agents densely populate the strategy space, and their actions are strongly correlated. This correlation leads to a worsening of the overall performance due to a “crowd” effect (Johnson et al. 1998): the agents in fact are too similar to each other and they all tend to choose the same party on the basis of the available information.

The third regime for $z \sim 1$ is where coordination produces a better-than-random performance. Here the agents are differentiated enough not to yield “crowd” effects, but at the same time sufficiently distributed over the strategy space so as not to produce a random-like behavior.

The literature on the MG in fact, has mostly focused on the criticality of the value z_c where σ is minimum, suggesting that a major change in the system behavior happens when this point is crossed (Challet and Marsili, 1999). As we will see in the following sections, this criticality survives to the introduction of the probabilistic learning rules.

4 Adaptive Learning

4.1 The Model

What happens to the system properties when additional heterogeneity is introduced in the population? In particular, what happens when agents are endowed with a probabilistic decision rule? In order to tackle the problem, in the following we investigate changes in the dynamics and asymptotic properties of a population of agents playing the MG as a function of changes in the nature of the agents’ learning models. Hence, we leave unaltered the setup previously described, and modify only the way in which agents update their strategies’ relative strength. In particular, we adopt the following probabilistic updating rule.

Recall the definition of $q_i(t)$ as the total number of points strategy i would have won if played until time t . Then each agent chooses among her strategies following the probability distribution:

$$p_i(t) = \frac{e^{\beta q_i(t)}}{\sum_j e^{\beta q_j(t)}}. \quad (9)$$

where the sum on j is over all the strategies possessed by the player. Note that, in general, different players will assign different probabilities to the same strategy due to the different strategy endow-

ments. Hence, the introduction of a probabilistic learning rule adds noise at the individual level and it also increases heterogeneity in the population.

The choice of a stochastic learning model is well supported by the available experimental evidence on adaptive behavior in games as well as by evidence from psychology experiments. In fact, most descriptive models of adaptive learning, whether belief-based or reinforcement-based, imply a probabilistic choice and updating of available actions [cf. Erev and Roth (1998), Camerer and Ho (1999)].

The parameter β can be considered as a sort of *sensitivity of choice to marginal information*: when it is high, the agents are sensitive even to little differences in the notional score of their strategies. In the limit for $\beta \rightarrow \infty$ the usual minority game rule is recovered. On the contrary, for low values of β a great difference in the strategies' strengths is required in order to obtain significant differences in probabilities.

The model indeed bears close similarities with a discrete time replicator dynamics (see Weibull (1995)). The connection is straightforward if one looks at the probability updating equation associated with Eq. 9:

$$p_i(t+1) = p_i(t) \frac{e^{\beta \delta q_i(t)}}{\sum_j p_j(t) e^{\beta \delta q_j(t)}}. \quad (10)$$

where $\delta q_i(t) = q_i(t+1) - q_i(t)$ are the points won by strategy i at time t . If one thinks of a continuous process $\delta q_i(t) = \dot{q}_i(t) \delta t$, where $\dot{q}_i(t)$ is the instantaneous “fitness” of strategy i , then the continuous time replicator dynamics equation is recovered keeping only the first terms in δt expansion:

$$\frac{\dot{p}_i(t)}{p_i(t)} = \dot{q}_i(t) - \sum_j p_j(t) \dot{q}_j(t) \quad (11)$$

4.2 Transient length

In everything that follows we will restrict our analysis to the case $N = 101$ and $s = 2$ and we will speak of the optimal value for memory length m_o with reference to the value of m which minimizes σ under this parameter choice⁶.

Let us consider the problem of defining the correct values for T_0 and T in Eq. 8 above. The central question is: How long must the system be left evolving before it reaches the asymptotically stable dynamics? Fig. 2 plots the average σ value based on the “deterministic” (i.e. $\beta = +\infty$) MG as a function of the time length T over which this average is taken with a transient $T_0 = T$. As it can be seen from the graph, the values used in the literature on the MG are generally sufficient to obtain a prediction correct to few percentage points. However, two properties are worth noticing:

- The system approaches the asymptotic value from above, intuitively suggesting that the system “learns” over time to self-organize
- For low values of m , in the “inefficient regime”, and for high value of m , in the “random regime”, the system reaches a stable dynamic quite fast. On the contrary, for values of m near the optimal value m_o , the system takes a longer time to self-organize

⁶The choice to set $s=2$ is justified by the fact that the system exhibits the same qualitative properties for any $s \geq 2$, while reducing to a trivial case for $s=1$.

Consider now the case in which the learning rule is the one described in Eq. 9, shown in Fig.3. For high values of β this learning rule approaches the standard one, and accordingly, the transient length is similar to the one found in such cases. However, as β decreases, the length generally increases. The increase is most dramatic for values of m near the optimal value m_o , and it progressively disappears for higher values of m . The interpretation of such a result stems from the meaning of β in terms of the learning rule. Supposing non trivial dynamics for m near m_o , the parameter β sets the time scale on which such dynamics are attained.

As an illustration, consider the following example: Let be $r(t) = p_1(t)/p_2(t)$ the ratio of the probabilities that an agent associates to his two strategies, and $\Delta q(t) = q_1(t) - q_2(t)$ the difference in their respective strengths. From Eq. 9 it follows that $r(t) = e^{\beta \Delta q(t)}$. Assuming that the difference in the two strategies performance holds constant over time (an assumption which is generally true in the initial transient regime where agents' behavior is basically random) we obtain $\Delta q(t) \sim t$: hence, from the equality above, a given difference in probability is obtained at a time which is inversely proportional to β .

In order to estimate the time scale over which the system long-run properties are attained, we use the following procedure. Holding all the parameters and the initial conditions constant, the system volatility can be expressed as a function of both the transient phase duration, and of the time length over which it is averaged, i.e. $\sigma = \sigma(T, T_0)$. Starting from a reference time T_r ,⁷ we compute the mean volatility progressively doubling T and T_0 , and thus obtaining a series of values $\sigma_n = \sigma(2^n T_r, 2^n T_r)$. When the relative variation $|\sigma_n - \sigma_{n-1}|/\sigma_n$ falls below a fixed threshold ϵ , we stop and take the last computed value of σ as an estimate of its asymptotic value. The corresponding time length $\hat{T}(\epsilon)$ will be an estimate of the time implied by the system to reach this asymptotic state. As can be seen in Fig. 6 the increase in \hat{T} when β is lowered is mainly concentrated around m_o , with shapes that suggest the presence of a regime discontinuity.

4.3 Allocative efficiency

In order to analyze the asymptotic properties of $\sigma(m)$ for different β 's, we use the procedure just described regarding the calculation of \hat{T} , i.e. we leave the system evolve until "stability" is reached⁸. The simulation results are plotted in Fig. 5. Interestingly, when β decreases, the performance level of the system generally increases. Moreover, such increase is larger the lower the value of m , and it becomes negligible for $m \geq m_0$. The observed behavior is consistent with the idea that for high values of m , the system dynamics tends to be determined by the initial distribution of strategies among players, while players themselves have no opportunity to attain a higher performance by adjusting their behavior. Recall that an increasing m means an increasing number of possible strategies over which players may initially draw. For a fixed N , the "ecology" of drawn strategies becomes thinner as compared to the notionally available ones. That phenomenon, it turns out, prevents the system from self-organizing. Note that this property is quite robust and largely independent from the particular learning rule adopted. Our results, in fact, are perfectly in line with previous simulation studies in the high m region.

⁷Note that the chosen value for T_0 is irrelevant as long as it is small compared to the typical time scale.

⁸Let us emphasize that the stable state does not imply point convergence to any state but simply long-run stability of the relevant time-averages (e.g., the mean volatility σ) even if the system continues to fluctuate in its limit state.

Conversely, for low values of m , the choice of the particular learning rule adopted does matter in terms of aggregate efficiency. In fact, when m is small, the original learning rule ($\beta = \infty$) produces a “crowd” effect (corresponding to large groups of agents choosing the same side) which, due to homogeneity in the initial strategy endowments, prevents the system from attaining a high degree of efficiency⁹. On the contrary, the introduction of a probabilistic learning rule for the strategy choice acts like a brake that dumps the amplitude of such correlated fluctuations. At the individual level, this corresponds to lower β 's, i.e. to higher degrees of “inertia”, as agents update their probabilities more slowly. In other words, as β decreases each agent behaves as if he was applying a sort of stochastic fictitious play approximation [e.g., Fudenberg and Levine (1998)]¹⁰, with an implicit assumption of stationarity on the distribution of other agents' choices. If the whole population shares the same β - as in the present model - the assumption is, in a way, self-fulfilling: a decrease in β makes the behavior of the population as a whole change at a slower pace. A slower probability updating at the individual level and the resulting more stable collective behavior, together, imply that σ is a non increasing function of β . In fact, if the system reaches a dynamical stability via an averaging procedure over the past outcomes, increasing the time scale over which averaging occurs cannot rule out previously attainable equilibria.

However, note that in order to capture the long run efficiency properties of the system for low values of β it is necessary to let the population play until stability is reached, according to the procedure described above. Our results, in fact, were not captured by previous simulation studies (Cavagna et al., 1999). Since the latter were all performed with a fixed time length their conclusion was that when β is small enough, the system behavior resembles the behavior of a random system. This finding was in fact due to both the increase in the transient length and the purely random initial dynamics which occur when β is decreased. Then, by fixing a time length, for small values of β the simulations capture throughout the adjustment phase, and the system behavior perfectly mirrors that of a collection of agents who choose at random (indeed our results of a fixed time simulation are plotted in Fig. 6 and are perfectly in line with the existing literature¹¹).

As can be seen in Fig. 5 the performance attainable in the MG via a collective organization of agents with limited information and limited ability to choose is, in general, surprisingly high. In the following we show that the level of efficiency achieved in the double limit $\beta \rightarrow 0$ and $m \rightarrow 0$ actually equals that attainable with hypothetical perfectly informed and perfectly rational agents endowed with a greater flexibility in choice.

Consider for instance a collection of agents with the following characteristics: each agent is assigned $S = 2$ strategies, and a vector of length 2^m containing the probability $p(h_m)$ of playing according to its first strategy after the occurrence of the history string h_m . Moreover, for each h_m , each agent knows the values of $N_0(h_m)$, $N_1(h_m)$ and $N_d(h_m)$ indicating respectively the number of agents for which their strategies both prescribe to play 0, both to play 1, or to play differently. Assuming that the game structure and the amount of information available to agents is common knowledge and that the agents are perfectly rational, the problem completely factorizes and, for each

⁹In some sense, one can interpret the “crowd” effect as a collective form of *overreaction* (Thaler, 1993).

¹⁰Note that fictitious play implies that a player best responds to the observed frequency of the opponent's play.

¹¹In other words, to discover the asymptotic properties of the system, the limits $T \rightarrow \infty$ and $\beta \rightarrow 0$ have to be performed in this order.

h_m , every agent in $N_d(h_m)$ will solve the game analytically choosing $p(h_m)$ in order to minimize

$$\frac{(N_1(h_m) - N_0(h_m))}{2} - p(h_m)N_d(h_m) \quad (12)$$

i.e. making the average fraction of the population choosing a given side as near to $N/2$ as possible. This choice will produce a volatility $\sigma \sim N_d/4 = N/8$ which is roughly similar to what obtained in the simulation shown in Fig. 5 in the low m , low β region¹². However, note that, as from Fig.5, these fully rational, fully informed players are (on average) “beaten” by a set of “self-organizing” agents with memory $m \sim m_o$, reaching a nearly double efficiency.

A final remark concerns the variance of the distribution of σ as a function of β for various m 's, as shown in Fig. 7: when β decreases the variance of σ decreases for any m . However it remains three times greater for $m = m_o$ suggesting a stronger dependence of the asymptotic performance on the initial strategy assignment which the system is not able to wash out. That is, significant *path-dependence effects* are present.

4.4 Informational efficiency

What we have been calling allocative efficiency basically highlights the collective ability of capturing the payoffs which the game notionally allows. A complementary issue regards the *informational efficiency* of the market process, i.e., the extent to which the future system outcomes are unpredictable, or, in other words, the absence of any arbitrage opportunities. Thus, let us analyze the informational content of the binary string H of successive winning sides. Let $p(0|h_l)$ be the probability that a 0 follows a given string h_l of all the possible 2^l strings of length l (as depicted in Fig. 8).

For the original “deterministic” MG the analysis performed in Savit et al. leads to the identification of two regimes: a “partially efficient regime” for $m < m_o$ in which $p(0|h_l) \sim .5, \forall h_l$ as long as $l \leq m$ and in which no informational content is left for strategies with memory less or equal to the one used by the agents; and an “inefficient regime” for $m > m_o$ in which the distribution of $p(0|h_l)$ is not flat, even for $l \leq m$, meaning there are good strategies that might exploit the market signal to obtain differential profits. For $l > m$ both the regions show a non trivial distribution $p(0|h_l)$ with an increasing degree of “ruggedness” as l increases. The effect of introducing some degree of behavioral randomness through the parameter β leads to the obvious effect of reducing the “ruggedness” of the distribution of $p(0|h_l)$ (see Fig 8).

In order to study the behavior of the system as β changes we introduce two related quantities which can be used to characterize the informational content of the time series. The first is the conditional entropy $S(l)$ defined as:

$$S(l) = - \sum_{h_l} p(h_l) \sum_{i \in \{0,1\}} p(i|h_l) \log p(i|h_l) \quad (13)$$

where the summation is intended over all the possible strings of length l and $p(h_l)$ is the frequency of a given string in the system history H . The maximum value $S(l) = 1$ is reached for a flat distribution $p(0|h_l) = .5$, and it corresponds to the impossibility of forecasting (in probability) the next outcome

¹²We are assuming $\Delta N = N_1(h_m) - N_0(h_m) < N_d(h_m)$. Notice that for random agents $\Delta N \sim \sqrt{N}$ and $N_d \sim N$ and that one can neglect the $\Delta N/N_d$ terms in the solution of Eq. 12 when N is large.

based on the previous l outcomes. The idea that the information content can be used to “exploit the market” leads to the definition of a second quantity $A(l)$:

$$A(l) = \sum_{h_l} p(h_l) \max \{p(0|h_l), p(1|h_l)\} \quad (14)$$

which is the average fraction of points won by the best strategy of memory l . This is a measure of the reward obtained by the best arbitrageur with memory l (whereas if no arbitrage opportunities are present $A(l)$ is equal to .5.)

Before analyzing the behavior of these quantities when β is varied, let us briefly consider as a sort of ideal benchmark the properties of a population composed of perfectly rational, perfectly informed, agents with common knowledge of strategy distributions. Not surprisingly, in these circumstances the problem factorizes for each past history and the dependence on m disappears. The history produced by such a system is a random series of 0 and 1. Indeed the number of agents choosing one side is distributed according to a binomial around $N/2$ with different widths for different h_m . In particular, this means that in such an extreme case the “memory” loses any predicting power and no arbitrage opportunity is left for agents with longer memory, i.e. no residual information is left in the time series and the behavior of agents makes the market perfectly efficient from an informational point of view. In the last resort, there is nothing to be learned from any history because agents know everything from the start and coordinate their mixed strategies accordingly. Under this assumption we expect $S \sim 1$ and $A \sim .5$.

Short of such an ideal case where the market loses its coordinating role, because agents *ex ante* generate the coordination “in their heads”, let us consider, for example, a population of “random agents”. Here, due to the unbalance in the initial strategy endowments we expect a non trivial structure to appear for every l ; thus $S < 1$ and $A > .5$.

In Fig. 9 we plot $S(l)$ and $A(l)$ for histories generated with a value of $m < m_0$, in the “partially efficient regime”. The effect of decreasing β shows up when $l > m$ but the information content for high l is never completely eliminated. The market becomes less efficient the larger the time scale l at which it is observed. In fact it can be shown under very general assumptions that certain strings in the history are more abundant than others (Savit et al., 1998) and the long-range correlation that was responsible for the “crowd” effect at high β survives as a non trivial structure in $p(0|h_l)$ for high l . All this applies despite the fact that to an agent with memory $l \leq m$ the market appears perfectly efficient regardless of the β value.

For values of $m > m_0$ in the “inefficient phase” the effect is in some sense reversed. As can be seen in Fig. 10 the effect of decreasing β is again negligible for $l \leq m$ but in the limit $\beta \rightarrow 0$ the curve becomes flat for $l > m$. This last result deserves some comments: the flatness in $l \geq m$ means that no gain is achieved from inspecting the time series with a very long memory $l \gg m$ because no more arbitrage opportunities are open for a longer memory agent than the best possible agent of memory m . The market can be considered to be, again, partially efficient in the sense that it generates an upper bound on the maximal attainable arbitrage capability which does not depend on the arbitrageur memory.

The particular form of the conditional entropy in Fig. 10 suggests that in the limit $\beta \rightarrow 0$ the system can be described as a Markov chain of memory m ¹³. The result can be explained

¹³Notice that by construction, in the learning rules considered here the past is not discounted, and the agents weigh

by noticing that when β is small only great differences in the past performances of strategies are relevant and in the limit $\beta \rightarrow 0$ only infinite differences become relevant. Putting it another way, the frequency of victories of the various strategies becomes constant implying the formation of a static hierarchical structure in the strategy space which at the end is responsible of the Markov character of the resulting history. The appearance of “best strategies” in $m > m_o$ region is well revealed by the plot of the average score by the best strategy versus the average point scored by the player (see Fig. 11): a correlation in fact appears between the performance of a player and the performance of its best strategy. Moreover, in the $m \sim m_o$ region a sub-population showing the same kind of high correlation coexists with another group that presents no correlation, composed of agents possessing two equally performing strategies.

Conversely, only in the low m region no strategy ends up being preferable to others and no player is bound to lose due only to his initial strategy endowment.

Notice, however, that equivalence between strategies does not necessarily imply equivalence in the agents’ performances. This is highlighted by the plot of the variances and the supports of the points distribution for different values of β and m in Fig. 12. Interestingly, only for low m and low β does equivalence in strategy performance imply a relatively uniform distribution of points over the population. In the other parameter regions learning does not eliminate performance heterogeneity over the population. Loosely speaking, the market self-organizes over an ecology of “mental models” and players, entailing the long-run coexistence of relative “winners” and “suckers”.

5 Conclusions and Outlook

One of the central questions we addressed in this work is the extent to which market dynamics generated by arbitraging activities, as represented in Minority Games, display generic properties, independent from particular behavioral assumptions. Our answer is largely negative: simple variations in the agents’ learning algorithms, we show, yield important modifications in the asymptotic properties of the system.

More specifically, we show that *less* sensitivity to marginal information, i.e., *more inertia* in the learning algorithm entails improved long-run collective performances, although at the expense of longer adjustment phases. Together, *performance asymmetries* across agents, as measured by the variance (or analogously) the support of the earnings distribution over the population, fall as inertia in the agents’ behavior increases.

In general, some degrees of randomness help in improving allocative and informational efficiencies of the market - as defined above. The major effect of randomness is that it performs like a brake on the system dynamics, thus preventing groups of players who densely populate the strategy space from acting in a strongly correlated way and thus from producing “crowd” effects which worsen the system performance¹⁴.

their strategies on the basis of all the game outcomes starting from the beginning of the simulation (however, see the Appendix for an analysis of the system properties when a time decaying factor is introduced).

¹⁴The introduction of randomness in individual behavior is indeed only one of possible ways to maintain behavioral heterogeneity in the population. For instance, the same effect has been obtained in De Cara et al. (2000) by substituting the “global” evaluation of strategies on the system history H with a “personal” evaluation in which each agent uses the binary string made up of her own record of victories. A diversification mechanism is again at work breaking the

Table 2 summarizes the different system properties - i.e. different market “regimes” - conditional on different ecologies of strategies (as captured by the parameters $z = 2^m/N$ and β respectively).

Indeed, one of our major conclusions, which refines over previous results in the Minority Game literature, is that market efficiency - in the complementary definitions proposed above - is only achieved in correspondence of an “optimal degree” of heterogeneity, whether in the agents’ decision behavior or in their underlying sets of beliefs. Moreover, our results suggest, more rationality - as approximated here by a greater ability of the agents to track novel environmental information - may well *lower* average performances (an analogous result is obtained in a different setup by Brock and Hommes (1989)).

The general sensitivity of system dynamics upon particular learning algorithms also indicates a natural way forward, beyond the exercises presented in this work, experimenting with cognitively less demanding learning rules. So, for example, it would be interesting to explore the properties of pure reinforcement learning whereby agents update only the strategies they play. That would also set a “zero-level” model in terms of degrees of required information and cognitive abilities - somewhat at the extreme opposite to the models studied so far in the Minority Game literature. And, somewhere inbetween, one might explore more refined learning models such as that in Easley and Rustichini (1999).

Moreover, beyond the strict setup of Minority Games so far, the impact of phenomena of social imitation still awaits to be studied¹⁵. And, more generally, the robustness of the foregoing conclusions ought to be checked in less “reduced form”, institutionally richer models, such as artificial markets of the type outlined in Marengo and Tordjman (1996), Arthur et al. (1997) and Chiaromonte and Dosi (1998) [e.g., Kirman (1999) for a review].

Finally, a complementary line of inquiry regards the analysis of behaviors and learning of human subjects under experimental settings isomorphic to the market interactions formalized above.

In the last resort, all these latter exercises, together with the results presented here ought to be considered as adding some pieces of evidence to the much broader effort aimed at identifying the variables which determine a “universality class”, if any, of market processes involving coordination *cum* heterogeneity, as distinguished from those characteristics which strictly depend upon specific distributions of beliefs and learning rules. Were our results confirmed in further studies, they would add strength to the conjecture that efficient coordination might not stem from the adherence of populations of agents to Nash equilibrium behaviors, but rather emerge out of persistent forms of heterogeneity in beliefs and behaviors within the population.

6 Appendix

Many authors, especially in the experimental literature [e.g., Erev and Roth, (1998)] add to the description of the learning process one more parameter, connected with the idea that agents weigh more the information they received in the recent history as compared to the one coming from far back in the past. This parameter typically takes the form of a decay factor. If $\epsilon_i(t)$ are the points scored by strategy i at time t and α ($0 < \alpha \leq 1$) the information decay factor, the updating rule for correlation among agents.

¹⁵A germane model of financial dynamics with stochastic choice and social imitation is Kirman (1993).

the total strength becomes

$$q_i(t+1) = \alpha q_i(t) + \epsilon_i(t) \quad (15)$$

and the associated updating rule for the probabilities:

$$p_i(t+1) = p_i^\alpha(t) \frac{e^{\beta q_i(t)}}{\sum_j p_j^\alpha(t) e^{\beta q_j(t)}}. \quad (16)$$

The effect of introducing such a “memory leakage” is twofold: First, it puts an upper limit to the maximal strength any strategy could reach, namely $1/(1-\alpha)$. Second, in presence of no information flux, the equiprobability between strategies is steadily restored. This effect implies that if one takes the limit $\beta \rightarrow 0$ keeping the value of α constant, the system will converge to a collection of random agents. In turn, this implies that agents, loosely speaking, have to collect a larger amount of information before they start behaving as a self-organized system.

The effect of introducing “forgetting” in the learning rule is easily understood: if the agents forget more rapidly than they learn they are always bounded to less efficient behavior. Indeed, as can be seen from Fig. 13, if the value of α is decreased the efficiency of the system is proportionally reduced.

References

- [1] Arthur, W.B., (1994), Inductive reasoning and bounded rationality, *Am. Econ. Review Papers and Proc.*, vol. 84, 406-412.
- [2] Arthur, W.B., J.H. Holland, B. Le Baron, R. Palmer and P. Taylor, (1997), Asset Pricing under Endogenous Expectations in an Artificial Stock Market, in Arthur, W.B., S.N. Durlauf, and D. Lane (Eds.) *The Economy as an Evolving Complex System II*, Reading, Mass., Addison Wesley.
- [3] Brock, W.A. and C.H. Hommes, (1998), Heterogeneous beliefs and routes to chaos in a simple asset pricing model, *Journal of Economic Dynamics and Control*, vol. 22, 1235-1274.
- [4] Camerer, C. and T. Ho, (1999), Experience Weighted Attraction Learning in Games: A Unifying Approach, *Econometrica*, 67,4, p.827.
- [5] Cavagna, A., J. P. Garrahan, I. Giardinà and D. Sherrington, (1999), A thermal model for adaptive competition in a market, *Phys. Rev. Lett.*, vol. 83
- [6] Challet, D. and M. Marsili, (1999), Phase Transition and Symmetry Breaking in the Minority Game, *Phys. Rev. E*, vol. 60
- [7] Challet, D. and Y.-C. Zhang, (1997), Emergence of cooperation and organization in an evolutionary game, *Physica A*, 226, 407.
Challet, D. and Y.-C. Zhang, (1998), On the minority Game: Analytical and Numerical Studies, *Physica A*, 256,514.
- [8] Chiaromonte, F. and G. Dosi, (1998), Modeling a Decentralized Asset Market: An Introduction to the Financial “Toy Room”, IIASA, Laxenburg, Austria, Interim Report (Available also as Working Paper, Pisa, St. Anna School of Advanced Studies, 1999).

- [9] De Bondt, W.F.M. and R.H. Thaler, (1995), Financial Decision-Making in Markets and Firms: A Behavioral Perspective, in: R. Jarrow et al (eds.), Handbooks in OR & MS, Amsterdam, Elsevier Science.
- [10] de Cara, M.A.R. and O. Pla and F. Guinea, (1999), Competition, efficiency and collective behaviour in the "El Farol" bar model, The European Physical Journal B10.
- [11] de Cara, M.A.R. and O. Pla and F. Guinea, (2000), Learning, competition and cooperation in simple game, The European Physical Journal B13.
- [12] Easley, D. and A. Rustichini, (1999), Choice without beliefs, *Econometrica*, 67, pp. 1157-1184
- [13] Erev, I. and A. Roth, (1998), Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria, *Am. Econ. Review*, 88, 848-881.
- [14] Fudenberg, D. and D. K. Levine, (1998), *The Theory of Learning in Games*, Cambridge, Mass., The MIT Press.
- [15] Johnson, N.F. and M. Hart and P. M. Hui, (1999), Crowd effects and volatility in a competitive market, *Physica A*, vol. 269
- [16] Kirman, A., (1993), Ants, Rationality and Recruitment, *Quarterly Journal of Economics*, 108, 137-156.
- [17] Kirman, A., (1999), Aggregate Activity and Economic Organization, *Revue Europeenne des Sciences Sociales*, 38, 189-230.
- [18] LeBaron, B. (2000), Agent-based computational finance: Suggested readings and early research, *Journal of Economic Dynamics and Control*, 24, pp. 679-702.
- [19] Marengo, L. and H. Tordjman, (1996), Speculation, Heterogeneity and Learning: A Simulation Model of Exchange Rate Dynamics, *Kyklos*, vol. 49, 407-437.
- [20] Mayer, D.J., Huyck, J., Battalio, R. and Soring T. (1992), Histories role in coordinating decentralized allocation decision: laboratory evidence on repeated binary allocation games, *Journal of Political Economy*, 100, pp. 292-316
- [21] Ochs, J. (1990) The coordination problem in the centralized markets: An experiment, *Quarterly Journal of Economics*, 105, pp.545-559.
- [22] Ochs, J. (1995) Coordination problems, in J.H.Kagel and A.E. Roth (Eds.) *Handbook of experimental economics*, Princeton: princeton University Press, pp. 195-251.
- [23] Savit, R., R. Manuca and R. Riolo, (1998), Adaptive Competition, Market Efficiency and Phase Transition, *Phys. Rev. Lett.*, vol. 82
- [24] Schelling, T.C. (1978) , *Micromotives and Macrobehavior*, New York, W.W. Norton and Company.
- [25] Thaler, R.H., (1993), *Advances in Behavioral Finance*, New York, Russel Sage Foundation.

[26] Weibull, J.W., (1995), *Evolutionary Game Theory*, Cambridge, Mass., The MIT Press.

history	action	history	action
000	1	100	1
001	0	101	0
010	0	110	1
011	1	111	0

Table 1: An example of strategy with $m = 3$.

	low z	high z
low β	high A.E., high I.E.	partial A.E., partial I.E.
high β	low A.E., high I.E.	partial A.E., low I.E.

Table 2: System properties: a summary (A.E. and I.E. stand for Allocative and Informational Efficiency, respectively.)

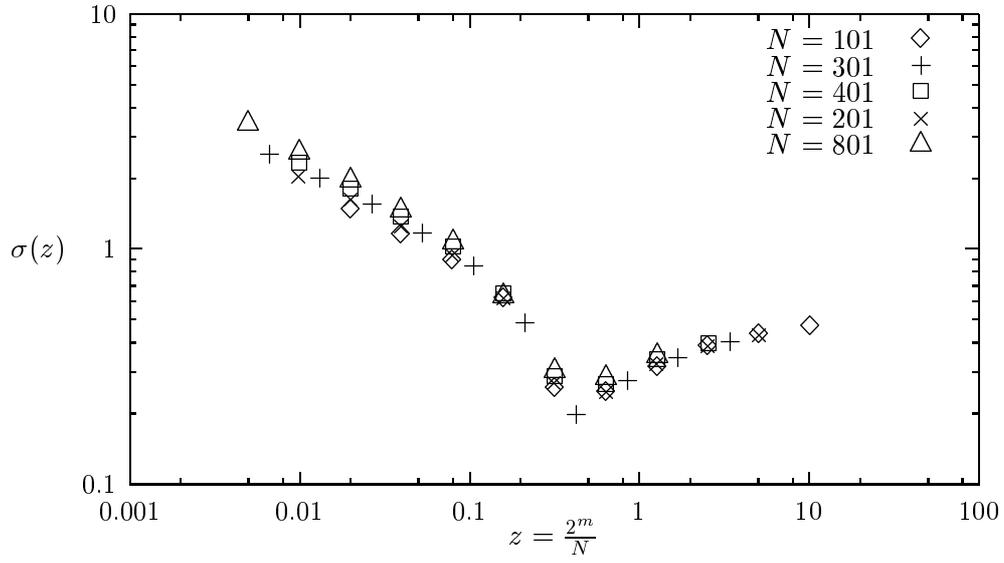


Figure 1: The volatility $\sigma(z)$ for $s = 2$ and different values of N and m .

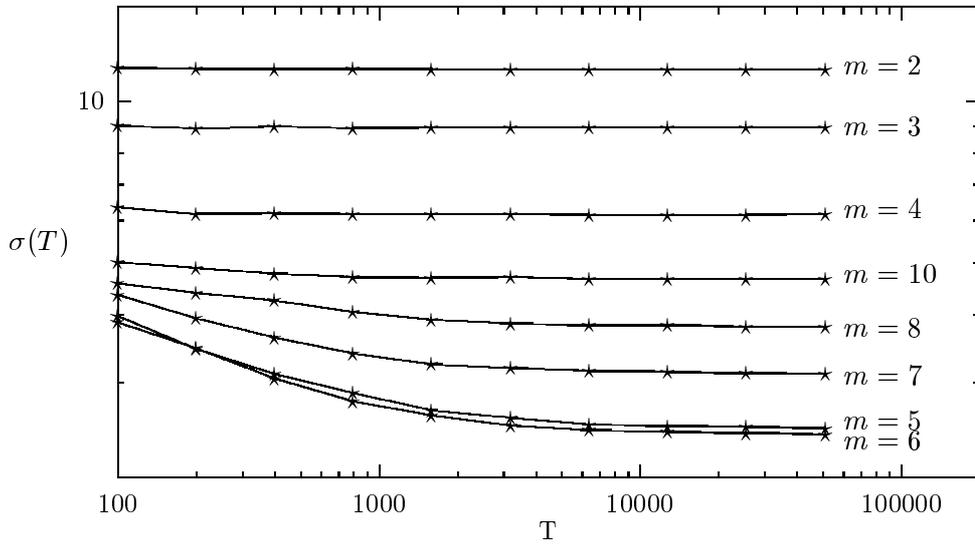


Figure 2: The mean σ along the run length for different m 's. The points are averages over a sample of 30 independent runs with $N = 101$ and $s = 2$.

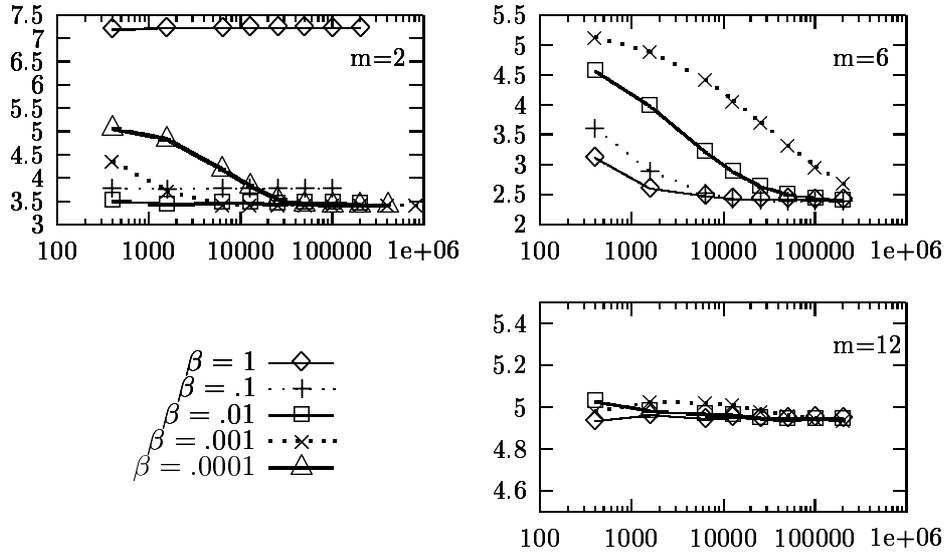


Figure 3: σ along the run length T for different β 's. The points are averages over a sample of 30 independent runs with a transient time $T_0 = T$.

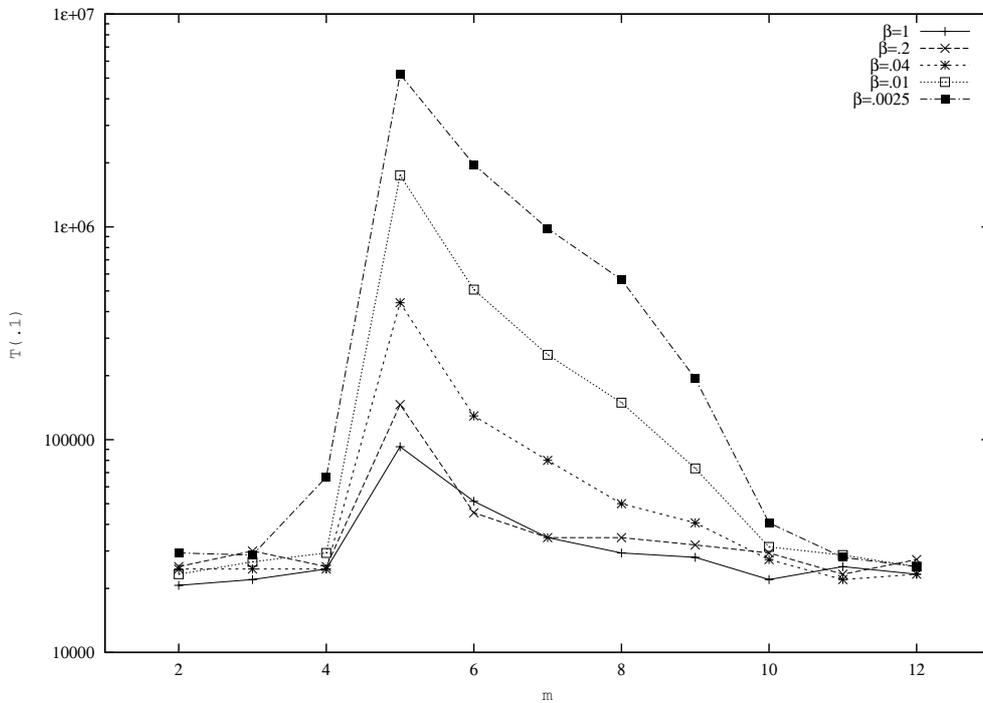


Figure 4: A (rough) estimate of the time \hat{T} required to reach the stable asymptotic σ value with an error not greater than a given threshold (The plots are made against m for different values of β).

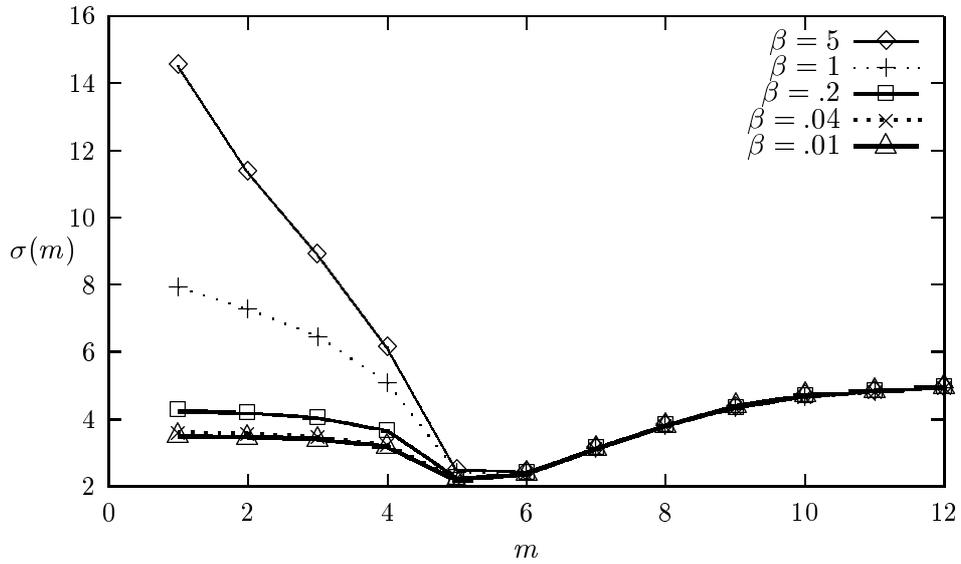


Figure 5: Volatility σ for $s = 2$, $N = 101$ and various m 's and β 's. The runs are performed by doubling the time length T until the last two values obtained differ by less than 1%.

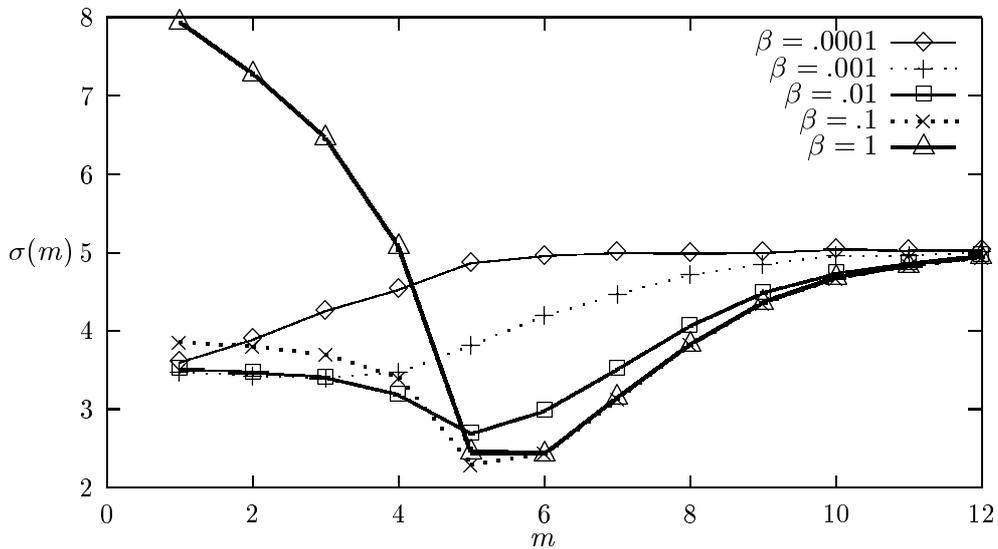


Figure 6: The volatility σ for $s = 2$, $N = 101$ and various m 's and β 's. The runs are performed with a fixed time length $T = T_0 = 10000$. When $\beta \rightarrow 0$ the system approaches a collection of randomly choosing agents.

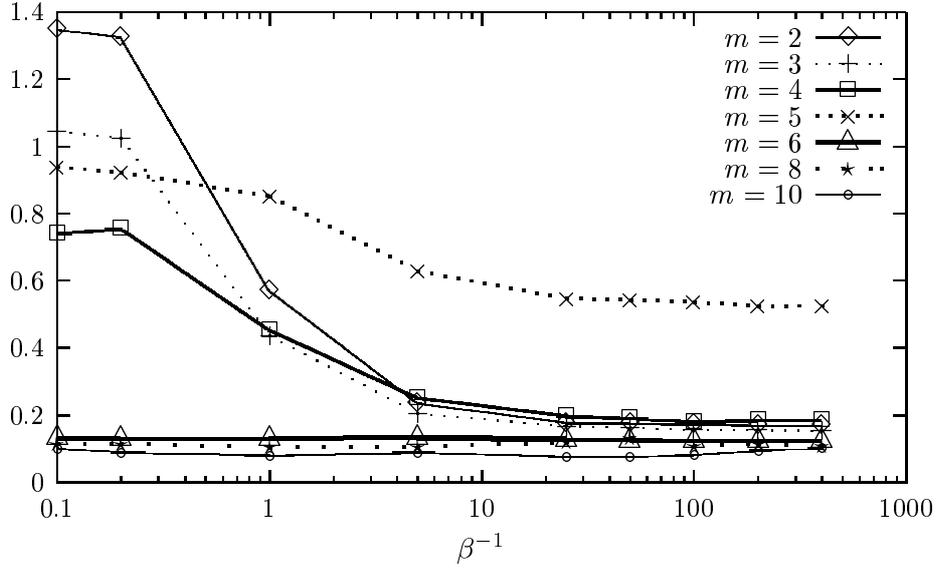


Figure 7: Variance of the distribution of σ over a sample of 50 independent runs. As β becomes small the point $m \sim m_0$ maintains a significantly larger variance.

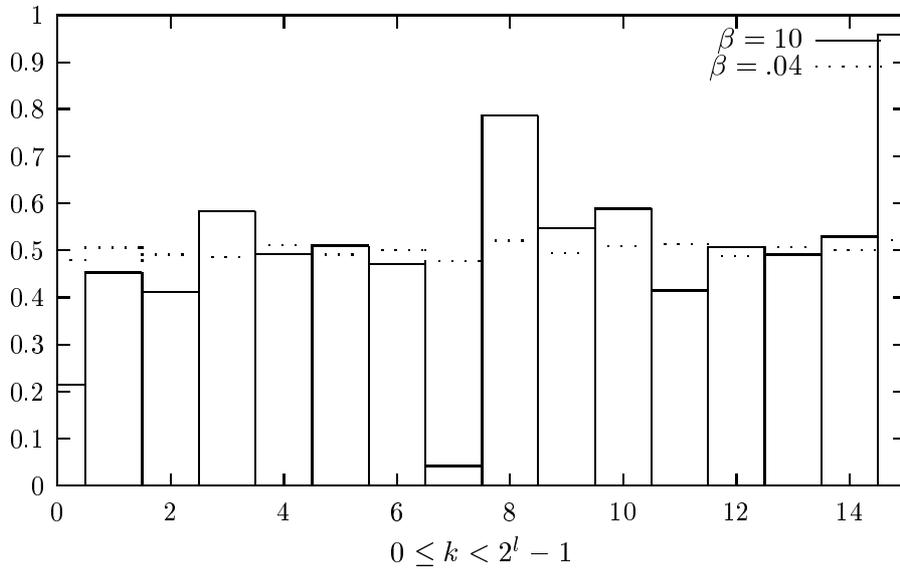


Figure 8: Probability $p(0|h_l)$ of obtaining a 0 following a given binary string h_l in the system history for $m = 3$ and $l = m + 1 = 4$. When β is reduced the distribution “flattens” and any structure disappears.

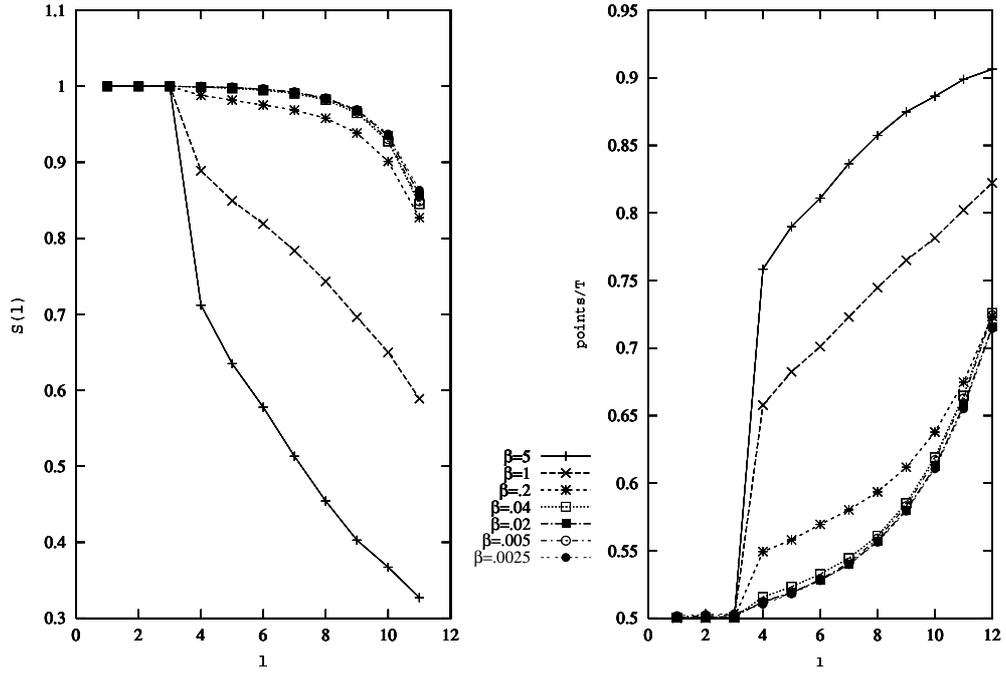


Figure 9: Conditional entropy $S(l)$ (left) and arbitrage opportunity $A(l)$ (right) as a function of the time depth l for $m = 3 < m_o$.

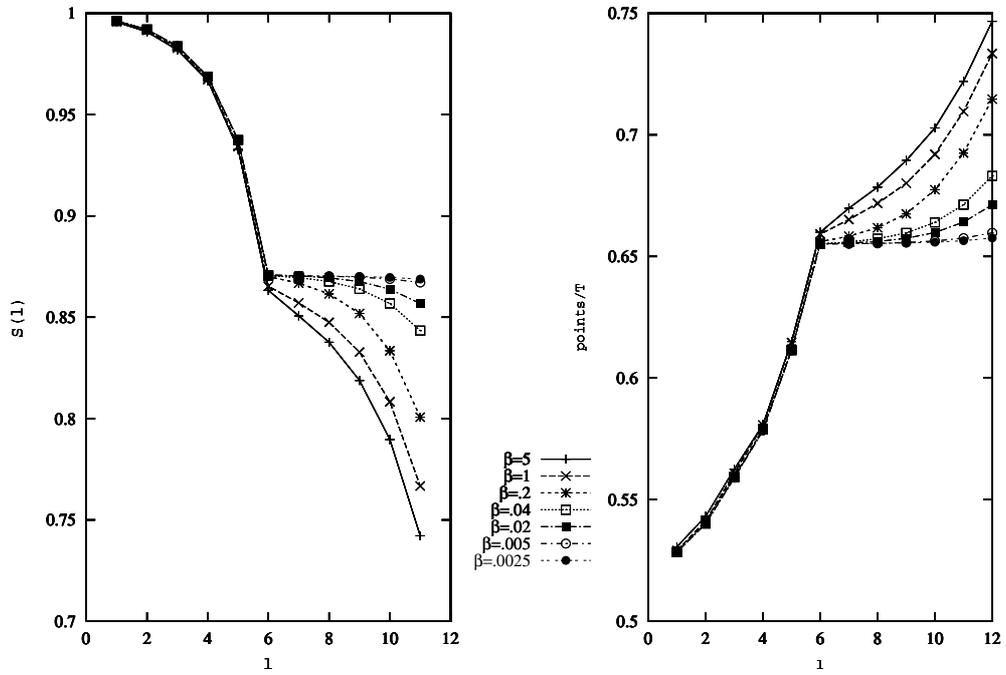


Figure 10: Conditional entropy $S(l)$ (left) and arbitrage opportunity $A(l)$ (right) as a function of the time depth l for $m = 6 > m_o$.

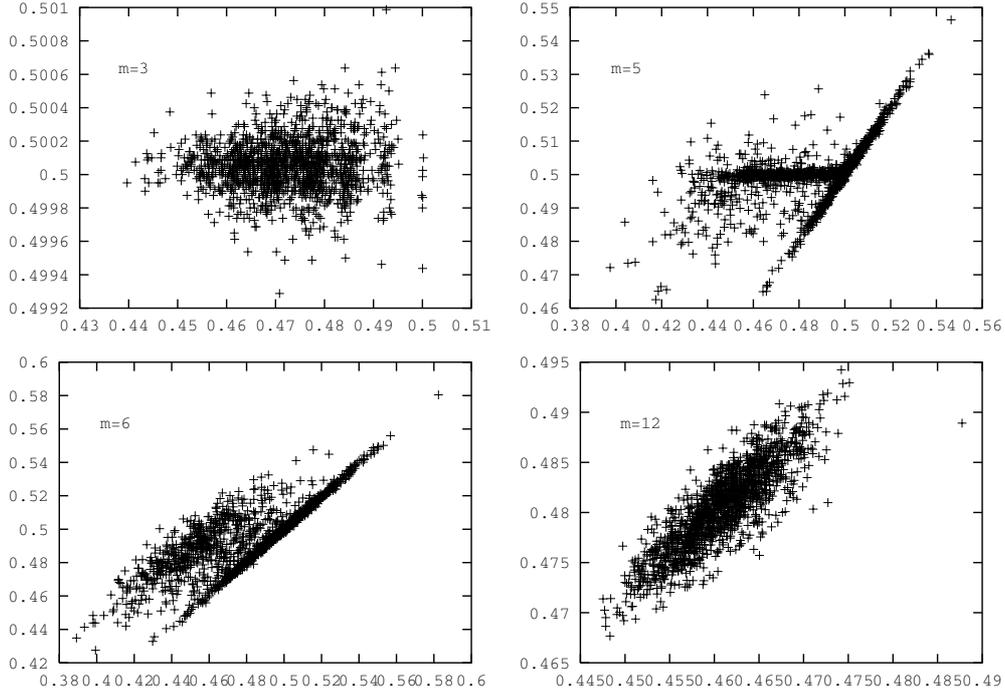


Figure 11: Plot, for each player, of the scoring rate of his best strategy against his own winning rate for a population of 101 players over 30 independent runs with $\beta = .04$.

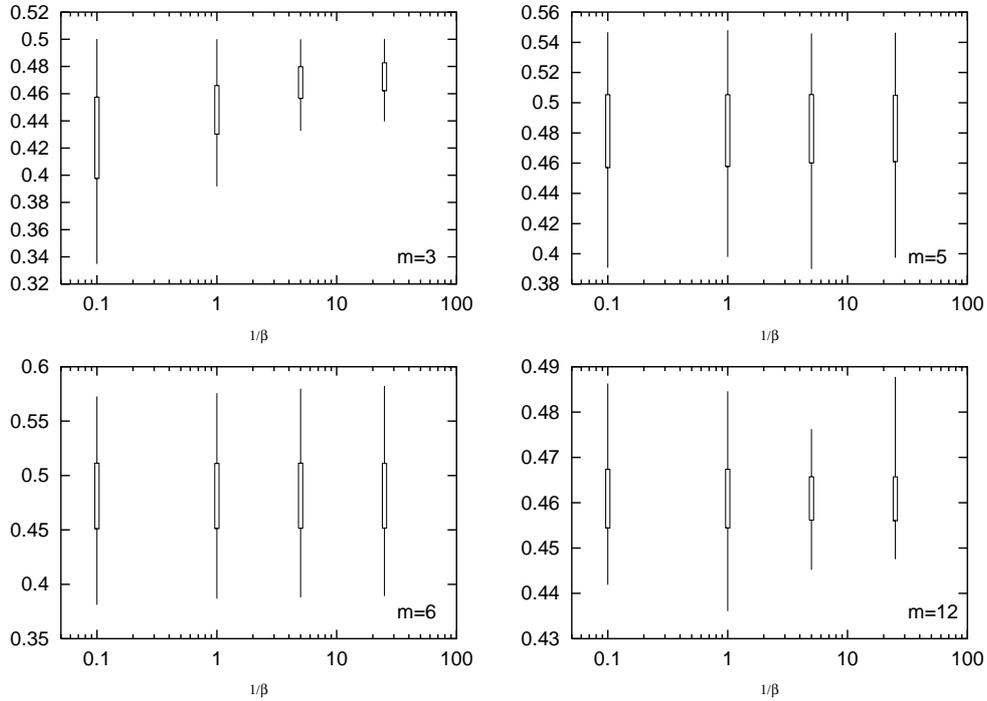


Figure 12: Variance (rectangle) and support (straight line) of the scored points distribution in a population of $N = 101$ with $s = 2$, over 30 independent runs. Notice that while in the high β simulations the distributions are similar in width for any m , when β is reduced the low m region emerges as the “socially optimal” one.

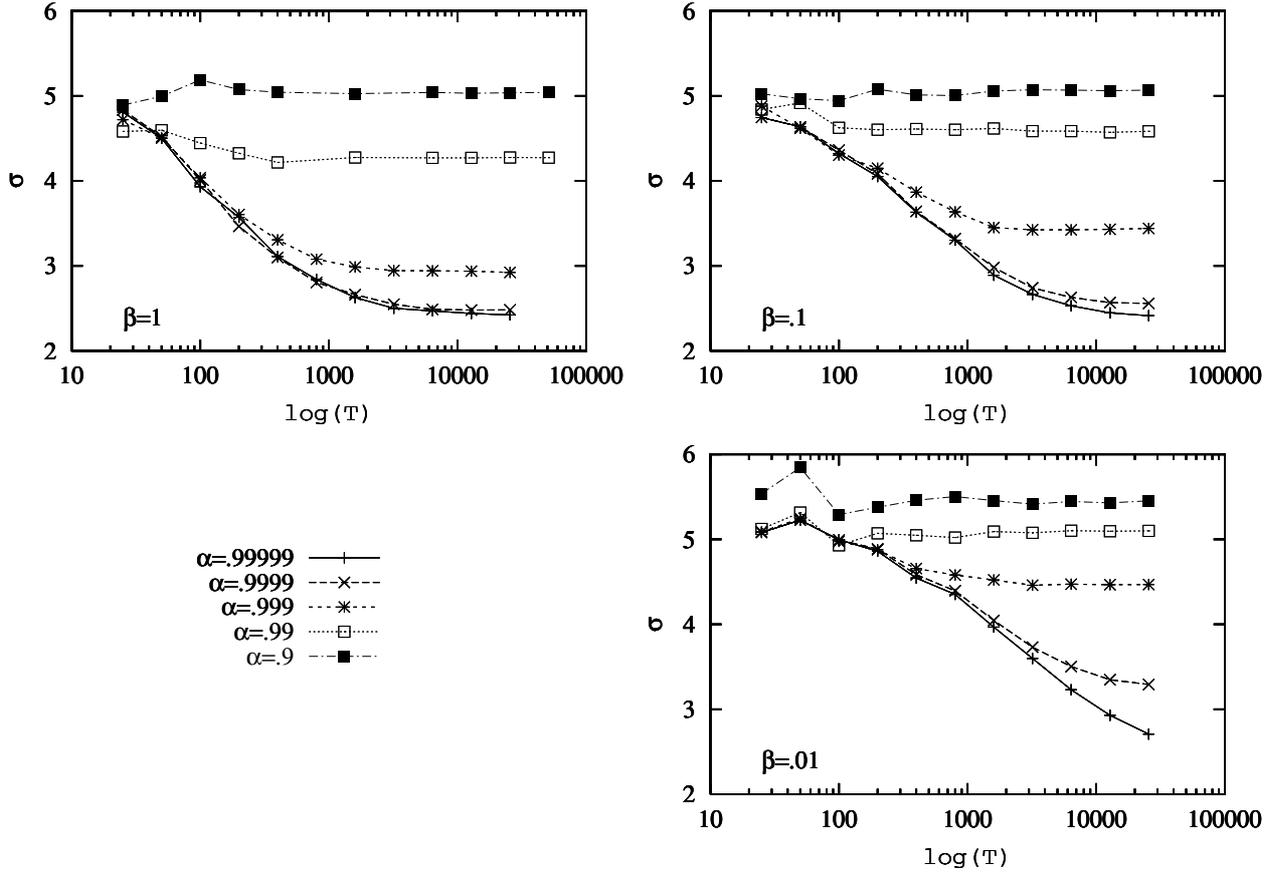


Figure 13: σ as a function of the run length T for different values of β and α . The simulations are performed with $m = 6$ where a greater sensitivity of the transient time length to the learning parameters β and α is expected (see Section 4.2).