Global Optimization of Minority Game by Smart Agents

Yan-Bo Xie¹, Bing-Hong Wang¹, and Tao Zhou¹
¹Nonlinear Science Center and Department of Modern Physics,
University of Science and Technology of China, Hefei, 230026, PR China
²Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan
(Dated: March 22, 2022)

We propose a new model of minority game with so-called smart agents such that the standard deviation \( \sigma^2 \) and the total loss in this model reach the theoretical minimum values in the limit of long time. The smart agents use trial and error method to make a choice but bring global optimization to the system, which suggests that the economic systems may have the ability to self-organize into a highly optimized state by agents who are forced to make decisions based on inductive thinking for their limited knowledge and capabilities. When other kinds of agents are also present, the experimental results and analyses show that the smart agent can gain profits from producers and are much more competent than the noise traders and conventional agents in original minority game.

PACS numbers: 02.50.Le, 05.65.+b, 87.23.Ge, 87.23.Kg

I. INTRODUCTION

The minority game (MG) models was introduced by Challet and Zhang in 1997 as a model for the competition for limited resources¹, which have attracted much attention in recent years. The basic scenario is easy to explain: there is a population of \( N \) players who, at each time step, have to choose either 0 or 1. Those who are in the minority win, the other lose (to avoid ambiguities, \( N \) is chosen to be odd). The agents make their decisions based on the most recent \( m \) outcomes, thus there are \( 2^m \) different histories. A strategy is defined as a table of \( 2^m \) choices (either 0 or 1) for the \( 2^m \) corresponding histories, so that there are \( 2^{2^m} \) different strategies in the strategy-space. Each agent randomly picks \( s > 1 \) strategies from the strategy-space in the beginning of the MG. To each strategy is associated an integral point, which initially takes the value 0 and will increase by 1 at each time step if it predicts the result correctly. Each agent uses the one with the highest point among his \( s \) strategies, if there are several strategies with the same highest point, one of those will be chosen randomly. A very important quantity in this model is the overall loss defined as

\[
L(t) = N_{\text{loss}}(t) - N_{\text{win}}(t) \geq 1
\]

where \( N_{\text{loss}} \) and \( N_{\text{win}} \) are, respectively, the number of losers and the winners at time \( t \). Apparently, the smaller \( L(t) \) is, the better the system performs. Another related quantity is called the standard deviation and defined as

\[
\sigma^2(t) = (n_0(t) - \bar{n})^2
\]

where \( n_0 \) is the number of agents who choose 0 and \( \bar{n} = N/2 \). It is easy to see that \( \sigma^2(t) = L^2(t)/4 \) and theoretically, the minimum value of \( \sigma^2(t) \) is 0.25.

One of the focuses of scientists' attention is the problem how to improve the performance of system, i.e. to reduce \( \sigma^2 \). Recently, some new kinds of agent are introduced²,³, by whom the overall performance of system is improved. A farther question is whether it is possible to achieve the global optimization in the framework of the MG model assuming that agents try to outsmart each other for their selfish gain and act based on inductive thinking¹.

Recently, a significant work is achieved by Reents, et al, who propose a stochastic minority game model in which \( \sigma^2 \) is minimized². In their model, an agent will not change his choice in the next time step if he wins in the present turn, on contrary, he will change his choice at probability \( p \). The value of \( p \) is the same for all the agents. When \( p < 1/N \), Reents et al, found that \( \sigma^2 \sim 0.25 \). However, the agents in real-life systems are not as clever as Reents, they do not know how to select a value of \( p \) and even do not know the total number of agents \( N \). Thus Reents's model may be not proper for the systems consisting of agent with inductive thinking.

Metzler and Horn have introduced the evolution into the stochastic minority game model⁴. Similarly to the evolutionary minority game model⁵, for an arbitrary agent \( i \), a probability \( p_i(t) \) and a score \( s_i \) is equipped⁶. The score \( s_i \) increases by 1 if the agent wins and decreases by 1 if the agent loses. When \( s_i \leq d < 0 \), the agent is deceased and replaced by a new agent with a reset score \( s_i = 0 \). If \( p_i(t) \) of the new agent is randomly distributed in \((0,1)\), the average value of \( p_i(t) \) in the final stationary state is found to be at the order of 1 and thus \( \sigma^2 \sim O(N^2) \). They also discussed the situation in which the new agent chooses \( p_i(t) \) by copying the value of \( p_j(t) \) of another agent who is randomly selected. Within this scheme, it is possible to see that \( p_i(t) \sim O(1/N) \) and \( \sigma^2 \sim 1 \) in sufficiently long time. However, it is still unreasonable to assume that an agent knows the information of all other agents. Furthermore, \( p \) is at the order of 1/N and thus \( \sigma^2 \) is greater than 0.25 in the final state. The best solution is still not achieved in their model.

In the present paper, we propose a new model of minority game with so-called smart agents such that the...
standard deviation $\sigma^2$ and the total loss in this model reach the theoretical minimum values in the limit of long time. The smart agents act based on inductive thinking but bring global optimization to the system. Experimental results and analyses show that when other kinds of agents are also present, the smart agent can gain profits from producers and are much more competent than the noise traders and conventional agents in original minority game.

II. MODEL AND NUMERICAL SIMULATION

Our model consists of $N$ agents with $N$ an odd integer. Each agent has only one strategy which evolves with the following rule: suppose at a given time step $t$, the memory (history) is $\mu$ and the strategy of the $i$-th agent is $s_i(t, \nu)$ for $\nu = 0, ..., 2^M - 1$. Also, each agent has a probability function $p_i(t, \nu)$ for $i = 1, ..., N$ and $\nu = 0, ..., 2^M - 1$. If the $i$-th agent wins at $t$, the strategy will not be changed; contrarily, with probability $1 - p_i(t, \mu)$, $s_i(t, \nu)$ is not changed, with probability $p_i(t, \mu)$, $s_i(t+1, \mu) = 1 - s_i(t, \mu)$, but $s_i(t+1, \nu) = s_i(t, \nu)$ for all other $\nu \neq \mu$.

The initial value of $p_i(t, \nu)$ is randomly selected in $(0, 1)$ and evolves by self-teaching mechanism, which is the simplest trial and error method. For a given time step $t$ with history $\mu$, consider the last time step $t'$ when the memory is also $\mu$. If the agent $i$ won at $t'$ or he loses but does not change $s_i(t', \mu)$, then no changes will occur. Otherwise, $p_i(t, \mu)$ will change according to the following rule:\[\]

$$p_i(t+1, \mu) = \begin{cases} \min(1, 2p_i(t, \mu)) & \text{agent } i \text{ wins at time } t \\ p_i(t, \mu)/2 & \text{agent } i \text{ loses at time } t \end{cases}$$

No changes will occur for all $p_i(t, \nu)$ with $\nu \neq \mu$.

Note that the evolution of $p_i(t, \mu)$ for different memories is essentially decoupled in our model. Therefore, mathematically speaking, the $m \neq 0$ case is a trivial generalization of the $m = 0$ case. The reason why we introduce different memories here is to mimic the behavior of the agents in real-life markets that the agents study the selection rules for different memories in order to find some regularities.

Figure 1 shows the simulation results, which indicate that the system will reach global optimization in sufficiently long time. We have checked that the property of time evolution of $\sigma^2(t)$ for the cases with more agents and larger memory is the same as that of $N = 101$ and $m = 0, 1, 2$.

![FIG. 1: Time evolution of $\sigma^2(t)$ for $N = 101$ smart agents with $m = 0$ (a), 1 (b), and 2 (c). The value of $\sigma^2(t)$ shown in this figure is the average of 10 independent experiments and the horizontal line represents $\sigma^2 = 0.25$.](image)

for two or more agents changing their strategies at the same time is negligibly small) thus the number of agents on the majority side is always $(N + 1)/2$. Therefore, the agent who changes the strategy is from the losing side to the losing side and $p_i(t)$ is reduced by a factor of 2. Since $p_i(t) \ll 1/N$, the probability that one agent will change his strategy is

$$\eta = 1 - \prod_{i \in W_t(t)} (1 - p_i(t)) \approx \sum_{i \in W_t(t)} p_i(t) \approx \frac{G(t)}{2}$$

where $W_t(t)$ is the set of losers at time $t$. Then we have the iterative equations for $G(t)$ and $H(t)$:

$$G(t + 1) = \eta \cdot \frac{2N - 1}{2N} G(t) + (1 - \eta) G(t) \quad (4)$$

$$H(t + 1) = \eta \cdot \frac{H(t)}{2} + (1 - \eta) H(t) \quad (5)$$

According to Eq. (4)&(5), one can find that $G(t) \sim t^{-1}$ and $H(t) \sim t^{-N/10}$, which is consistent with the simulation results shown in figure 2.
Before it switches to another case, the equilibrium state stays in one case for a period of time, called the life time. The life times of two cases are different. Assume that the probability $p_i$ of agent $i$ is independent of $i$, then the life time of the former case is $\tau_1 = 2/(N_s - \Delta + 1)(p)$ and the latter case is $\tau_2 = 2/(N_s + \Delta + 1)(p)$, where $p$ denotes the average value of $p_i$. The overall gain of the smart agents at each time step is equal to

$$
\Sigma = \frac{1}{\tau_1 + \tau_2} \left[ \left( \frac{N_s + \Delta - 1}{2} - \frac{N_s - \Delta + 1}{2} \right) \tau_1 + \left( \frac{N_s - \Delta - 1}{2} - \frac{N_s + \Delta + 1}{2} \right) \tau_2 \right]
\approx \frac{1}{N_s + 1} [\Delta^2 - 1 - N_s]
$$

Therefore, $\Sigma > 0$ when $\Delta < N_s < \Delta^2 - 1$. The average profit gained by each smart agent at each time step is

$$
\Sigma \approx \frac{1}{N_s(N_s + 1)} [\Delta^2 - 1 - N_s]
$$

According to Eq.(7), when $N_s < \Delta^2 - 1$, each smart agent can gain profits from producers. Suppose the number of smart agent $N_s$ is not fixed, if $N_s < \Delta^2 - 1$, some new smart agents, if available, will join the game since they can gain profits from producers. Thus there will be eventually $N_s \approx \Delta^2 - 1$ smart agents in the market, whose profits are approximatively equal to 0 with slight fluctuation. This process can be considered as an example for the efficient market hypothesis (EMH), which is hotly controversial in the recent years\[13\]. But in real-life financial market, the number of producers is alterable, thus the equilibrium state can rarely be reached.

When $m > 0$, the number of possible histories is $2^m > 1$. For a given history $\mu$, suppose $N_{p0}(\mu)$ producers always choose 0 and $N_{p1}(\mu)$ producers always choose 1. Then $\Delta(\mu) = N_{p0}(\mu) - N_{p1}(\mu)$ is a function of $\mu$. Since different history $\mu$ is essentially decoupled in our model and the number of smart agents $N_s$ is fixed, there may be three cases under history $\mu$: (i) $|\Delta(\mu)| \geq N_s$, each smart agent can gain one point at each time step; (ii) $\Delta^2(\mu) - 1 > N_s > |\Delta(\mu)|$, the smart agents can averagely gain profit from the producers; (iii) $\Delta^2(\mu) - 1 < N_s$, the smart agent cannot gain profit and are characterized by the overall loss described by Eq.(1).

The above picture is confirmed by the numerical simulation result shown in Figure 3(a). One can find that $\sigma^2$ decreases as $t$ increases and decays to 0.25 when $t$ is sufficiently large. Figure 3(b) plots the time dependence of the mean gain for smart agents:

$$
A_s(t) = \frac{N_{s\text{win}}(t) - N_{s\text{lose}}(t)}{N_s}
$$

where $N_{s\text{win}}$ and $N_{s\text{lose}}$ denote the number of smart agents who win and lose, respectively. Initially, $A_s(t)$ is negative, but as $t$ increases, $A_s(t)$ becomes positive. Therefore, the smart agents can gain profits from producers in the regime $\Delta^2(\mu) - 1 > N_s$.

### III. SMART AGENTS IN MIXED MARKET

Challet et al classified the agents into three different types\[11\]: producers who have only one strategy, speculators (conventional agents in original minority game) who have two or more strategies, and the noise traders who make their choices by random tosses. In this section, we will investigate how smart agents perform in mixed market\[12\].

Firstly, let us look into how the smart agents compete with the producers. Assume that there are $N_p$ producers and $N_s$ smart agents with $N_p + N_s$ an odd integer, each producer has only one fixed strategy. For simplicity, we shall first discuss the case of $m = 0$. Suppose $N_{p0}$ producers always choose 0, and $N_{p1}$ producers always choose 1. If $\Delta = N_{p0} - N_{p1} > N_s(< -N_s)$, then $N_s$ smart agents must choose $1(0)$ in the equilibrium state and win at each time step. When $N_s > \Delta > 0$ (the case $N_s > -\Delta > 0$ is analogic), the situation is slightly complicated. From the discussion in section 2, it is not difficult to see that the overall loss of $N_p + N_s$ agents is minimized in the equilibrium state. Namely, there will be either $(N_s - \Delta + 1)/2$ smart agents choosing 0 and $(N_s + \Delta - 1)/2$ smart agents choosing 1 or $(N_s - \Delta - 1)/2$ smart agents choosing 0 and $(N_s + \Delta + 1)/2$ smart agents choosing 1. In the former case, the agents choosing 0 are losers, while in the latter case, the agents choosing 0 are winners. The equilibrium state is described by the transition between two cases.

![FIG. 2: Time dependence of $G(t)$ (a) and $H(t)$ (b), where $N = 101$ and $m = 0$. The slopes of the two curves in figure a & b are $-1.01(\approx -1)$ and $-102(\approx -N)$, respectively.](image-url)
Assume that there are minority game model\[1\], and smart agents are present. Figure 3(a) shows the time dependence of $\sigma^2$, where $N_p = 200$, $N_s = 801$, $m = 1$ and $\Delta(0) = \Delta(1) = 200$. The value of $\sigma^2(t)$ and $A_s(t)$ shown in these two figures is the average of 32 independent experiments and the horizontal line in figure (a) represents $\sigma^2 = 0.25$.

Secondly, let’s consider the case in which the noise traders and smart agents are present. Assume that there are $N_n$ noise traders and $N_s$ smart agents with $N_n + N_s$ an odd integer. Figure 4(a) plots the time dependence of $\sigma^2$, one can find that $\sigma^2$ decreases as $t$ increases, but does not reach the theoretical Optimization 0.25 in the limit of long time. This result is not difficult to understand for the existence of noise traders will bring more fluctuations into the system. Figure 4(b) and 4(c) exhibit the time dependence of $A_s$ and $A_n$ respectively, where $A_n$ is the mean gain of noise traders:

$$A_n(t) = \frac{N_{\text{win}}(t) - N_{\text{lose}}(t)}{N_n}$$

$N_{\text{win}}$ and $N_{\text{lose}}$ denote the number of the noise traders who win and lose, respectively. Apparently, the smart agents perform much better than the noise traders do.

At last, we have studied the case in which the conventional agents, who take the actions based on the original minority game model\[1\], and smart agents are present. Assume that there are $N_s$ smart agents and $N_m$ conventional agents with $N_s + N_m$ an odd integer. Figure 5(a) shows the time dependence of $\sigma^2$. One sees that $\sigma^2$ decreases with time but also does not reach the theoretical Optimization 0.25 in the limit of long time. This result implies that the conventional agents also introduce fluctuations, though its magnitude is less than the noise traders.
We propose a new model of minority game with so-called smart agents, who use trial and error method to make a choice. When only the smart agents are present, it is found that the overall loss is minimized to the theoretical limit as \( \sigma^2 \rightarrow 0.25(t \rightarrow \infty) \). Notice that although those smart agents are independent and only trying to do their best for their selfish gain based on inductive thinking, the Global Optimization is achieved in our model. The result suggests that the economic systems may have become very small when the time is sufficiently large. It is worthwhile to emphasize that, the smarts agents perform much better than the conventional agents in mixed market. Imagine an agent trying to figure out the regularity of the financial market. Assume at time \( t_1 \), he has the selection rules for all possible histories, i.e., he has a strategy. At a later time \( t_2 \), he finds that the selection rules for some histories do not give profits. Therefore, he may change the selection rule for these history, but not for the other histories which still give him profits. This is in contrast with the original MG model in which an agents selects the strategy with the highest virtual point. When he changes the strategy, he may change many selection rules although they still make profits. We think that is the reason why the conventional agents are less competent than smart agents.

Acknowledgments

This work was supported by the State Key Development Programme of Basic Research of China (973 Project), the National Natural Science Foundation of China under Grant No.70271070, the China-Canada University Industry Partnership Program (CCUIPP-NSFC No.70142005), the Doctoral Fund from the Ministry of Education of China, and the National Science Council (Taiwan) under Grant No. NSC 92-2112-M-001-063.

\[ \eta(4), \text{note that } \eta(4) \text{ is easily to be understood}. \]

\[ \int G^{-2}dG = \frac{t}{\mu} dt, \text{then } \int G^{-2}dG \sim \int \frac{dt}{\mu}, \text{and we can obtain the order of } G(t): G(t) \sim t^{-1}. \] Analogously, we have \( \int \frac{dt}{\mu} \sim \int (\frac{t}{\mu} + C) dt, \text{and then } H(t) \sim t^{-N}, \text{where } C \text{ is a constant.} \]

[1] D. Challet and Y.-C. Zhang, Physica A246, 407 (1997); Y.-C. Zhang, Europhys. News 29, 51 (1998); D. Challet and Y.-C. Zhang, Physica A256, 514 (1998).
[2] Y. Li, R. Riolo and S. Savit, Physica A276, 234 (2000).
[3] H.-J. Quan, B.-H. Wang, Z.-N. Tang and J. Liu, Preprint Cond-mat/0406165 (to be published in Int. J. Mod. Phys. B).
[4] H. Simon, Models of Bounded Rationality (MIT Press, Cambridge, 1997); W. B. Arthur, Amer. Econ. Assoc. Papers Proc. 84, 406 (1994).
[5] G. Reents, R. Metzler and W. Kinzel, Physica A299, 253 (2001).
[6] R. Metzler and C. Horn, Physica A329, 484 (2003).
[7] N. F. Johnson, P.-M. Hui, R. Jonson and T. S. Lo, Phys. Rev. Lett., 82, 3360 (1999); T. S. Lo, S. W. Lim, P.-M. Hui and N. F. Johnson, Physica A287, 313 (2000).
[8] If agent \( i \) wins at time \( t \), he will think that the change at \( t' \) was advisable and therefore \( p_i(t+1, \mu) \) is increased; otherwise he will think that the change was too harried, therefore \( p_i(t+1, \mu) \) is reduced.
[9] In Eq.(4), note that \( \eta = G(t)/2, \) we have \( G^{-2}dG = \frac{1}{\mu^2} dt, \) then \( \int G^{-2}dG \sim \int \frac{dt}{\mu}, \) and we can obtain the order of \( G(t) \): \( G(t) \sim t^{-1}. \) Analogously, we have \( \int \frac{dt}{\mu^2} \sim \int (\frac{t}{\mu} + C) dt, \) and then \( H(t) \sim t^{-N}, \) where \( C \) is a constant.