Deceptive Strategies for the Evolutionary Minority Game

Garrison W. Greenwood

***Abstract*— The evolutionary minority game is extensively used to study adaptive behavior in a population of interacting agents. In time the agents self-organize despite the fact agents act independently in choosing how to play the game and do not know the play of any other agent. In this paper we study agents who collude with each other to play the same strategy. However, nothing prevents agents from being deceptive and playing a different strategy instead. It is shown that deceptive strategies can be profitable if the number of deceptive agents is small enough.**

I. Introduction

The minority game (MG) was originally introduced by Challet and Zhang [1] to help study adaptive behavior in a population of interacting agents. At each time step an odd number of agents $N \geq 3$ simultaneously chooses to be in either room '0' or room '1'. Only those on the minority side win. Winners receive a reward while the losers receive nothing. Each agent has s strategies to choose from but the strategy that has performed best since the beginning of the game is the one used at each time step. Each agent can remember the outcome of the previous M plays. The agents do not know how big the minority was; agents only know which of the two choices was in the minority.

The MG model is simple, yet remarkably capable of modeling many real-world situations. For example, in financial markets traders must choose between buying and selling. Drivers must choose between one of two possible routes to get to work. In both examples being in the minority is most advantageous: buyers can find real bargains if most traders are selling and drivers on the route with less traffic get to work faster. Every agent makes a choice, but the payoff depends on what choices the other agents make.

Agents often consider recent history before making decisions about their next play. For instance, a driver might pick route A instead of route B because the traffic has been less on route A over the last three days. Table I shows one possible agent strategy for a memory size of $M = 3$. The two choices are denoted by '0' and '1' and the table indicates what the next choice (or play) should be for a given history of M outcomes. For example, if the past three winners (minorities) were '1 0 1', then this strategy says on the next time step the agent should play '0'. It is worth noting that searching for an optimal strategy—assuming one exists—is nearly impossible because the strategy search space is hyper-exponential in size[1].

Garrison Greenwood is with the Electrical & Computer Engineering Department at Portland State University, Portland, OR 97207–0751 USA (email: greenwd@ece.pdx.edu)

[1]With a memory size of M there are 2^M possible previous outcomes yielding 2^{2^M} possible strategies. A search space for $M = 5$ has over 4 billion strategies.

Challet and Zhang [2] later introduced the so-called evolutionary minority game (EMG) where agents are allowed to alter their strategies. Each agent has $s = 2$ strategies. Both strategies are evaluated at each time step but only the best performing one is actually used. Agents can adapt by changing to a different strategy every few time steps. These new strategies are created by first cloning good performing strategies and then mutating them with a small probability. (More details are given in the next section.)

Previous 3 outcomes	Next Play
0 0 0	0
0 0 1	1
0 1 0	1
0 1 1	0
1 0 0	1
1 0 1	0
1 1 0	0
1 1 1	1

TABLE I
ONE POSSIBLE STRATEGY WITH MEMORY SIZE $M = 3$.

A thorough literature search found little prior EA-based work on the MG and what does exist hasn't contributed anything noteworthy. The EA-based methods kept the basic Challet and Zhang EMG model and only changed how the replacement strategies were found. For example, Sysi-Aho et. al [3] used 1-pt crossover on the s strategies to get the replacements. Yang et. al [4] used a cut-and-splice operator instead of crossover. Neither paper offered any new insights into minority game behavior.

Johnson et. al [5] used an entirely different to evolving strategies. Each agent a_j used a fixed strategy but it also has a gene giving the probability of using that strategy. Suppose the previous three outcomes were '0 0 1' and a_j plays the strategy shown in Table I. Then a_j plays '1' with probability p_j and plays '0' with probability $1 - p_j$. Agents can change their gene values through an evolutionary process. Each time step the agent scores +1 (-1) if he wins (loses). If the score drops below a pre-defined $d < 0$, then $p_j \leftarrow p_j + \triangle p_j$ where $\triangle p_j$ is a zero mean Gaussian distributed random variable with reflective binary conditions to keep the new p_j on the unit interval[2]. Surprisingly they discovered, despite the initial distribution of probabilities, over time the probabilities evolve to the two extremes—i.e., agents play either $p = 0$ or $p = 1$ (see Figure 1).

[2]If $p = 0.1$ and the random variable value $\triangle p = -0.3$, then a reflective boundary condition makes the new probability $p = 0.2$.

978-1-4244-4815-9/09/$25.00 ©2009 IEEE

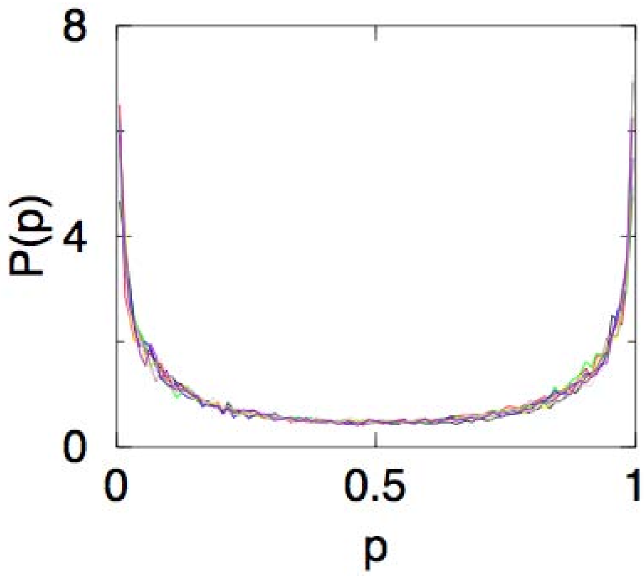

Fig. 1. Distribution of probabilities $P(p)$ for different M values. The distribution is insensitive to the M value. Results shown are for $N = 101$ and $d = -4$. (See text.)

The MG has been extensively studied with most of the published work appearing in physics journals. Most MG versions don't permit agents to disclose any information about their strategies—i.e., the only global information is the previous M outcomes of the game—although some researchers have studies the role of social networks (see Section II-B). In this paper we take a whole new approach to the MG. Here we allow agents to collude by agreeing to play the same strategy. We also allow deception by letting agents agree with their co-conspirators to play one way, but then, in their own self-interest, play just the opposite way. We will see that under certain circumstances collusion with deception can be profitable.

II. The EMG With and Without Collusion

A. The Basic EMG Model

The basic EMG model gives each agent s randomly generated strategies. Most researchers have found $s = 2$ is sufficient for a serious investigation and the dynamics don't change significantly for higher values. Agents have a memory of size M to keep track of the recent M outcomes of the game. This personal memory forms the basis of the agent's strategy on what to play in the future. Table I shows a strategy for $s = 1$. Adding more "next play" columns will accommodate $s > 1$ strategies.

After each time step a strategy receives a virtual point if the suggested play was the minority play. Strategies do not have to be actually used to receive virtual points. Agents always play the strategy with the highest number of virtual points and they receive a real point if the strategy physically played puts the agent in the minority. The agent's number of real points serves as a measure of fitness. Every τ time steps the worst fit agent is replaced by a clone of the best fit agent. One of the strategies in the clone can be mutated with a small probability.

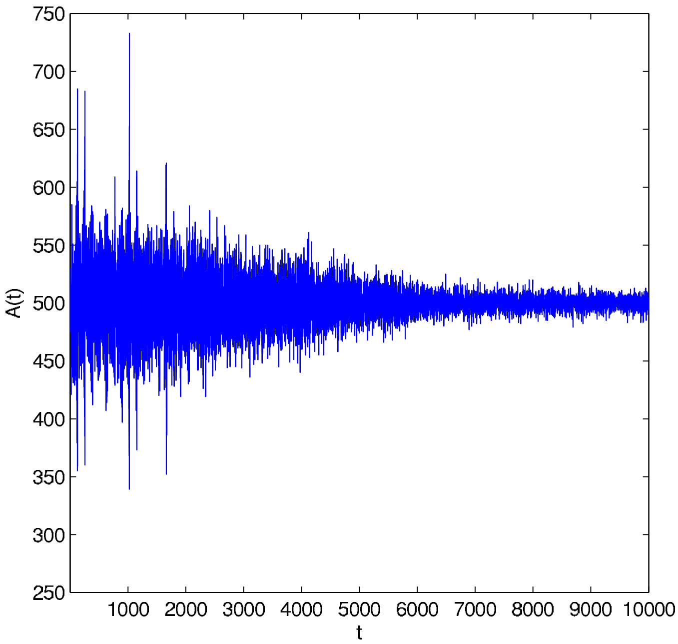

Fig. 2. Temporal attendance for the genetic approach with $N = 1001$. The tapering indicates a learning process.

This process introduces some interesting game dynamics. We can visualize these dynamics by plotting the attendance in room '0' for a population of $N = 1001$ agents. As shown in Figure 2, at first the attendance fluctuates wildly but later on the agents self-organize and the attendance varies over a narrow range centered around $N/2$. Learning has emerged thereby improving the average gain for all agents. This self-organization is completely unexpected since it requires cooperation and coordination which is achieved without any explicit contact between agents.

The mean of the time series shown in Figure 2 is $N/2$ and the standard deviation, also called the *volatility*, is

$$\sigma \approx \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(A(t) - \frac{N}{2}\right)^2} \tag{1}$$

where $A(t)$ is the attendance at time step t and T is the total number of time steps. For a completely random choice game—i.e., where each agent randomly picks a room with equal probability—the standard deviation is $\sqrt{N/4}$. Studies have shown σ for the EMG is higher than that of the random choice game for small memory sizes ($M < 6$) [6]. However, the EMG σ will be smaller for $M \geq 6$. Consequently, $M = 6$ is used in our investigation.

Figure 3 shows the volatility is remarkably small when the agents are self-organized. What is significant about a small attendance range (or equivalently, a small volatility σ)? The largest possible minority set size is N/2, so a small attendance range around N/2 means the minority is consistently close

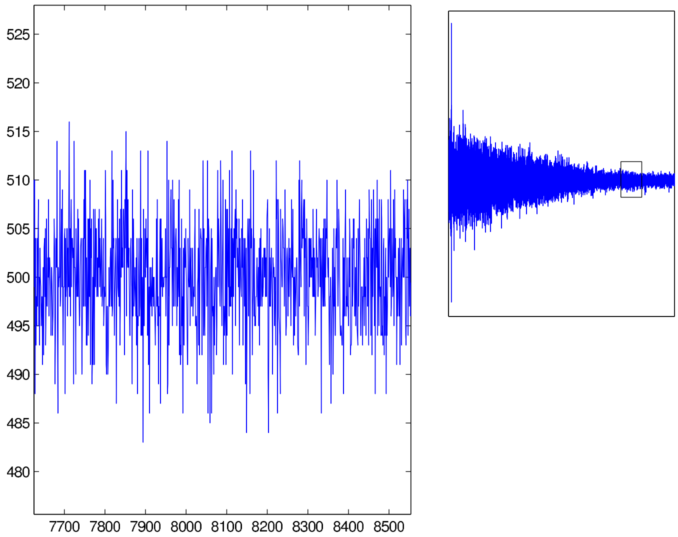

Fig. 3. Attendance with a self-organized population of 1001 agents. The figure on the left is a blowup of the boxed area in the figure on the right.

to its maximum size. Suppose the EMG models a financial market with 0/1 representing buy/sell decisions, Large fluctuations imply a highly volatile market where the minority size is small. Such a market is inefficient because relatively few investors are in the minority, which means relatively few investors are making any money. On the other hand, small fluctuations imply an efficient market with low risk where many investors can be successful despite having very different investment strategies.

B. The EMG with Collusion

For the most part researchers have assumed the past M outcomes is the only information publicly available to all agents. Agents know if room '0' was the minority play, but they do not know the exact number of agents who played room '0'. A few researchers investigated small groups of agents, organized into social networks, who disclose their choices to each other (e.g., see [7], [8], [9]). They found, on average, social networks are beneficial so long as they are not too large.

Our work described here is distinguished from these previous social network studies in two very important ways. First, in those previous studies agents knew the choices other agents in their network made and exploited that information to influence—but not necessary change—how they might play at the next time step. In other words, an agent's strategy was based on both global information (the last M outcomes of the game) and local information (how members of their social network last played). Agents played one of their s individual strategies as normal. Second, agents were honest in the sense they truthfully disclosed what they played on the previous time step; no deception was permitted or tolerated.

Agents in the real world act in their own self interest. Sometimes agents enter into a collusion with other agents hoping to profit from their secret cooperation. Unfortunately, integrity is not obligatory in these agreements [10]. Collusions frequently arise in financial markets. However, the outcomes are not always predictable and occasionally even backfire. In 1998 thirty security firms paid 910 million dollars to settle a class-action suit alleging price-fixing on America's Nasdaq stock exchange [11]. The sharp rise in oil prices during the summer of 2008 was widely attributed to the actions of a group of oil speculators [12].

In our work agents collude by all agreeing beforehand to play the identical strategy. This common strategy is the only shared knowledge among the colluding agents (other than the last M EMG outcomes). Agents do not report what they actually play, which is usually done in social networks. One other significant difference is honesty is not enforced in the social network. Colluding agents are told what they *should* play next, but agents may independently decide otherwise. In other words, an agent may deliberately deceive the others by agreeing to play as they do, but he is under no obligation to actually play as they do. Agents don't report what they actually play to the others, so traitorous behavior is kept secret.

An EMG with collusion is played as follows. Let $\mathcal{C}$ be a subset of agents that participate in a collusion with $|\mathcal{C}| = K$ and $K \ll N$. a_i is the leader of the collusion and a_{-i} represents all of the other agents in $\mathcal{C}$. a_i plays as normal—i.e., he selects the best of his s strategies and picks either room '0' or '1' based on the last M outcomes of the game. a_i broadcasts the designated room to play to a_{-i} at the beginning of each time step. All agents in $\mathcal{C}$ have previously agreed to play this designated room.

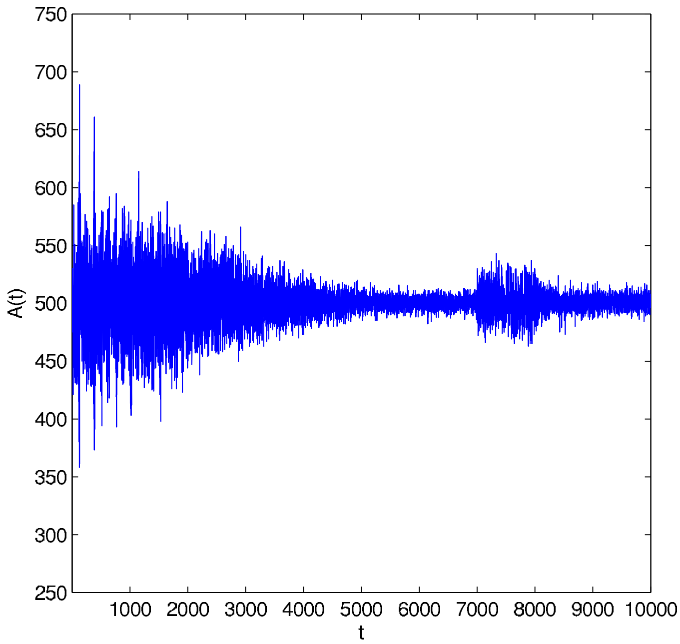

Fig. 4. Temporal attendance with collusion between time steps 7000 and 8000.

The term "general population" refers to the $N - K$ agents who are not participating in the collusion. An analogy frequently used is to think of this general population as traders in a financial market where the rooms '0' and '1' represent 'buy' or 'sell' actions. In a self-organized state the market is efficient with nearly $N/2$ agents making the same decision at each time step. That means a small number of agents who agree to all make the same decision at future time steps could easily upset this delicate balance. Figure 4 shows how collusion effects attendance. The collusion occurs between time steps 7000 to 8000, which is well after self-organization has taken place. Observe how σ increases when the collusion starts but eventually decreases once collusion stops as the agents return to a self-organized state.

The increased volatility caused by the collusion suggests agents who conspire to pick the same room may not do all that well because, by all choosing the same room, they might wind up in the majority. (We will come back to this point later.) An agent could thus conclude his chances of winning would improve by being deceptive. That is, the agent *says* he will go along with the group and pick the same room they do, but in reality he picks the other room.

To test the deception hypothesis an experiment was conducted with $N = 1001$ agents where $K = 30$ agents colluded. Agents play as normal except between time steps 7000 and 8000 where collusion took place; the general population continued to play as normal. Agents got +1 if they were in minority and -1 if they were in the majority. Initially there is no deception.

The average gain of the colluding agents and the average gain of the general population is plotted in Figure 5. A bifurcation occurs when the collusion begins showing that colluding agents on average do poorer than the general population indicating they consistently wind up in the majority. Conversely, the general population average increases during the collusion. This phenomena is easily explained. If the majority of the colluding agents are in the majority, then more than half the general population can be in the minority. A simple example is useful. Suppose 500 agents (out of the 1001) are in the minority. With $K = 30$ colluding agents there are 971 agents in the general population. Since the colluding agents consistently lose, they are in the majority. If all 30 colluding agents are in the majority, then all 500 agents in the minority are in the general population. Thus, the general population average is positive since $500 > 971/2$.

The dynamics change considerably when deception is allowed. If the non-deceptive colluding agents are consistently in the majority, then the deceptive agents are consistently in the minority. Figure 6 shows deception is not only a profitable strategy, but the profit can be considerable so long as the number of deceptive agents is not too large. Notice the deceptive agents outperform the general population.

Figure 3 can help explain why non-deceptive agents tend to wind up in the majority. Once the population is self-organized the minority size is near $N/2$. Notice in Figure 3, where $N = 1001$, attendance $A(t)$ is within ± 15 of the average $N/2$. Our experiments were conducted with $K = 30$, which is about 3% of the population. Certainly some of the colluding agents will be in the minority at any given time step *if they play their default strategy*. But these are colluding agents and they are forced to play a defined strategy if they live up to the agreement they made. Since the volatility is so small, only a fraction of the colluding agents would have to change their decision to flip the majority to the other room. These colluding agents would now all be in the majority and any deceptive agents would now be in the minority.

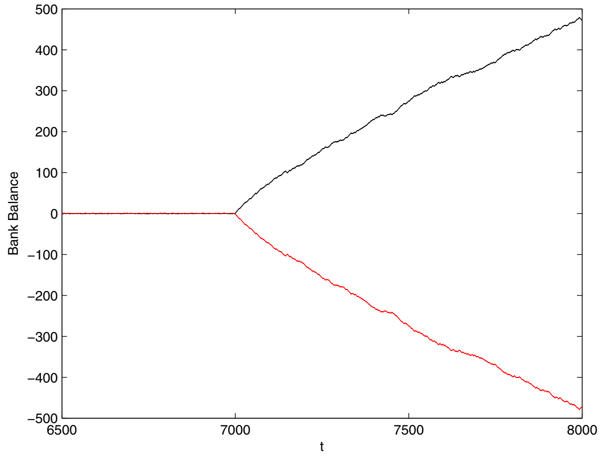

Fig. 5. Bifurcation

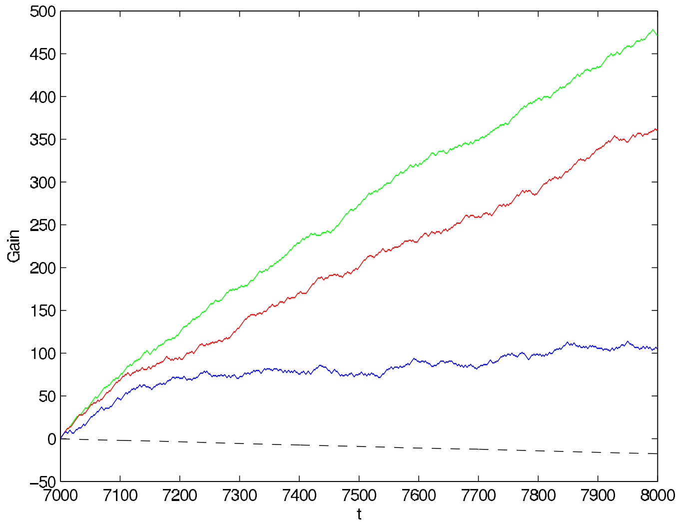

Fig. 6. Average returns for 3 (green), 7 (red) and 10 (blue) deceptive agents with 30 total colluding agents. The average return for the general population is shown as a dashed line.

C. Mixed Strategies

Of course any colluding agent, being equally intelligent, may surmise his chances of winning would improve by being deceptive. Remember agents do not know what choices other agents make, so any traitorous behavior is kept a secret. However, practically there are limits on how many agents can be deceptive. (If *every* agent chooses deception, then they *all* lose.) Figure 6 shows that the gains tend to decrease as the number of deceptive agents increases. This situation is expected because deceptive and non-deceptive agents cancel each other, which lowers the average gain of the deceptive agents.

Nevertheless, deception does have potentially high payoffs. It therefore might make more sense to adopt a mixed strategy where each colluding agent decides on deception with some probability. To test this hypothesis we added a gene to each agent whose value represents the probability of being deceptive. Each agent also keeps track of how successful this mixed strategy is and adapt the probability if he loses to often.

An experiment was set up where agents could choose deception from a discrete set of probabilities $p \in \mathcal{P} = \{0, 1/4, 1/2, 3/4, 1\}$. This allows an agent to never be deceptive ($p = 0$), always be deceptive ($p = 1$) or sometimes be deceptive. Each agent also keeps track of how successful this mixed strategy is by incrementing a register if he wins and decrementing the register if he loses. If the register value drops below some $d < 0$, then the agent resets the register value to zero and adapts the deception probability by randomly choosing some other probability in $\mathcal{P}$. Figure 7 shows the long-term evolution of the probabilities. This shows agents tend to eventually play the extremes, which supports the findings of Johnson et. al [5] (c.f. Figurefig6).

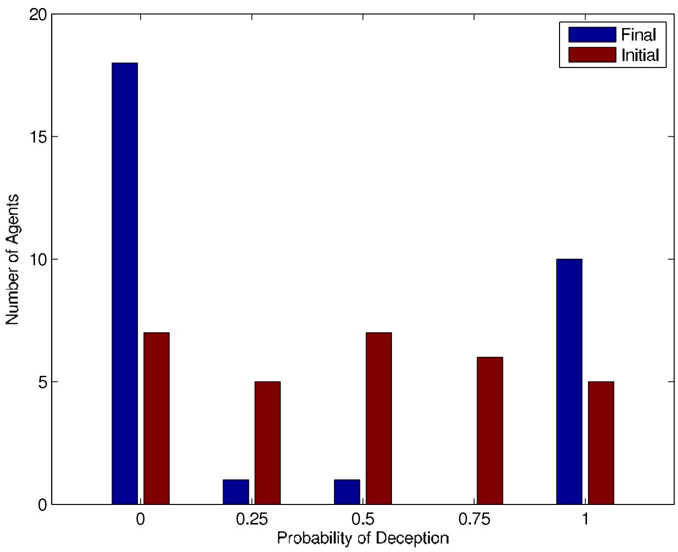

Fig. 7. Histogram of the long-term probability distribution. The red bars show the initial distribution for $N = 101$ agents. The blue bars show the final distribution.

D. Variable Payoffs

So far the payoff for winning is a reward of a fixed quantity (+1) and all winners get the same reward. However, in some games the payoff varies with the size of the minority—i.e., the smaller the minority, the larger the payoff. A good example are the lotteries run daily by many states in the United States. In Oregon state, for instance, the powerball lottery uses two sets of numbers: 59 in one set and 39 in the other. A player chooses five numbers from the first set and one number from the second set. Players pay a dollar to play. The lottery administrator randomly draws five numbers from the first set and one number from the second set. Any player who correctly picks all six numbers wins the jackpot[3]. The jackpot payoff is at least twenty million dollars. Multiple winners must split the payoff, so there is a distinct advantage in being part of a really, really small minority (preferably, one).

Two variable payoffs were considered [13]: $\mathcal{R}(x) = \frac{N}{2} - 2$ and $\mathcal{R}(x) = \exp(-\gamma \frac{x}{N})$ where $\mathcal{R}(x)$ is the reward for a minority size of x, $\gamma > 0$ is a constant and N is the number of players. Both payoffs provide very little reward as $x \rightarrow N/2$ but provide large rewards as x gets small. Figures 8 and 9 show the payoffs for 3, 5 and 10 deceptive agents. As expected, the larger the number of deceptive agents the smaller the payoff.

Figure 9 shows a monotonic increase in the gains of the deceptive agents except for the 10 deceptive agent case where initially there is some gain but eventually the agents start losing—even though they frequently wind up in the minority! The explanation is easy to understand. Agents don't necessarily win every time, which means to show a monotonically increasing gain the rewards for winning must be greater than the penalties for losing. This situation does occur when the reward and penalty is $+1$ and -1, respectively, so long as the agent wins more times than he loses. However, with variable payoffs the reward may not be high enough—even if the agent *does* win more often than losing! With only 5 (or less) defective agents the minority size is small enough to make up for the frequent losses. But with 10 deceptive agents this compensation doesn't occur, which means over time the average gain is flat or even possibly negative.

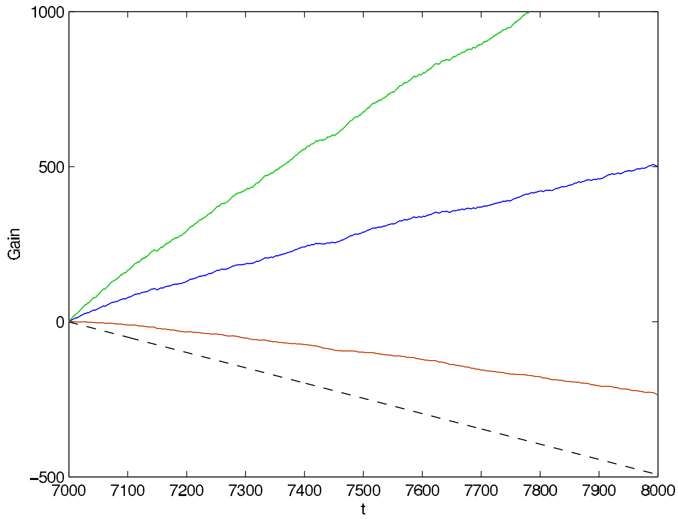

Fig. 8. Average payoff amongst 3 (green), 5 (blue) and 10 (red) deceptive agents for a $\frac{N}{x} - L$ payoff. The dashed line shows the average payoff of the general population. Collusion begins at time step 7000.

III. Discussion

The results shown above can be explained using the idea of "crowds" in the MG, an idea introduced by Hart et. al [14].

[3]Actually, players can get a payoff for choosing 3 or 4 numbers correctly, but the payoff is only a tiny fraction of jackpot payoff. In this paper we only consider winners to be those players who correctly picked all 6 numbers.

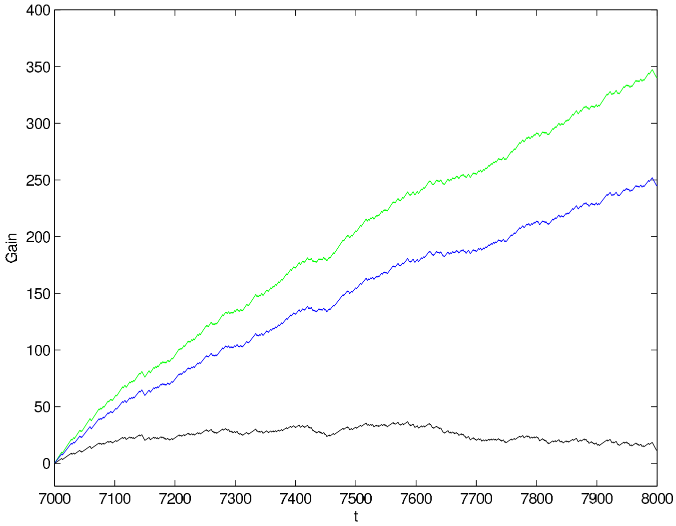

Fig. 9. Average payoff amongst 3 (green), 5 (blue) and 10 (black) deceptive agents for a $\exp(-\gamma \frac{x}{N})$ payoff with $\gamma = 0.75$. Collusion begins at time step 7000.

In the basic game there are 2^{2^M} total strategies but it is possible to get similar standard deviation in the attendance, and hence, similar dynamics, with a reduced strategy space of 2^M strategies. Two strategies are *anti-correlated* if they have maximal Hamming distance $d_H = 2^M$ and are *uncorrelated* if they have Hamming distance $d_H = 2^{M-1}$. A correlated and anti-correlated strategy forms a strategy pair $\mathcal{G}$. The reduced strategy space consists of uncorrelated strategy pairs—i.e., the strategies between $\mathcal{G}$ and $\mathcal{G}\prime$ are uncorrelated. Hence, strategies taken from distinct strategy pairs are really very different, which is important for getting into the minority. Hart et. al defined a *crowd* as the n_R agents playing the same strategy R and the *anti-crowd* as the $n_{\bar{R}}$ agents playing just the opposite strategy. This gives the strategy pair $\mathcal{G} \equiv (R, \bar{R})$.

Colluding agents play a strategy pair since the non-deceptive agents play a correlated strategy and the deceptive agents play the anti-correlated strategy. But our work differs significantly from Hart et. al. They used a restricted strategy space where every strategy is taken from $\{\mathcal{G}\}$. Conversely, in our work only the colluding agents play a strategy pair. The rest of the agents are free to choose any strategy they wish from a considerably larger strategy space of size 2^{2^M}. This larger strategy space means if a non-colluding agent plays a strategy U, there is nearly a zero probability any other agent plays $\bar{U}$. (Recall our studies were conducted with $M = 6$, which means the non-colluding agents choose from $2^{64} - 2$ strategies. With the Hart et. al studies, which uses the reduced strategy space, there would be only $2^7 = 128$ total strategies to choose from.) Moreover, the non-colluding agents can evolve new strategies whereas the strategies in Hart et. al are fixed. Nevertheless, we can still borrow some of their analysis techniques to help explain the behavior of colluding populations.

With crowds the volatility σ from (1) now becomes [14]

$$\sigma = \sqrt{\frac{1}{4} \sum_{\mathcal{G} \equiv (R,\bar{R})} |n_R - n_{\bar{R}}|^2} \tag{2}$$

In our work there is only one strategy pair (the colluding agents). Hence, (2) can be partitioned into

$$\sigma = \sqrt{\frac{1}{4} \sum_{\mathcal{G} \equiv (R,\bar{R})} |n_R - n_{\bar{R}}|^2 + \frac{1}{4} \sum_{U \notin \{\mathcal{G}\}} |n_U - n_{\bar{U}}|^2} \tag{3}$$

where the first term represents the colluding agents and the second term representing the non-colluding, general population agents. Given the large size of the strategy space it is reasonable to assume and $n_{\bar{U}} = 0$ for all of the $N - K$ non-colluding agents. Some of these agents are clones so $n_U \geq 1$. There are K total colluding agents. Let R be the strategy the colluding agents agree to play. Then n_R is the number of agents playing R, $n_{\bar{R}}$ the number of deceptive agents and $K = n_R + n_{\bar{R}}$. Then (3) can be re-written as

$$\sigma \approx \sqrt{\frac{1}{4}(K - 2n_{\bar{R}})^2 + \frac{1}{4} \sum_{U \notin \{\mathcal{G}\}} |n_U|^2} \tag{4}$$

Remember the non-colluding agents are self-organized so their contribution to the overall volatility is relatively small. There are colluding agents whenever $K > 0$ and (4) indicates the volatility should increase whenever collusion is present. This explains the increase seen between time steps 7000 and 8000 in Figure 4.

A larger volatility means a larger majority size. Non-deceptive colluding agents all pick the same room and thus are more likely to define the majority room. This means deceptive agents are more likely to be in the minority and thus gain more profit. When $n_{\bar{R}} = 1$ the volatility is highest, making it more likely the deceptive agents are in the minority. Conversely, every deceptive agent cancels a non-deceptive agent so increasing the number of deceptive agents decreases the volatilty. In fact, with the minimum volatility occurring when $n_{\bar{R}} = K/2$. This explains why a small number of deceptive agents on average gain more than a large number as depicted in Figures 8 and 9.

IV. Conclusions

We have investigated how collusion and collusion with deception affects returns in the EMG. Our results indicate

1) Collusion does not provide any inherent benefit. Indeed, without deception colluding agents tend to do on average worse than non-colluding self-organized agents.
2) Deception among colluding agents can be beneficial. Deceptive agents do better on average than both non-colluding agents and non-deceptive colluding agents so long as the number of deceptive agents is not too large.
3) Deception loses its advantage in lotteries unless the payoff for winning is sufficiently high enough to overcome losses. In fact, in some instances repeated wins may not

be sufficient to return a positive gain unless the minority size is quite small.

References

[1] D. Challet and Y. Zhang. Emergence of cooperation and organization in an evolutionary game. *Physica A*, 246:407–418, 1997.

[2] D. Challet, M. Marsili, and R. Zecchina. Statistical mechanics of systems with heterogeneous agents: minority games. *Phys. Rev. Lett.*, 84:1824–1827, 2000.

[3] M. Sysi-Aho, A. Chakraborti, and K. Kaski. Intelligent minority game with genetic crossover strategies. *Eur. Phys. J. B*, 34:373–377, 2003.

[4] W. Yang, B. Wang, Y. Wu, and Y. Xie. Searching good strategies in evolutionary minority game using variable length genetic algorithm. *Physica A*, pages 583–590, 2004.

[5] N. Johnson, P. Hui, R. Jonson, and T. Lo. Self-organized segregation within an evolving population. *Phys. Rev. Lett.*, 82(16):3360–3363, 1999.

[6] R.Savit, R. Manuca, and R. Riolo. Adaptive competition, market efficiency, and phase transitions. *Phys. Rev. Lett.*, 82(10):2203–2206, 1999.

[7] M. Remondino and A. Cappellini. Minority game with communication: an agent based model. In *Proc. 16th Euro. Simulation Symp.*, pages 155–160, 2004.

[8] H. Chau, F. Chow, and K. Ho. Minority game with peer pressure. *Physica A*, 332:483–495, 2004.

[9] I. Caridi and H. Ceva. Minority game with interactions. *Physica A*, 339:574–582, 2004.

[10] B. Flyvberg, M. Holm, and S. Buhl. Underestimating costs in public works projects: error or lie? *J. Amer. Planning Assoc.*, 68(3):279–295, 2002.

[11] Collusion in the stockmarket. *Economist*, 346:71, 1998. Issue 8051.

[12] D. Just and R. Just. Monopoly power, futures market manipulation and the oil price bubble. *J. Agricultural & Food Ind. Org.*, 6:Article 2, 2008.

[13] Y. Li, A. Vandeemen, and R. Savit. The minority game with variable payoffs. *Physica A*, 284(1–4):461–477, 2000.

[14] M. Hart, P. Jefferies, N. Johnson, and P. Hui. Crowd-anticrowd theory of the minority game. *Physica A*, 298(3–4):537–544, 2001.