Strategy updating rules and strategy distributions in dynamical multiagent systems

Shahar Hod[1] and Ehud Nakar[2]

[1]*Department of Condensed Matter Physics, Weizmann Institute, Rehovot 76100, Israel*
[2]*The Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel*

In the evolutionary version of the minority game, agents update their strategies (gene value p) in order to improve their performance. Motivated by the recent intriguing results obtained for prize-to-fine ratios, which are smaller than unity, we explore the system's dynamics with a strategy updating rule of the form $p \to p \pm \delta p$ ($0 \leq p \leq 1$). We find that the strategy distribution depends strongly on the values of the prize-to-fine ratio R, the length scale δp, and the type of boundary condition used. We show that these parameters determine the amplitude and the frequency of the temporal oscillations observed in the gene space. These regular oscillations are shown to be the main factors which determine the strategy distribution of the population. In addition, we find that the agents characterized by $p = \frac{1}{2}$ (a coin-tossing strategy) have the best chances of survival at asymptotically long times, regardless of the value of δp and the boundary conditions used.

The minority game (MG) is a successful model describing a population of competing and evolving individuals. The evolutionary version of the model (EMG) was introduced and widely explored by Johnson *et al.* [1]. This complex system has been explored extensively in the last few years, see e.g., Refs. [2–21] and references therein. The present work is mainly motivated by the recent results of Refs. [20,15].

In this toy model, a population of N agents with limited information and capabilities repeatedly competes for a limited global resource, or to be in the minority. The desire to be in a minority group is found in many real life situations, such as financial markets, traffic jams, or among a group of predators (who wish to hunt in areas with fewer competitors).

At each round of the game, every individual has to choose whether to be in room "0" (e.g., choosing to sell an asset) or in room "1" (e.g., choosing to buy an asset). At the end of each turn, the agents belonging to the smaller group (the minority) are the winners, each of them gains R points (the "prize"), whereas the others loose a point (the "fine"). The agents share a common look-up table containing the outcomes of recent occurrences. This allows the determination of a "predicted trend" in the system, which is followed by each agent with probability p, known as the agent's "gene" value.

In the evolutionary formulation of the model (EMG), the agents are allowed to evolve their strategies based on the past experiences. If an agent's score falls below some value d, he modifies its gene value. In this sense, each agent tries to learn from his past mistakes and to adjust his strategy in order to perform better.

A remarkable conclusion deduced from the EMG [1] is that a population of competing agents tends to *self-segregate* into opposing groups characterized by extreme behavior. It was realized that in order to flourish in such situations, an agent should behave in an extreme way ($p = 0$ or $p = 1$) [1]. On the other hand, in many real life situations, the prize-to-fine ratio may take a variety of different values [15,13]. A different kind of strategy may be more favoralbe in such situations. In recent studies, it was found [15] that an intrigu-

ing phase transition exists in the model; "Confusion" and "indecisiveness" take over when the prize-to-fine ratio falls below some critical value, in which case the agents characterized by a coin-tossing strategy ($p = \frac{1}{2}$) perform better than the extreme ones. In such circumstances, the agents tend to *cluster* around $p = \frac{1}{2}$ (see Fig. 1 of Ref. [15]) rather than self-segregate into two opposing groups.

In Ref. [15], we have considered a *uniform* strategy updating rule in which the new strategy is chosen uniformly within the range $0 \leq p \leq 1$. Burgos, Ceva, and Perazzo [20] have recently considered the same model problem with an updating rule of the form $p \to p \pm \delta p$, where $\delta p < \frac{1}{2}$, and found that the population tends to form an *M*-shaped strategy distribution in the $R < 1$ case. In the present work, we further explore this system and provide some insights that extend and link the results of Ref. [15] to those of Ref. [20].

First, we would like to stress the importance of the chosen *boundary conditions* in the case of an updating rule of the form $p \to p \pm \delta p$ [1]. Figure 1 displays the long-time averaged gene distribution $P(p)$ of the agents for two different types of boundary conditions: Periodic and reflective. One finds that for periodic boundary conditions, the population tends to cluster at intermediate gene values. The curve between the two peaks, located at $p = \delta p$ and $p = 1 - \delta p$, is almost *flat*, while the agents with extreme gene values ($p \simeq 0$ and $p \simeq 1$) perform much worse (we shall shortly demonstrate that the gene distribution may also have an inverse-*U* shape, depending on the precise values of R and δp). On the other hand, the gene distribution is almost flat for reflective boundary conditions.

The underlying mechanism that is responsible for this important difference is the temporal oscillations observed in the winning probabilities of the agents [15,16]. Figure 2 displays the time dependence of the winning probability of a $p = 0$ agent (the winning probability of a central agent, with $p = \frac{1}{2}$, is practically constant in time). We consider three distinct cases, characterized by (i) $\delta p = 0.1$ with periodic boundary conditions, (ii) $\delta p = 0.1$ with reflective boundary conditions, and (iii) uniform updating rule. One finds smaller oscillation amplitudes and longer periods for reflective
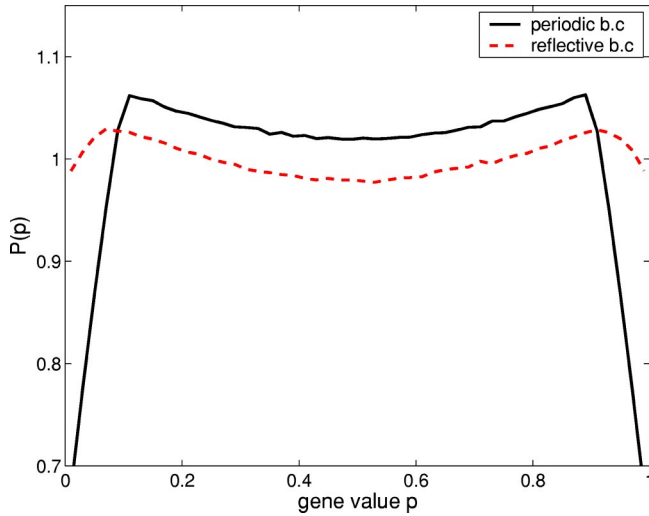
FIG. 1. The strategy distribution $P(p)$ for periodic boundary conditions (solid line) and reflective boundary conditions (dashed line). The results are for $N = 10\,001$ agents, $R = 0.8$, $d = -4$, and $\delta p = 0.1$. Each point represents an average value over 10 runs and 100 000 time steps per run.

boundary conditions, as compared to the case of periodic boundary conditions. This implies that for reflective boundary conditions, the performance of extreme agents ($p=0$ and $p=1$) becomes quite similar to the performance of central agents (characterized by $p=\frac{1}{2}$), implying a flatter gene distribution for these boundary conditions. On the other hand, for periodic boundary conditions, one finds that the temporal oscillations are much more similar to the uniform case studied in Refs. [15,16] (as compared to the case of reflective boundary conditions). Indeed, the ratio $P(\frac{1}{2}):P(0)$ for periodic boundary conditions is very similar to the corresponding ratio in the uniform case (compare Fig. 1 with Fig. 1 of Ref. [15]).
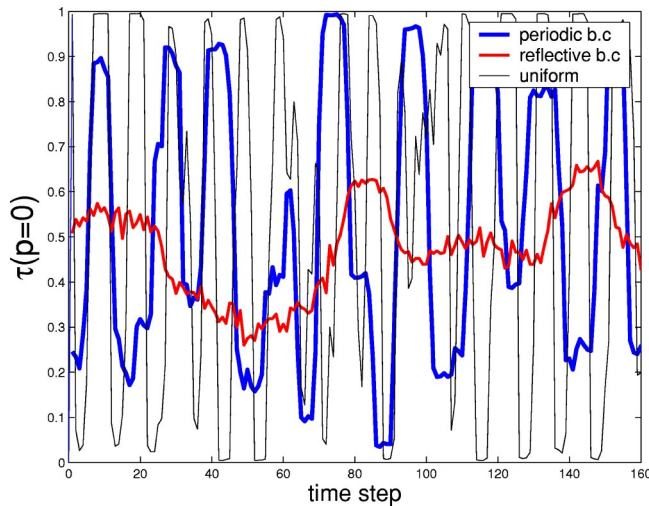

FIG. 2. Temporal dependence of the winning probabilities $\tau(p=0)$ for three distinct cases: (i) $\delta p = 0.1$ with periodic boundary conditions, (ii) $\delta p = 0.1$ with reflective boundary conditions, and (iii) uniform updating rule. The results are for $N = 10\,001$ agents, $R = 0.8$, and $d = -4$.
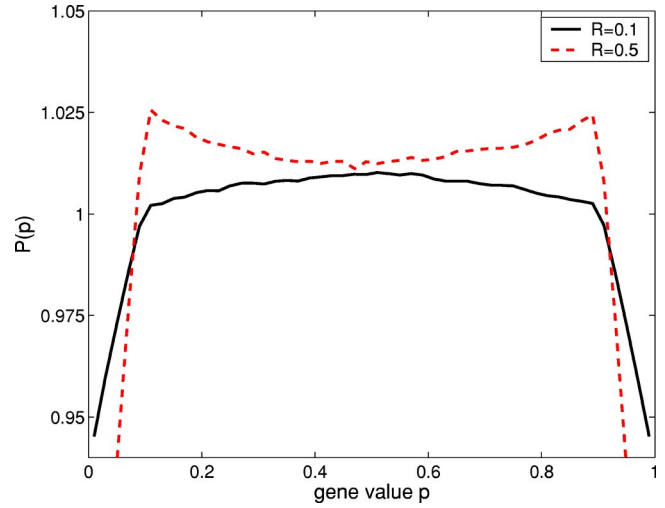

FIG. 3. The strategy distribution $P(p)$ for different values of the prize-to-fine ratio: $R = 0.1$ and $R = 0.5$. The results are for $N = 10\,001$ agents, $d = -4, \delta p = 0.1$, and periodic boundary conditions. Each point represents an average value over 10 runs and 100 000 time steps per run.

Figure 3 shows the strategy distribution of the population for different prize-to-fine ratios, and with $\delta p \ll 1$. The results demonstrate the existence of a stable phase characterized by an inverse-U shaped gene distribution. However, unlike the uniform case [15], the critical value of R which separates the inverse-U distribution from the M-shaped one does not equal 1 (in the $N \to \infty$ limit).

In Fig. 4, we display $P(p)$ for different δp values with periodic boundary conditions. We find that the peaks of the strategy distribution (for prize-to-fine ratios which are large enough to allow an M-shaped gene distribution) occurs at $p = \delta p$ and its symmetric counterpart $1 - \delta p$. Regardless of the value of δp, the agents do *not* self-segregate—the ex-
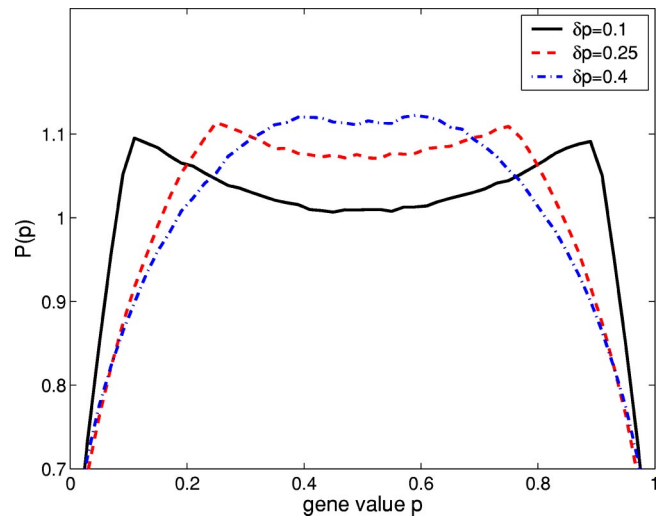

FIG. 4. The strategy distribution $P(p)$ for different δp values: $\delta p = 0.1$, 0.25, and 0.4. The results are for $N = 10\,001$ agents, $R = 0.9, d = -4$, and periodic boundary conditions. Each point represents an average value over 10 runs and 100 000 time steps per run.
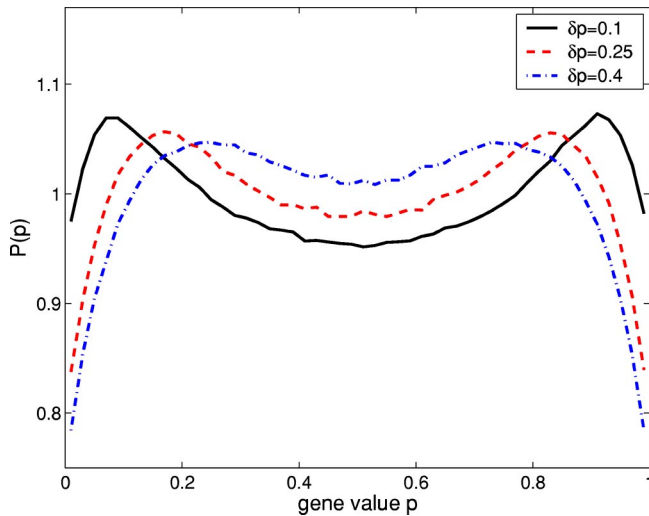
FIG. 5. The strategy distribution $P(p)$ for different δp values: $\delta p = 0.1$, 0.25, and 0.4. The results are for $N = 10\,001$ agents, $R = 0.9, d = -4$, and reflective boundary conditions. Each point represents an average value over 10 runs and 100 000 time steps per run.

FIG. 7. The efficiency of the system as a function of the length scale, δp. Horizontal line represents the efficiency for a coin-tossing situation. The results are for $N = 10\,001$ agents, $R = 0.7, d = -4$, and reflective boundary conditions.

treme strategies ($p = 0$ and $p = 1$) perform worst. The strategy distribution moves smoothly into an inverse-U shape in the limit of $\delta p = \frac{1}{2}$ [15]. Figure 5 displays the same results for reflective boundary conditions, where $\delta p = 1$ is equivalent to the uniform updating rule [15].

Figure 6 displays the average life span $\langle L(p) \rangle$ of the agents. In order to get a better picture of the life span distribution, we also plot $\langle L(p) \rangle + \sigma_L(p)$ as a function of the gene value p. Here, $\sigma_L(p)$ is the root mean square separation of the life spans. In this case, one finds an inverse-U shaped distribution (with the peak occurring at $p = \frac{1}{2}$). This implies that the agents characterized by $p = \frac{1}{2}$ (a coin-tossing strat-
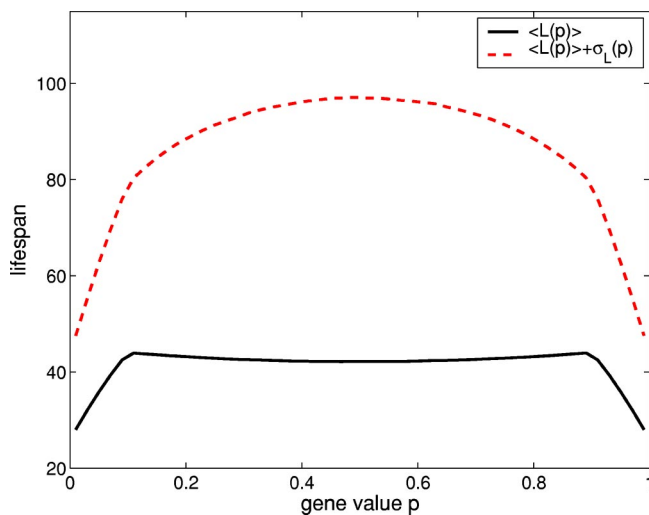

FIG. 6. The average life span $\langle L(p) \rangle$ (solid curve) and $\langle L(p) \rangle + \sigma_L(p)$ (dashed curve) of the agents. The results are for $N = 10\,001$ agents, $R = 0.8, d = -4, \delta p = 0.1$, and periodic boundary conditions. Each point represents an average value over 10 runs and 100 000 time steps per run.
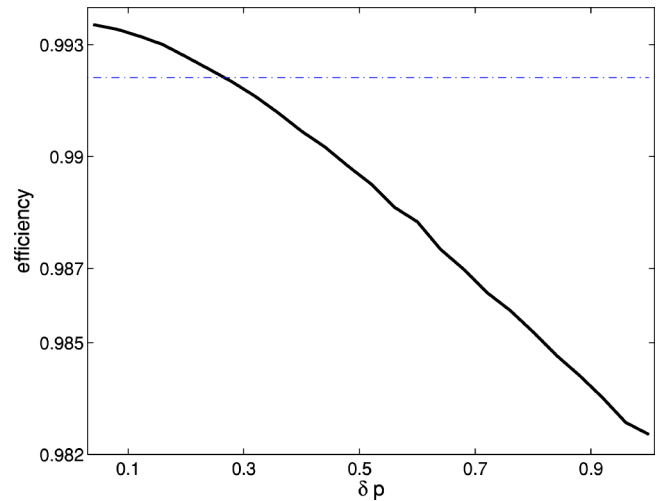
egy) have the best chances of survival in asymptotically long times, as predicted analytically in Ref. [17]. This important feature is explained by the global currents in the gene space, which *reduce* the value of $\sigma_L(p = 0)$, and have a negligible effect on $\sigma_L(p = \frac{1}{2})$ [16,17]. We emphasize that these results hold true for both periodic and reflective boundary conditions.

The efficiency of the system is defined as the number of agents in the minority room divided by the maximal possible size of the minority group, $(N - 1)/2$. Figure 7 displays the efficiency as a function of the length scale δp. The system's efficiency is a monotonic decreasing function of δp. This is caused by the fact that larger δp values imply *larger* temporal oscillations in the occupation numbers of the rooms, thus decreasing the number of agents in the winning group (and increasing the number of agents in the losing room).

We would like to stress that different complex systems display different updating rules and different boundary conditions. For instance, in a system in which an agent whose score falls below the threshold d quits the game (and replaced by a *new* agent), the relevant updating rule is the uniform one. In systems with periodic boundary conditions, the agents identify the $p = 0$ strategy with the $p = 1$ strategy—this reflects a psychological effect in which humans tends to replace one extreme strategy with the other extreme strategy [Extreme agents (with $p = 0$) may prefer another extreme strategy ($p = 1$) on taking the cautious strategy $p = \frac{1}{2}$.] In biological systems, in which this psychological effect is not strong, the relevant boundary conditions are the reflecting ones.

Finally, we would like to address the last point raised in Ref. [20]. It is claimed that the fluctuations in the average gene value $\langle p \rangle$ have been considered in Ref. [14]. However, the *oscillatory* behavior of $\langle p \rangle$, which is a highly important feature of the system's dynamics, was *not* observed in Ref. [14]. Rather, Burgos *et al.* [14] found a nonoscillatory value for $\langle p \rangle - \frac{1}{2}$, see Eq. (15) of Ref. [14]. We have shown, on the

other hand, that the quantity $\langle p \rangle - \frac{1}{2}$ displays temporal oscillations with a *well-defined* frequency and amplitude [15,16]. It is important to distinguish between the *regular* temporal oscillations of the physical quantities (such as $\langle p \rangle$) discussed in Refs. [15,16], as opposed to thermal fluctuations discussed in Ref. [14]. Thermal fluctuations of a thermodynamic system are essentially random in nature, whereas we have found regular oscillations that are characterized by a well-defined frequency and amplitude. The *oscillatory* nature of $\langle p \rangle$ [15,16] has been proven to be an essential feature which is responsible for the dynamical phase transition (from self-segregation to clustering) observed in the EMG [17]. We would like to emphasize that these oscillations exist also for the complex systems with a strategy updating rule of the form $p \rightarrow p \pm \delta p$, regardless of the value of δp and the type of boundary conditions used (see Fig. 2).

[1] N.F. Johnson, P.M. Hui, R. Jonson, and T.S. Lo, Phys. Rev. Lett. **82**, 3360 (1999).

[2] B. Huberman and R. Lukose, Science **277**, 535 (1997).

[3] D. Challet and C. Zhang, Physica A **246**, 407 (1997); **256**, 514 (1998); **269**, 30 (1999).

[4] R. D'Hulst and G.J. Rodgers, Physica A **270**, 514 (1999).

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

[6] H. Ceva, Physica A **277**, 496 (2000).

[7] E. Burgos and H. Ceva, Physica A **284**, 489 (2000).

[8] T.S. Lo, P.M. Hui, and N.F. Johnson, Phys. Rev. E **62**, 4393 (2000).

[9] P.M. Hui, T.S. Lo, and N.F. Johnson, e-print cond-mat/0003309.

[10] M. Hart, P. Jefferies, N.F. Johnson, and P.M. Hui, e-print cond-mat/0003486; Phys. Rev. E **63**, 017102 (2000).

[11] T.S. Lo, S.W. Lim, P.M. Hui, and N.F. Johnson, Physica A **287**, 313 (2000).

[12] Y. Li, A. VanDeemen, and R. Savit, e-print nlin.AO/002004.

[13] E. Burgos, H. Ceva, and R.P.J. Perazzo, Physica A **294**, 539 (2001); Phys. Rev. E **64**, 016130 (2001); e-print cond-mat/0212635.

[14] E. Burgos, H. Ceva, and R.P.J. Perazzo, Phys. Rev. E **65**, 036711 (2002).

[15] S. Hod and E. Nakar, Phys. Rev. Lett. **88**, 238702 (2002).

[16] E. Nakar and S. Hod, Phys. Rev. E **67**, 016109 (2003).

[17] S. Hod, e-print cond-mat/0212055.

[18] R. Metzler and C. Horn, e-print cond-mat/0212481.

[19] A. Soulier and T. Halpin Healy, Phys. Rev. Lett. (to be published), e-print cond-mat/0209451.

[20] E. Burgos, H. Ceva, and R.P.J. Perazzo, e-print cond-mat/0301518.

[21] See http://www.unifr.ch/econophysics/minority