Eur. Phys. J. B **60**, 369–384 (2007)
DOI: 10.1140/epjb/e2007-00353-6

**THE EUROPEAN PHYSICAL JOURNAL B**

# "Illusion of control" in Time-Horizon Minority and Parrondo Games

J.B. Satinover[1] and D. Sornette[2,a]

[1] Laboratoire de Physique de la Matière Condensée, CNRS UMR6622 and Université des Sciences, Parc Valrose, 06108 Nice Cedex 2, France
[2] Department of Management, Technology and Economics, ETH Zurich, 8032 Zurich, Switzerland

Received 4 April 2007 / Received in final form 30 July 2007
Published online 22 December 2007 – 

**Abstract.** Human beings like to believe they are in control of their destiny. This ubiquitous trait seems to increase motivation and persistence, and is probably evolutionarily adaptive [J.D. Taylor, S.E. Brown, Psych. Bull. **103**, 193 (1988); A. Bandura, *Self-efficacy: the exercise of control* (WH Freeman, New York, 1997)]. But how good really is our ability to control? How successful is our track record in these areas? There is little understanding of when and under what circumstances we may over-estimate [E. Langer, J. Pers. Soc. Psych. **7**, 185 (1975)] or even lose our ability to control and optimize outcomes, especially when they are the result of aggregations of individual optimization processes. Here, we demonstrate analytically using the theory of Markov Chains and by numerical simulations in two classes of games, the Time-Horizon Minority Game [M.L. Hart, P. Jefferies, N.F. Johnson, Phys. A **311**, 275 (2002)] and the Parrondo Game [J.M.R. Parrondo, G.P. Harmer, D. Abbott, Phys. Rev. Lett. **85**, 5226 (2000); J.M.R. Parrondo, *How to cheat a bad mathematician* (ISI, Italy, 1996)], that agents who optimize their strategy based on past information may actually perform worse than non-optimizing agents. In other words, low-entropy (more informative) strategies under-perform high-entropy (or random) strategies. This provides a precise definition of the "illusion of control" in certain set-ups a priori defined to emphasize the importance of optimization.

**PACS.** 89.75.-k Complex systems – 89.65.Gh Economics; econophysics, financial markets, business and management – 02.50.Le Decision theory and game theory

## 1 Introduction

The success of science and technology, with the development of ever more sophisticated computerized integrated sensors in the biological, environmental and social sciences, all illustrate the quest for control as a universal endeavor. The exercise of governmental authority, the managing of the economy, the regulation of financial markets, the management of corporations, and the attempt to master natural resources, control natural forces and influence environmental factors all arise from this quest. Langer's phrase, "illusion of control" [3] describes the fact that individuals appear hard-wired to over-attribute success to skill, and to underestimate the role of chance, when both are in fact present. Whether control is genuine or merely perceived is a prevalent question in psychological theories. The following presents two rigorously controlled mathematical set-ups demonstrating generic circumstances in which optimizing agents perform worse than their non-optimized strategies, or than non-optimizing or random agents.

[a] e-mail: dsornette@ethz.ch

## 2 Minority Games

### 2.1 Definition and summary of main results for the Time-Horizon MG (THMG)

We first study a variant of Minority games (MGs), which constitute a sub-class of market-entry games. MGs exemplify situations in which the "rational expectations" mechanism of standard economic theory fails. This mechanism in effect asks, "what expectation model would lead to collective actions that would on average validate the model, assuming everyone adopted it?" [10]. In minority games, a large number of interacting decision-making agents, each aiming for personal gain in an artificial universe with scarce resources, try to anticipate the actions of others on the basis of incomplete information. Those who subsequently find themselves in the minority group gain. Therefore, expectations that are held in common negate themselves, leading to anti-persistent behavior both for the aggregate behavior and for individuals. Minority games have been much studied as repeated games with expectation indeterminacy, multiple equilibria and inductive optimization behavior.

Consider the Time-Horizon MG (THMG), where $N$ players have to choose one out of two alternatives at each time step based on information represented as a binary time series $A(t)$. Those who happen to be in the minority win. Each agent is endowed with $S$ strategies. Each strategy gives a prediction for the next outcome $A(t)$ based on the history of the last $m$ realizations $A(t-1)$, ..., $A(t-m)$ ($m$ is called the memory size of the agents). Each agent holds the same number $S$ of (in general different) strategies among the $2^{2^m}$ total number of strategies. The $S$ strategies of each agent are chosen at random once and for all at the beginning of the game. At each time $t$, in the absence of better information, in order to decide between the two alternatives for $A(t)$, each agent uses her most successful strategy in terms of payoff accumulated in a rolling window of finite length $\tau$ up to the last information available at the present time $t$ (the case of a limitlessly growing $\tau$ corresponds to the standard MG; the term "Time Horizon MG" refers to the case of a fixed and finite $\tau$). This is the key optimization step. If her best strategy predicts $A(t) = +1$ (resp. $-1$), she will take the action $a_i(t) = -1$ (resp. $+1$). The aggregate behavior $A(t) = \Sigma_{i=1}^{N} a_i(t)$ is then added to the information set available for the next iteration at time $t+1$. The corresponding instantaneous payoff of agent $i$ is given by $-\text{sign}[a_i(t)A(t)]$ (and similarly for each strategy for which it is added to the $\tau - 1$ previous payoffs). As the name of the game indicates, if a strategy is in the minority ($a_i(t)A(t) < 0$), it is rewarded. In other words, agents in THMG try to be anti-imitative. The richness and complexity of minority games stem from the fact that agents strive to be different. Previous investigations have shown the existence of a phase transition marked by agent cooperation and efficiency between an inefficient regime (worse than random) and a random-like regime as the control parameter $\alpha \equiv 2^m/N$ is increased: in the vicinity of the phase transition at $\alpha_c = 2^{m_c}/N \approx 0.34$ (for both the THMG and MG proper), the size of the fluctuations of $A(t)$ (as measured by its normalized variance $\sigma^2/N$) falls below the random coin-toss limit for large $m$'s (assuming fixed $N$) when agents always use their highest scoring strategy [6]. In other words, for a range of $m$ (given $N$, $S$), agent performance is better than what strategy performance would be in a game with no agents optimizing. The phenomenon discussed here is that when optimizing, and averaged over all actual agents and strategies in a given realization, agents in the TH variant of the MG nonetheless generally underperform the mean of their own measured strategy performance and do so in all phases for reasonable lengths of $\tau$ (as also the mean over all strategies in a given realization. For any given realization, however, a minority of agents outperform their strategies and the majority of other agents. Some may also achieve net positive gain, if rarely.). In the MG proper, however, $\tau$ is unbounded and a stationary state is reached at some very large $\tau_{eq} \geq 2^m \times 200$ where a subset of agents "freeze" their choice of strategy: one virtual strategy score attains a permanently higher value than any other. These frozen agents in general do outperform the mean of all strategies in a given realization as well as the mean of their own $S$ original strategies: they perform precisely as well as their best. We focus on results in the THMG with an eye towards real-world markets in which because the time series being predicted are non-stationary, trading strategies are weakened if they incorporate an unbounded (and uniformly-weighted) history of prior strategic success or failure: remote history is less important than recent history and beyond a certain point is meaningless. Unless specifically stated otherwise, throughout this paper, whenever we compare agent to strategy performance, we always mean the performance of agents' strategies as measured by the accumulation of hypothetical points averaged over *all* agents in the system and the set of *all* of their strategies. Furthermore, in selecting a strategy the agents do not take account of the impact of their choice on the probable minority state — that is, they do not consider that their own selection of action reduces the probability that this action will be the minority one (we refer to such agents as "standard").

### 2.2 Statement of our main results on the "illusion of control" in the THMG

Our main result may be stated concisely from the perspective of utility theory: throughout the space of parameters ($N$, $m$, $S$, $\tau \ll \tau_{eq}$), the mean payoff of agents' strategies (as calculated by each agent averaged over all strategies and agents in a realization) not only surpasses the mean payoff of supposedly-optimizing agents (averaged over all given agents), but the respective cumulative distribution functions (CDF) of payoffs show a first-order stochastic dominance of strategies over agents. Thus, were the option available to them, agents would behave in a risk-averse fashion (concave utility function) by switching randomly between strategies rather than optimizing. This result generalizes when comparing optimizing agents with $S > 1$ strategies with agents having only one strategy (or equivalently $S$ identical strategies), when the single strategies are actually implemented. (This takes into account any difference in strategy performance that may arise from the simple fact of a strategy actually being deployed). The same result is also found when comparing optimizing agents with agents flipping randomly among their $S$ strategies. Agents are supposed to enhance their performance by choosing adaptively between their available strategies. In fact, the opposite is true: by our metric, the optimization method would appear to agents as strictly a method for worsening performance! (In the MG proper the situation is more complex. As detailed in [6], agents with two identical strategies — equivalent to having only one [and called therein "producers"] — always have net gain $\leq 0$. But this gain may be on average either greater or less than for agents that optimize among more than one strategy [called "speculators"]. Which is true depends inter alia on the proportion of producers to speculators: a very small proportion of producers will outperform speculators. There is an expected proportion of producers that arises from the average over many different random possible initial allocations of strategies among agents (i.e., quenched disorder). Given this expected mean proportion, and averaging over *all* the agents in each initial allocation, the mean performance of strategies is better than that of agents. This is true, however, only when agents do not choose their strategies at each step taking into account the impact of that selection.)

Let us restate our result for the THMG in the language of a financial market with traders trying to outperform the overall market. We argue that in using the THMG as a model for traders' actions, the following is the case: every trader attempting to optimize by selecting his "best performing strategy" measures that performance virtually, not by contrast to an imagined setting where all traders select fixed strategies at random (to whose results he would have no access anyway). Even though the virtual performances of each of his basket of strategies might never have been implemented in reality, if he found that his real performance under a selection process was worse than the virtual performance of the strategies he had been selecting among, he would abandon the selection process. This would be true for most agents and not true only for a small minority (If *every* trader were to do the same, of course, then one would end up with the random or fixed choice game as discussed below. This forms the usual standard of comparison for strategy performance in the MG literature). This resonates with the finding of Doran and Wright [7], who report that two-thirds of all finance professors at accredited, four-year universities and colleges in the US (arguably among the most sophisticated and informed financial investors) are passive investors who think that the traditional valuation techniques are all *un*important in the decision of whether to buy or sell a specific stock (in particular, the CAPM, APT and Fama and French and Carhart models).

### 2.3 Quantitative statement and tests

In the THMG, the "illusion of control" effect is observed for all $N$, $m$, $S$ and $\tau \ll \tau_{eq}$. We use the Markov chain formalism for the THMG [9,11] to obtain the following theoretical prediction for the gains, $\Delta W_{Agent}$ averaged over all agents and $\Delta W_{Strategy}$ averaged over all strategies respectively, of agents and of all strategies in a given realization [12]:

$$\langle \Delta W_{Agent} \rangle = \frac{1}{N} \left| \vec{A}_D \right| \cdot \vec{\mu} \qquad (2.1)$$

$$\langle \Delta W_{Strategy} \rangle = \frac{1}{2N} \left( \hat{\mathbf{s}}_\mu \cdot \vec{\kappa} \right) \cdot \vec{\mu} \qquad (2.2)$$

Brackets denote a time average. $\mu$ is a $(m+\tau)$-bit "path history" [9] (sequence of 1-bit states); $\vec{\mu}$ is the normalized steady-state probability vector for the history-dependent $(m+\tau)(m+\tau)$ transition matrix $\hat{\mathbf{T}}$, where a given element $T_{\mu_t,\mu_{t-1}}$ represents the transition probability that $\mu_{t-1}$ will be followed by $\mu_t$; $\vec{A}_D$ is a $2^{(m+\tau)}$-element vector listing the particular sum of decided values of $A(t)$ associated with each path-history; $\hat{\mathbf{s}}_\mu$ is the table of points accumulated by each strategy for each path-history; $\vec{\kappa}$ is a $2^{(m+\tau)}$-element vector listing the total number of times each strategy is represented in the collection of $N$ agents. As shown in the supplementary material, $\hat{\mathbf{T}}$ may be derived from $\vec{A}_D$, $\hat{\mathbf{s}}_\mu$ and $\vec{N}_U$, the number of undecided agents associated with each path history. Thus agents' mean gain is determined by the non-stochastic contribution to $A(t)$ weighted by the probability of the possible path histories. This is because the stochastic contribution for each path history is binomially distributed about the determined contribution. Strategies' mean gain is determined by the change in points associated with each strategy over each path-history weighted by the probability of that path.

We find excellent agreement between the numerical simulations and the analytical predictions (2.1) and (2.2) for the THMG. For instance, for $m = 2$, $S = 2$, $\tau = 1$ and $N = 31$, $\langle \Delta W_{Agent} \rangle = -0.22$ for both analytic and numerical methods (payoff per time step averaged over time and over all optimizing agents) compared with $\langle \Delta W_{Strategy} \rangle = -0.06$ also (similarly averaged over all strategies) for both analytic and numerical methods. In this numerical example, the average payoff of individual strategies is larger than for optimizing agents by 0.16 units per time step. The numerical values of the predictions (2.1) and (2.2) are obtained by implementing each agent individually as a coded object.

In the THMG, the mean per-agent per-step payoff $\langle \Delta W_{Non\text{-}Opt} \rangle$ accrued by non-optimizing agents (they have only one fixed strategy, or equivalently their $S$ strategies are identical; a.k.a. "producers") is larger than the payoff $\langle \Delta W_{Agent} \rangle$ of optimizing agents (a.k.a. "speculators"). In general, this comparative advantage decreases with their proportion but much less rapidly than in the MG proper [6]. For example, with $m = 2$, $S = 2$, $\tau = 1$ and $N = 31$, and 2500 random initializations and $n$ optimizing agents, $\langle \Delta W_{Non\text{-}Opt} \rangle - \langle \Delta W_{Agent} \rangle = (2.380, 2.270, 2.289, 2.275, 2.145, 2.060, 2.039, 1.994, 1.836, 1.964) \times 10^{-3}$ for $n = 1, 2, \ldots, 10$. More generally, the following ordering holds: payoff of individual strategies > payoff of non-optimizing agents > payoff of optimizing agents. The first inequality is due to the fact that not all individual strategies are implemented and the theoretical payoff of the non-implemented strategies does not take into account what their effect would have been (had they been implemented). Implementation of a strategy tends to decrease its performance (this is similar to the market impact of trading strategies in financial markets associated with slippage and market friction). Non-optimizing agents by definition always implement their strategies. However, the higher payoff of non-optimizing compared with optimizing agents shows that the illusion-of-control effect is not due to their actually being deployed, but is a genuine observable effect.

### 2.4 Generalizations

The amplitude of the illusion-of-control effect in the THMG highlights important differences between the MG proper, in which $\tau$ is sufficiently large so as to allow the system to attain equilibrium with many "frozen" agents ($\sim 10^4 - 10^6$ time steps) and the THMG in which $\tau$ is arguably of a length comparable to real-world investment "lookbacks". The effect also highlights the distinction between optimizing agents with $S$ maximally distinct strategies (in the sense of Hamming distance) and non-optimizing agents with $S$ identical strategies — a distinction with different characteristics in the THMG than in the MG proper.

It is helpful to generalize the latter distinction by characterizing the degree of similarity between the $S$ strategies of a given agent using the Hamming distance $d_H$ between strategies (the Hamming distance between two binary strings of equal length is the number of positions for which the corresponding symbols are different, normalized on the unit interval). Non-optimizing agents with $S$ identical strategies correspond to $d_H = 0$. In contrast, optimizing agents with $S$ maximally distinct strategies have large $d_H$'s. Since agents in the THMG with $d_H = 0$ outperform agents with large $d_H$, it is natural to ask whether the ranking of $d_H$ could be predictive of the ordering of agents' payoffs. But first it is important to clarify differences with respect to $d_H$ in the MG versus the THMG.

The first difference to emphasize is that in the MG, where the system runs to equilibrium, one of the chief features of the stationary state attained is that some (and sometimes many) of the optimizing agents (with their strategy $d_H > 0$) "freeze". That is, the effective $\tau$ is long enough so that one of the $S$ strategies for some agents will attain permanently the largest number of "virtual" points. It will then always be deployed, somewhat similar it might seem, to an agent with $S$ identical strategies from the start.

The second difference is that at equilibrium in the MG, the relation between agent performance and $d_H$ inverts at the critical point $\alpha_c$ [14]: on average, for $\alpha > \alpha_c$, agents with larger $d_H$ outperform those with smaller $d_H$ — and outperform the mean over the selected strategies. This reversal is due to the freezing of a subset of agents. Over the very long run-up to equilibrium, frozen agents have the opportunity to choose what is in fact a better strategy. It is unsurprising that a larger Hamming distance between strategies offers more opportunity for such a differentiation to occur. Conversely, for agents with $d_H = 0$, such selection is impossible.

Note that for extremely short $\tau$ (e.g., 1, 10), the phase-transition does not occur: rather, mean agent performance increases monotonically and approaches asymptotically that of mean strategy performance. As $\tau$ increases a number of things happen. First, the phase transition at $\alpha_c$ appears and grows increasingly sharp. Second, the overall scale of agent return (comparably, volatility of $A(t)$, i.e., $-\sigma^2/N$) as a function of $m$ varies periodically with a period equal to $2 \times 2^m$ for real histories (but does not vary for random ones) [12]. Third, so long as $\tau$ remains "reasonable", a reversal of the relation between $d_H$ and agent performance does not occur (i.e., the larger the $d_H$, the smaller the agent gain in wealth, for all $\alpha$). "Reasonable" lengths for $\tau$ in the THMG, from the perspective of relative stationarity in real-world financial time-series, cannot be denoted without taking into account the regime: for a modest number of agents (e.g., 31) at small $m$ (e.g., for $N = 31$, $m < 4$), $200 \times 2^m$ time-steps is sufficient to reach equilibrium. But near the phase transition, many more steps are required, on the order of $5000 \times 2^m \geq 80\,000$, equivalent to 320 years of daily price data, assuming that a time step equals one trading day. For $\tau$ on the order of 1000, no reversal of the relation between small $d_H$ and better agent performance occurs. Figure 1 provides a typical example of the distribution of agent returns by $d_H$. A similar distribution with negative linear slope occurs for all reasonable values of $m$ and $\tau$ short of $\tau_{eq}$.

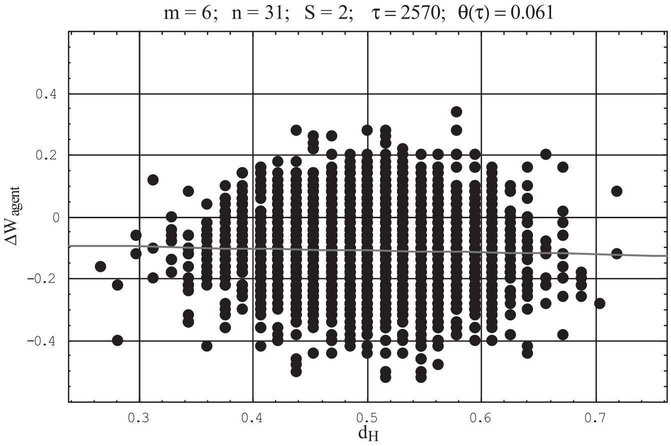


**Fig. 1.** Typical distribution of agent returns by Hamming distance between component strategies. A simple linear fit with slope $\theta(t) = -0.061$ demonstrates decreasing mean agent performance with increasing $d_H$; the time-horizon $\tau = 2570$ is even longer than "reasonable"; it is more than long enough for a sharp phase transition to be present at $3 \leq m \leq 4$ (but still well short of equilibrium, i.e., $\tau < \tau_{eq} \simeq 12\,800$); the memory $m = 6$ is well past the phase transition after which in the MG proper the relation between $d_H$ and $\Delta W_{agent}$ inverts.

A non-zero $d_H$ implies that there are at least two strategies among the $S$ strategies of the agent which are different. But, if $d_H$ is small, the small difference between the $S$ strategies makes the optimization only faintly relevant and one can expect to observe a payoff similar to that of non-optimizing agents, therefore larger than for optimizing agents with large $d_H$'s. This intuition is indeed confirmed by our calculations: the payoff per time step averaged over all agents is a decreasing function of $d_H$, as originally discussed in [13,14] for the MG at equilibrium (and for $\alpha < \alpha_c$).

The illusion-of-control effect suggests that the initial set-up of the THMG in terms of $S$ fixed strategies per agent is evolutionarily unstable (when agents do not select strategies taking account of their impact). It is thus important to ask what happens when agents are allowed to replace strategies over time based on performance. A number of authors have investigated this issue in the MG, adding a variety of longer-term learning mechanisms on top of the short-term adaptation that constitutes the basic MG [14–20]. Inter alia, reference [16] demonstrates that if agents are allowed to replace strategies over time based on performance, they do so by ridding themselves of those composed of the more widely Hamming-distant tuples. Agents that start out composed of identical strategies do not change at all; those composed of strategies close in Hamming space change little. Similarly, the authors of [15] explicitly fixed agents with tuples of identical strategies and found they performed best. Another important finding in [15] is that the best performance attainable is equivalent to that obtained by agents choosing their strategies at random. Note that learning only confers a *relative* advantage. In general, agents that learn out-perform agents that don't. This is certainly true for this privileged subset of agents among standard ones. But the performance of learning agents approaches a maximum most closely attained by agents where the Hamming distance between strategies is 0. These agents neither adapt (optimize) nor learn. One might say that when learning is introduced, the system learns to rid itself of the illusory optimization method that has been hampering it. (Note that if one compares optimizing agents' performance to the performance of a separate system composed entirely of non-optimizing agents, there are regimes in $m$ for which the optimizing agents do better: the standard metric of comparison in the MG literature. This could arise "in reality" only if traders deliberately ignored the evidence most perceive, namely, that the mean of their own strategies appear to be outperforming the optimization process that chooses deliberately among them. We emphasize "most" here, because a smaller proportion of traders' would in fact see their optimization process succeeding. Again, mean agent performance underperforms mean strategy performance when averaged over all agents and all of the strategies represented in a given quenched disorder.)

There are exceptions, of course. Carefully designed privileges and certain kinds of learning can yield superior results for a subset of agents, and occasionally for all agents. But the routine outcome is that both optimization and straightforward learning cannot improve on simple chance, as measured by agents' own assessment of their strategies' respective virtual performance. The fact that the optimization method employed in the THMG yields the opposite of the intended consequence, and that learning eliminates the method, leads to an important question. We pose it carefully so as to avoid introducing either privileged agents or learning: is the illusion-of-control so powerful in this instance that inverting the optimization rule could yield equally unanticipated and opposite results? The answer is yes: if the fundamental optimization rule of the MG is symmetrically inverted for a limited subset of agents who choose their worst-performing strategy instead of their best, those agents systematically outperform both their strategies and other agents. They also can attain positive gain. Thus, the intuitively self-evident control over outcome proffered by the THMG "optimization" strategy is most strikingly shown to be an illusion. Even learning and evolutionary strategies generally at best rid the system of any optimization method altogether. They do not attain the kind of results obtained simply by allowing some agents to reverse the method altogether. We discuss elsewhere the phenomena that arise as the proportion of agents choosing their best performing strategy and of agents choosing their worst performing strategy are varied for different parameters of the THMG and MG proper. We emphasize the fact that extensive numerical studies confirm that the phenomenon here indicated persist over a very wide range of parameters in the MG and over all parameter values in the THMG. Hence, having a portfolio of $S$ strategies to choose from is in the THMG always counter-productive, and in the MG often so: diversification + optimization performs on average worse than a single fixed strategy.

Let us also mention briefly a related work by Menche and de Almeida. In the standard MG, the only public information are recent first places, while Menche and de Almeida [22] introduce a history of second places in the agents' set of strategies, thus providing more information to the agents about the state of the game and about the quality of their strategies. They find that the resulting performance of the system becomes significantly better and the phase transition into the uncorrelated phase is strongly suppressed. Note that this variation grants agents greater computational capacity than agents in the standard MG. For $S = 2$ it comes close to the simple variation we explore, where some agents choose their worst strategy instead of their best without changing the computational complexity of the game nor of individual agents.

### 2.5 Illusion of control and the crowding-out mechanism

Intuitively, the illusion-of-control effect in MG results from the fact that a strategy that has performed well in the past becomes crowded out in the future due to the minority mechanism: performing well in the recent past, there is a larger probability for a strategy to be chosen by an increasing number of agents, which inevitably leads to its demise. This argument in fact also applies to all the strategies that belong to the same reduced set; their number is $2^{2^m}/2^m$, equal to the ratio of the cardinality of the set of all strategies to the cardinality of the set of reduced strategies. Thus, the crowding mechanism operates from the fact that a significant number of agents have at least one strategy in the same reduced subset among the $2^m$ reduced strategy subsets. Optimizing agents tend on average to adapt to the past but not the present. They choose an action $a(t)$ which is on average out-of-phase with the collective action $A(t)$. In contrast, non-optimizing agents average over all the regimes for which their strategy may be good and bad, and do not face the crowding-out effect. The crowding-out effect also explains simply why anti-optimizing agents over-perform: choosing their worst strategy ensures that it will be the least used by other agents in the next time step, which implies that they will be in the minority. The crowding mechanism also predicts that the smaller the parameter $2^m/N$, the larger the illusion-of-control effect.

Indeed, as one considers larger and larger values of $2^m/N$, it becomes more and more probable that agents have their strategies in different reduced strategy classes, so that a strategy which is best for an agent tells nothing about the strategies used by the other agents, and the crowding out mechanism does not operate. Thus, regions of successful optimization, if they occur at all, are more likely at higher values of $2^m/N$ (see Appendix A for further details).

### 2.6 Robustness of the "illusion of control" phenomenon: THMG versus MG

It could be argued that the phenomenon of "illusion of control" that we report is very specific because we consider the THMG, and not the standard MG. In the standard MG, some agents are frozen as a result of lengthy optimization, so that some of these agents are able to win more than half of the time. This appears similar to the references we discuss wherein learning takes place and agents learn to "rid" themselves of strategy choice. But in the MG proper the "ridding" takes place at the more fundamental level of the basic optimization procedure and reflects a genuine (non-illusory) control that appears along with the phase transition (requiring an especially lengthy run-up to equilibrium). That some agents are able to win more than half the time when they are frozen is only in part analogous to when there exists a subset of select agents — for example ones that take into account their impact; or agents with two identical strategies (equivalent therefore to being frozen from the start, albeit without any preceding selection process); or agents that choose their worst strategy. From one perspective, the existence of genuine control in certain phases of the MG proper is an artefact of an "unreasonably" long equilibration process (and an equilibrium state arguably not found in real-world markets). From the opposite perspective, the illusion of control in the THMG is an artefact of "transients" relatively early in the system's equilibration process.

In any event, we observe that in the THMG, agents with two identical strategies on average outperform those selecting among strategies, but do not do better than a 0.5 win rate — again, averaged over all those that do in fact do better than 0.5 and those that don't. We also observe that agents that always choose their worst strategy (when they are a subset among a majority that choose their best as usual) have a better than 0.5 win rate on average for a number of parameter values. As detailed in [6,22], when all agents take into account their impact, the agents do now outperform their strategies. However, the game settles into a Nash equilibrium [24] which is arguably an entirely different situation, one in which the "illusion" of control is no longer pertinent as the dynamics are in this case deterministic. (A more realistic situation occurs when, for example, only some traders account correctly for impact, or when some or all account for impact only imperfectly. Depending on the extent of impact-accounting, the system may remain frustrated but the illusion of control may still disappear.)

Any agent accounting for impact looks back at the prior vote imbalance and determines what the imbalance would have been had she used each given strategy (not what its score would have been using just the strategy she actually did use). Similar methods are used by real-world traders taking positions large enough to have an impact on price. A consequence of taking into account impact is that, with a certain probability usually smaller than one, at any time-step, an agent will select some strategy other than the "best" (i.e., as computed in the standard way). From this perspective, accounting for impact is similar (but not identical) to the computationally simpler act of standard agents in fact choosing other than the best strategy, always. For $S = 2$, this is the same as choosing the worst. Although we have reported on the effect of choosing the worst strategy for $S = 2$, the same principle holds for $S > 2$: a subset of agents choosing their worst strategy outperforms, on average, those that choose their next-to-worst, etc. When a small subset of agents take into account their impact, on average these perform better than those that do not. However, they do not perform as well as a similarly sized subset of agents choosing their worst strategy. The dynamics of a game composed entirely of agents choosing their worst strategy is not similar to the Nash-equilibrated structure of a game with all agents accounting for impact.

It is true in particular that there is a complex relation between maximum/minimum typical system fluctuation/cooperation and length of strategy score memory ($\tau$), which we do not discuss in detail: $\sigma^2/N$ is periodic in $2 \times 2^m$ [12]. Nonetheless, as many simulations that we have performed illustrate for both real and random histories (they are available from the authors upon request), the underperformance of standard agents vis-à-vis the mean of all strategies represented in a given $\hat{\mathbf{\Omega}}$, and over many $\hat{\mathbf{\Omega}}$, is found for all $\tau$ up to (and greater than) the equilibrium number of steps at which point the THMG becomes equivalent to the MG. The difference between strategy performance (as we define it) and agent performance declines roughly exponentially with $\tau$ but remains positive. When a critical point is present (sufficiently long $\tau$ but still well short of $\tau_c$) it reaches its positive minimum at the critical point. We find that strategy out-performance is greater for random histories at this point than for real histories. The phase transition central to the MG is most evident for $\tau$ long and is attenuated for $\tau$ short.

Even restricted to the THMG, the phenomena we are most interested in are essentially as prominent for $\tau \approx 1000$, say, as for $\tau \approx 1$. The rule of thumb for reaching the stationary state in the standard MG is to iterate for about $200 \times 2^m$ time steps (it takes even more time close to the critical point $\alpha_c$). Thus, for our simulations with $m = 2$, values of $\tau \approx 1000$ and above probe the stationary regime of the standard MG and confirm the robustness of the illusion-of-control effect. It is reasonable to argue that for real-life trading situations, which are generally non-stationary, a 100 or 1000 time-unit "look back" is of significant interest (Real world $\tau$'s $\leq 100$ are not unusual). "Look backs" long enough to achieve equilibrium, even for only tens of agents in the MG are (at least arguably) less likely to happen in reality. (On the other hand, studies that employ tick data may arguably require look-backs on the order of a MG $\tau_c$ and may be treated as at equilibrium.) If the subset of such agents taking into account their impact is large enough to be meaningful ($\sim$1/3 or more), one can see that the performance of these agents distributes itself as symmetrically as possible around zero. The remaining agents "perturb" this equilibrium. When the proportion of agents accounting for impact is large enough (depending on the other system parameters), the system as a whole settles into a deterministic equilibrium and there is no longer a phase transition at critical $\alpha_c$. This equilibrium is achieved more readily when $\tau$ is large yet need not be too large.

However, in a THMG composed entirely of impact-accounting agents, with $N = 31$, $S = 2$, a near equilibrium state is attained for $10 > \tau > 100$. Also, for $\tau = 1$ or 10, strategies outperform their agents as we have described. For $\tau \geq 100$, the reverse is true. In the MG (with standard agents that do not account for their impact), whenever the fluctuations of the global choice are better than with random ones, the agents perform globally better than in a game composed entirely of non-optimising standard ones. As discussed in [6], which agents perform best during which phase is highly sensitive to the precise ratio of "producers" (agents with $S$ identical strategies, hence non-optimizing) to "speculators" (agents with at least two different strategies), and to the degree to which agents have correlated actions as averaged over all histories. Frozen speculators in general perform best of all. We stress that this is not inconsistent with our observation that in the aggregate — not examining these "microscopic" differences among types and proportions of agents — standard agents at all $\tau$ nonetheless under-perform the mean of all strategies in a given quenched disorder averaged over many different such configurations.

### 2.7 First-entry games and symmetric evolutionary stable equilibria

The above discussion leads to the conclusion that there is often a profound clash between optimization on the one hand and minority payoff on the other hand: an agent who optimizes identifies her best strategy, but in so doing by her "introspection", she somehow knows the fate of the other agents, that it is probable that the other agents are also going to choose similar strategies, ... which leads to their underperformance since most of them will then be in the majority. It follows then that an optimizing agent playing a standard minority game should optimize at a second order of recursion in order to win: her best strategy allows her to identify the class of best strategies of others, which she thus must avoid absolutely to be in the minority and to win (given that other players are just optimizing at the first order as in the standard MG). Generalization to ever more complex optimizing set-ups, in which agents are aware of prior-level effects up to some finite recursive level, can in principle be iterated ad infinitum.

Actually, the game theory literature on first-entry games shows that the resulting equilibria depend on how agents learn [25]: with reinforcement learning, pure equilibria involve considerable coordination on asymmetric outcomes where some agents enter and some stay out; learning with stochastic fictitious plays leads to symmetric equilibria in which agents randomize over the entry decisions. There may even exist asymmetric mixed equilibria, where some agents adopt pure strategies while others play mixed strategies. We consider the situation where agents use a boundless recursion scheme to learn and optimize their strategy so that the equilibrium corresponds to the fully symmetric mixed strategies where agents randomize their choice at each time step with unbiased coin tosses. Consider a THMG game with N agents total, $N_R$ of which employ such a fully random symmetric choice. The remaining $N_S = N - N_R$ "special" agents (with $N_R \gg N_S$) will all be one of three possible types: agents with $S$ fixed strategies that choose their best (respectively worst, referred to above as anti-optimizing) performing strategy to make the decision at the next step and agents with a single fixed strategy. Our simulations confirm that these three types of agents indeed under-perform on average the optimal fully symmetric purely random mixed strategies of the $N_R$ agents (see Fig. A.5 of the Appendix). Here, pure random strategies are obtained as optimal, given the fully rational fully informed nature of the competing agents. The particular results are sensitive to which strategies are available to the special agents and to their proportion. Their underperformance in general requires averaging over all possible strategies and $S$-tuples of strategies. (In the Appendix, we show sample numerical results for $N_S = 1$.)

## 3 Parrondo games

We now turn to an entirely different kind of game of the Parrondo type. In the Parrondo effect (PE) [8], individually fair or losing games are combined either periodically or randomly to yield a winning game. That random alternation wins seems especially counterintuitive. The PE was originally conceptualized as the game-theoretic equivalent of the "flashing ratchet" effect: a charged particle that executes symmetric Brownian motion in a ratchet-shaped potential drifts unidirectionally if the potential is flashed on and off either at random or periodically [26–28]. It has been proposed as a potential explanation for aspects of random-walk diffusion [29], diffusion-mediated transport [30], spin systems [31], enzyme synthesis and gene recombination [32] and to be applied in investment strategies and portfolio optimization [33–35].

### 3.1 Single-player capital-dependent Parrondo effect

The original Parrondo Effect (PE) combines two "capital-dependent" games. A single player has (discrete)-time-dependent capital $X(t)$, $t = 0, 1, 2, \ldots$ The time evolution of $X(t)$ is determined by tossing biased coins. If game $A$ is played, the player's capital changes by $+1$ ("win") with probability $p$ and by $-1$ ("loss") with probability $1-p$. If game $B$ is played, the changes are determined by:

| | Prob. of win | Prob. of loss |
|---|---|---|
| $X(t)/3 \in \mathbb{Z}$ | $p_1$ | $1-p_1$ |
| $X(t)/3 \notin \mathbb{Z}$ | $p_2$ | $1-p_2$ |

For $p = 1/2-\varepsilon$, $p_1 = 1/10-\varepsilon$ and $p_2 = 3/4-\varepsilon$ ($\varepsilon > 0$). If either game $A$ or game $B$ is played exclusively, they both lose. In other words, $\langle X(t)\rangle$ decreases with $t$. But if the games are alternated at random $\langle X(t)\rangle$ increases. This is because the capital-dependent parameter $Mod[X(t), M]$ (here with $M = 3$) can drive the system into a sufficient frequency of the winning component of game $B$ (e.g., $B_2$) to cause the PE. Winning by playing losing games is only a seeming paradox as the possible values of $X(t)$ when both games are played are not equiprobable. Instead, they take on values that, for a range of biases in the coins, are favorable to the player, given the peculiar rules of game $B$. (One may also devise probabilities such that both games are winning, yet the combined game is losing, and so on. A more general definition of the PE includes such "negative" effects as well. This is more fully explicated in Appendix B and in Ref. [36]).

### 3.2 Single-player history-dependent Parrondo effect

Reference [37] extends the basic PE. Game $A$ remains as described above (a simple biased coin toss). Game $B$ is replaced with a history- as opposed to capital-dependent coin (game) defined by the respective winning/losing probabilities of four biased coins. A specific bias is associated with each of the four possible two-step binary histories (00, 01, 10, 11) of the player's wins (1) or losses (0). The choice of coin follows the history dependent rule:

| Before last $t-2$ | Last $t-1$ | History | Coin (Game) at $t$ | Prob. of win at $t$ | Prob. of loss at $t$ |
|---|---|---|---|---|---|
| Loss | Loss | 00 | $B_1$ | $q_1$ | $1-q_1$ |
| Loss | Win | 01 | $B_2$ | $q_2$ | $1-q_2$ |
| Win | Loss | 10 | $B_3$ | $q_3$ | $1-q_3$ |
| Win | Win | 11 | $B_4$ | $q_4$ | $1-q_4$ |

Both games $A$ and history-dependent games of type $B$ can be expressed as Markov transition matrices. But in this case $X(t)$, the evolution of the capital, is non-Markovian. To relate the capital to history one may therefore define the Markov chain

$$\vec{Y}(t) = \begin{pmatrix} X(t) - X(t-1) \\ X(t-1) - X(t-2) \end{pmatrix} \tag{3.1}$$

with the set of four states $\{(-1,-1), (-1,+1), (+1,-1), (+1,+1)\}$ with associated conditional probabilities and probability state vector $\{\pi_1(t), \pi_2(t), \pi_3(t), \pi_4(t)\} \equiv \vec{\pi}(t)$.

The transition matrix for game $B$ is therefore

$$\mathbf{B} = \begin{pmatrix} 1-q_1 & 0 & 1-q_3 & 0 \\ q_1 & 0 & q_3 & 0 \\ 0 & 1-q_2 & 0 & 1-q_4 \\ 0 & q_2 & 0 & q_4 \end{pmatrix} \tag{3.2}$$

and $\vec{\pi}(t+1) = \mathbf{B}\vec{\pi}(t)$. The stationary state distribution $\vec{\pi}_{st}$ obeys

$$\mathbf{B}\vec{\pi}_{st} = \vec{\pi}_{st} \tag{3.3}$$

with

$$\vec{\pi}_{st} = \frac{1}{N}\begin{pmatrix} (1-q_3)(1-q_4) \\ (1-q_4)\,q_1 \\ (1-q_4)\,q_1 \\ q_1 q_2 \end{pmatrix} \tag{3.4}$$

($N$ is a normalization factor; we assume a similar set of equations exists for $\mathbf{A}$, a simple coin-toss).

As explained in reference [37], though game $B$ as a whole is losing, the values of $\{q_1, q_2, q_3, q_4\}$ in $\mathbf{B}$ are such that $B_2$ and $B_3$ are independently losing, $B_1$ and $B_4$ winning. Then even if game $A$ is losing ($p < 1-p$), it perturbs the losing cycles of $B$ such that for certain values of $p$ and $\{q_1, q_2, q_3, q_4\}$ the winning games in $B$ dominate. This can occur when

$$\begin{gathered} (1-q_4)(1-q_3) > q_1 q_2 \\ (2-q_4-p)(2-q_3-p) < (q_1+p)(q_2+p)\,. \end{gathered} \tag{3.5}$$

For example, in reference [36], $p = 1/2-\varepsilon$ and $\{q_1, q_2, q_3, q_4\} = \{9/10-\varepsilon, 1/4-\varepsilon, 1/4-\varepsilon, 7/10-\varepsilon\}$. Then the conditions of (3.5) are met when $0 < \varepsilon < 1/168$.

Reference [38] extends the history-dependent PE further by showing that it may arise when game $A$ is redefined to have the same history-dependent structure as (3.2). A more complex set of equations define the conditions under which two losing games of this kind, each with four coins, generate winning results under random alternation. From this perspective the simple coin toss form for game $A$ in reference [37] may be reformulated with a specific set of parameters that fall within the more general parameter space analyzed in reference [38], viz.:

$$\mathbf{A} = \begin{pmatrix} 1-\left(\frac{1}{2}-\varepsilon\right) & 0 & 1-\left(\frac{1}{2}-\varepsilon\right) & 0 \\ \frac{1}{2}-\varepsilon & 0 & \frac{1}{2}-\varepsilon & 0 \\ 0 & 1-\left(\frac{1}{2}-\varepsilon\right) & 0 & 1-\left(\frac{1}{2}-\varepsilon\right) \\ 0 & \frac{1}{2}-\varepsilon & 0 & \frac{1}{2}-\varepsilon \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2}+\varepsilon & 0 & \frac{1}{2}+\varepsilon & 0 \\ \frac{1}{2}-\varepsilon & 0 & \frac{1}{2}-\varepsilon & 0 \\ 0 & \frac{1}{2}+\varepsilon & 0 & \frac{1}{2}+\varepsilon \\ 0 & \frac{1}{2}-\varepsilon & 0 & \frac{1}{2}-\varepsilon \end{pmatrix}. \quad (3.6)$$

For $\varepsilon = 0.005$, we obtain the following game matrices:

$$\mathbf{A} = \begin{pmatrix} 0.505 & 0 & 0.505 & 0 \\ 0.495 & 0 & 0.495 & 0 \\ 0 & 0.505 & 0 & 0.505 \\ 0 & 0.495 & 0 & 0.495 \end{pmatrix};$$

$$\mathbf{B} = \begin{pmatrix} 0.105 & 0 & 0.755 & 0 \\ 0.895 & 0 & 0.245 & 0 \\ 0 & 0.755 & 0 & 0.305 \\ 0 & 0.245 & 0 & 0.695 \end{pmatrix} \quad (3.7)$$

i.e., $\vec{\pi}^{(A)} = \{0.495, 0.495, 0.495, 0.495\}$ and $\vec{\pi}^{(B)} = \{0.895,$ $0.245, 0.245, 0.695\}$.

Solving the eigenvalue equation (3.3) for $\mathbf{B}$ and the equivalent for $\mathbf{A}$, we obtain the respective steady state probabilities for the two independent games:

$$\vec{\pi}_{st}^{(A)} = \begin{pmatrix} 0.255 \\ 0.250 \\ 0.250 \\ 0.245 \end{pmatrix}; \quad \vec{\pi}_{st}^{(B)} = \begin{pmatrix} 0.231 \\ 0.274 \\ 0.274 \\ 0.220 \end{pmatrix} \quad (3.8)$$

and the respective independent probabilities for winning:

$$P_{win}(A) = \vec{\pi}_{st}^{(A)} \bullet \vec{\pi}^{(A)} = 0.495$$
$$P_{win}(B) = \vec{\pi}_{st}^{(B)} \bullet \vec{\pi}^{(B)} = 0.494. \quad (3.9)$$

Naively, one might presume that with a mixing ratio of 1:1, a random alternation of the games would yield a winning probability equal to the mean of their winning probabilities, but this is not so, i.e. $P_{win}(1/2A, 1/2B) \neq 1/2[P_{win}(A) + P_{win}(B)] = 0.4945$. Instead, the winning probability is determined by the probabilities and steady-state vector of the mean of the transition matrices. As detailed more generally in Appendix B and reference [36], the winning probability of a combination of Markovian transition (game) matrices is not generally equal to the mean of their independent winning probabilities. Thus:

$$\vec{\pi}^{\left(\frac{1}{2}A, \frac{1}{2}B\right)} = \left\{\tfrac{1}{2}(p_1+q_1), \tfrac{1}{2}(p_2+q_2), \tfrac{1}{2}(p_3+q_3), \tfrac{1}{2}(p_4+q_4)\right\} = \{0.695, 0.370, 0.370, 0.595\} \quad (3.10)$$

so that

$$\frac{1}{2}(\mathbf{A}+\mathbf{B}) = \begin{pmatrix} 0.305 & 0 & 0.630 & 0 \\ 0.695 & 0 & 0.370 & 0 \\ 0 & 0.630 & 0 & 0.405 \\ 0 & 0.370 & 0 & 0.595 \end{pmatrix} \quad (3.11)$$

and

$$P_{win}\left(\tfrac{1}{2}A, \tfrac{1}{2}B\right) = \vec{\pi}_{st}^{(\frac{1}{2}A, \frac{1}{2}B)} \bullet \vec{\pi}^{(\frac{1}{2}A, \frac{1}{2}B)} = 0.501. \quad (3.12)$$

The winning probability is in this instance greater than either Pwin(A) or Pwin(B).

### 3.3 Multiple-player capital-dependent Parrondo effect and its reversal under optimization

Many variants of the PE have been studied, including capital-dependent multi-player PG (MPPG) [38,39]: At (every) time-step $t$, a constant-size subset of all participants is randomly re-selected actually to play. All participants keep individual track of their own capital but do not alternate games independently based on it. Instead, this data is used to select which game the participants must use at $t$. The chosen game is the one which, given the individual values of the capital at $t-1$ and the known matrices of the two games and their linear convex combination, has the most positive expected *aggregate* gain in capital, summed over all participants. This rule may be thought of as a static optimization procedure — static in the sense that the "optimal" choice appears to be known in advance. It appears exactly quantifiable because of access to each player's individual history. If the game is chosen at random, the change in wealth averaged over all participants is significantly positive. But when the "optimization" rule is employed, the gain becomes a loss significantly greater than that of either game alone. The intended "optimization" scheme actually reverses the positive (collective) PE. The reversal arises in this way: the "optimization" rule causes the system to spend much more time playing one of the games, and individually, any one game is losing.

### 3.4 Single-player capital-dependent Parrondo effect and its reversal under optimization

Here, we present a more natural illustration of the illusion-of-control in Parrondo games: while the MG is intrinsically collective, PGs are not. Neither the capital- nor the history-dependent variations require a collective setting for the PE to appear. Thus, the effect is most clearly demonstrated in a single-player implementation with two games under the most natural kind of optimization rule: at time $t$, the player plays whichever game has accumulated the most points (wealth) over a sliding window of $\tau$ prior time-steps from $t-1$ to $t-\tau$. Under this rule, a "current reversal" (reversal of a positive PE) appears. By construction, the individual games $A$ and $B$ played individually are both losing; random alternation between them is winning (the PE effect), but unexpectedly, choosing the previously best-performing game yields losses slightly less than either $A$ or $B$ individually: the PE is almost entirely eliminated. Furthermore, if instead the previously *worst* performing game is chosen, the player does better than either game and even much better than the PE from random game choice.

Under the choose-best optimization rule, two matrices $\mathbf{A}$ and $\mathbf{B}$ do not form a linear convex sum. Instead, the combined game is represented by an $8 \times 8$ transition matrix $\mathbf{M}$ with conditional winning probabilities:

$$m_j = \frac{1}{2}\left\{\pi^{(A)}_{\alpha(j)}\left[1+\pi^{(A)}_{\beta(j)}-\pi^{(B)}_{\beta(j)}\right] +\pi^{(B)}_{\alpha(j)}\left[1-\pi^{(A)}_{\beta(j)}+\pi^{(A)}_{\beta(j)}\right]\right\} \quad j=1,2,\dots 8. \quad (3.13)$$

The indices on the individual conditional probabilities for game $A$ and $B$, $\pi^{(A)}_i$, $\pi^B_i$; $i=1,2,\dots 4$ are converted to indices $\alpha(j)$ and $\beta(j)$ with $j=1,2,\dots,8$ by the following:

$$\alpha(j) = Mod\,[j-1,4]+1,$$
$$\beta[j] = \tfrac{1}{2}\left(j - Mod\,[j-1,2]+1\right). \quad (3.14)$$

For the "choose worst" rule, equation (3.13) is replaced by:

$$m_j = \frac{1}{2}\left\{\pi^{(A)}_{\alpha(j)}\left[1-\pi^{(A)}_{\beta(j)}+\pi^{(B)}_{\beta(j)}\right] +\pi^{(B)}_{\alpha(j)}\left[1+\pi^{(A)}_{\beta(j)}-\pi^{(A)}_{\beta(j)}\right]\right\} \quad j=1,2,\dots 8. \quad (3.15)$$

Alternated at random in equal proportion under the "choose best rule", $P^{best(A,B)}_{win} = 0.496$, while if "choose worst" is used, $P^{worst(A,B)}_{win} = 0.507$ (Compare to Eqs. (3.9) and (3.12)). The mechanism for this illusion-of-control effect characterized by the reversing of the PE under optimization is not the same as for the MG, as there is no collective effect and thus no-crowding out of strategies or games. Instead, the PE results from a distortion of the steady-state equilibrium distributions $\vec{\pi}^{(A)}_{st}$ and $\vec{\pi}^{(B)}_{st}$ into a vector $\vec{\pi}^{(1/2A,1/2B)}_{st}$ which is more co-linear to the conditional winning probability vector $\vec{\pi}^{(1/2A,1/2B)}$ than in the case of each individual game (this is just a geometric restatement of the fact that the combined game is winning). One can say that each game alternatively acts at random so as to better align these two vectors on average under the action of the other game. Choosing the previously best performing game amounts to removing this combined effect, while choosing the previously worst performing game tends to intensify it.

## 4 Conclusions

We have identified two classes of mechanisms operating in Minority games and in Parrondo games in which optimizing agents obtain suboptimal outcomes compared with non-optimizing agents. These examples suggest a general definition: the "illusion of control" effect occurs when low-entropy strategies (i.e. which use more information) under-perform random strategies (with maximal entropy). The illusion of control effect is related to bounded rationality as well as limited information [41] since, as we have shown, unbounded rational agents learn to converge to the symmetric mixed fully random strategies. It is only in the presence of bound rationality that agents can stick with an optimization scheme on a subset of strategies. Our robust message is that, under bounded rationality, the simple (large-entropy) strategies are often to be preferred over more complex elaborated (low-entropy) strategies. This is a message that should appeal to managers and practitioners, who are well-aware in their everyday practice that simple solutions are preferable to complex ones, in the presence of the ubiquitous uncertainty.

More examples should be easy to find. For instance, control algorithms, which employ optimal parameter estimation based on past observations, have been shown to generate broad power law distributions of fluctuations and of their corresponding corrections in the control process, suggesting that, in certain situations [42], uncertainty and risk may be amplified by optimal control. In the same spirit, more quality control in code development often decreases the overall quality which itself spurs more quality control leading to a vicious circle [43]. In finance, there are many studies suggesting that most fund managers perform worse than random [44] and strong evidence that over-trading leads to anomalously large financial volatility [45]. Let us also mention the interesting experiments in which optimizing humans are found to perform worse than rats [46]. We conjecture that the illusion-of-control effect should be widespread in many strategic and optimization games and perhaps in many real life situations. Our contribution is to put this question at a quantitative level so that it can be studied rigorously to help formulate better strategies and tools for management.

We are grateful to Riley Crane, Gilles Daniel and especially Yannick Malevergne for helpful feedback on the manuscript.

## Appendix A: Analytic methods and simulations for the Minority Game

### A.1 The Time-Horizon Minority Game: choosing the best strategy

In the simplest version of the Minority Game (MG) with $N$ agents, every agent has $S = 2$ strategies and $m = 2$. In the Time Horizon Minority Game (THMG), the point (or score) table associated with strategies is not maintained from the beginning of the game and is not ever growing. It is a rolling window of finite length $\tau$ (in the simplest case $\tau = 1$). The standard MG reaches an equilibrium state after a finite number of steps $t_{st}$. At this point, the dynamics and the behavior of individual agents for a given initial quenched disorder in the MG are indistinguishable from an otherwise identical THMG with $\tau \geq t_{st}$.

The fundamental result of the MG is generally cast in terms of system volatility: $\sigma^2/N$. All variations of agent and strategy reward functions depend on the negative sign of the majority vote. Therefore both agent and strategy "wealth" (points, whether "real" or hypothetical) are inverse or negative functions of the volatility: the lower the value of $\sigma^2/N$, the greater the mean "wealth" of the "system", i.e., of agents. However, this mean value is only rarely compared to the equivalent value for the raw strategies of which agents are composed. Yet agents are supposed to enhance their performance by choosing adaptively between their available strategies. In fact, the opposite is true in the THMG: the optimization method is strictly a method for worsening performance, not necessarily for every agent, but averaged over all agents and all strategies in a given $\hat{\mathbf{\Omega}}$, averaged over many $\hat{\mathbf{\Omega}}$.

To emphasize the relation of the THMG to market-games and the illusion of optimization, we transform the fundamental result of the THMG from statements on the properties of $\sigma^2/N$ to change in wealth, i.e., $\langle \Delta W/\Delta t\rangle$ for agents and $\langle \Delta W/\Delta t\rangle$ for strategies. We again use the simplest possible formulation — if an agent's actual (or strategy's hypothetical) vote places it in the minority, it scores +1 points, otherwise −1. Formally: at every discrete time-step $t$, each agent independently re-selects one of its $S$ strategies. It "votes" as the selected strategy dictates by taking one of two "actions," designated by a binary value:

$$a_i(t) \in \{1, 0\}, \quad \forall\, i, t. \tag{A.1}$$

The state of the system as a whole at time $t$ is a mapping of the sum of all the agents' actions to the integer set $\{2N_1 - N\}$, where $N_1$ is the number of 1 votes and $N_0 = N - N_1$. This mapping is defined as:

$$A(t) = 2\sum_{i=1}^{N} a_i(t) - N = N_1 - N_0. \tag{A.2}$$

If $A(t) > \frac{N}{2}$, then the minority of agents will have chosen 0 at time $t$ ($N_0 < N_1$); if $A(t) < \frac{N}{2}$, then the minority of agents will have chosen 1 at time $t$ ($N_1 < N_0$). The minority choice is the "winning" decision for $t$. This is then mapped back to $\{0, 1\}$:

$$D_{sys}(t) = -\mathrm{sgn}[A(t)] \quad \therefore D_{sys}(t) \in \{-1, +1\} \to \{0, 1\}. \tag{A.3}$$

For the MG, binary strings of length $m$ form histories $\mu(t)$, with $\dim[\mu(t)] = m$. For the THMG, binary strings of length $m + \tau$ form paths (or "path histories"), with $m + \tau = \dim(\mu_t)$, where we define $\mu(t)$ as a history in the standard MG and $\mu_t$ as a path in the THMG. Note that for memory m, there are $2^{2^m}$ possible strategies from which agents select $S$ at random. However as first detailed in reference [5], the space of strategies can be minimally spanned by a subset of all possible strategies. This reduced strategy space [RSS] has dimension $2^{m+1}$. As in reference [11] we may represent this quenched disorder in the allocation of strategies among agents (from the RSS) by a $\dim = \prod_{s=1}^{S} 2^{m+1}$ tensor, $\hat{\mathbf{\Omega}}$ (or from the full strategy space by a $\dim = \prod_{s=1}^{S} 2^{2^m}$ tensor). The $2^{m+1}$ (or $2^{2^m}$) strategies are arranged in numerical order along the edges of $\hat{\mathbf{\Omega}}$. Each entry represents the number of agents with the set of strategies indicated by the element's position. Then as demonstrated in [9], any THMG has a Markov chain formulation. For $\{m, S, N\} = \{2, 2, 31\}$ and using the RSS, the typical initial quenched disorder in the strategies attributed to each of the $N$ agents is represented by an $8 \times 8$ matrix $\hat{\mathbf{\Omega}}$ and its symmetrized equivalent $\hat{\mathbf{\Psi}} = 1/2(\hat{\mathbf{\Omega}} + \hat{\mathbf{\Omega}}^T)$. Positions along all $S$ edges of $\hat{\mathbf{\Omega}}$ represent an ordered listing of all available strategies. The numerical values $\Omega_{ij\ldots}$ in $\hat{\mathbf{\Omega}}$ indicate the number of times a specific strategy-tuple has been selected. (E.g., for two strategies per agent, $S = 2$, $\Omega_{2,5} = 3$ means that there are 3 agents with strategy 2 and strategy 5.) Without loss of generality, we may express $\hat{\mathbf{\Omega}}$ in upper-triangular form since the order of strategies in an agent has no meaning. The example (A.4) is a typical such tensor $\hat{\mathbf{\Omega}}$ for $S = 2$, $N = 31$

$$\hat{\mathbf{\Omega}} = \begin{pmatrix} 1 & 2 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 3 & 1 & 1 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{A.4}$$

Actions are drawn from a reduced strategy space (RSS) of dimension $2^m$. Each action is associated with a strategy $k$ and a path $\mu_t$. Together they can be represented in table form as a dim(RSS)$\times$dim($\mu_t$) binary matrix with elements converted for convenience from $\{0,1\} \rightarrow \{-1,+1\}$, i.e., $a_k^{\mu_t} \in \{-1,+1\}$. For $m=2$, $\tau=1$, $m+\tau=\dim(\mu_t)=3$, there are $2^3$ possible histories and $r=2^2$ reduced strategies (and $2^r$ strategies in total). In this case, the table of dimension dim(RSS) $\times$ dim($\mu_t$) coding for all possible reduced strategies and paths reads:

$$\hat{\mathbf{a}} \equiv \begin{pmatrix} -1 & -1 & -1 & -1 \\ -1 & -1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ -1 & +1 & +1 & -1 \\ +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & +1 & +1 & +1 \end{pmatrix}. \tag{A.5}$$

The change in wealth (point gain or loss) associated with each of the $2^r=8$ strategies for the 8 paths (=allowed transitions between the 4 histories) at any time $t$ is then:

$$\delta\vec{S}_{\mu(t),\mu(t-1)} = \left(\hat{\mathbf{a}}^{T}\right)_{\mu(t)} \{2Mod[\mu(t-1),2]-1\}. \tag{A.6}$$

$Mod[x,y]$ is "$x$ modulo $y$"; $\mu(t)$ and $\mu(t-1)$ label each of the 4 histories $\{00,01,10,11\}$ hence take on one of values $\{1,2,3,4\}$. Equation (A.6) picks out from (A.5) the correct change in wealth over a single step since the strategies are ordered in symmetrical sequence.

The change in points associated with each strategy for each of the allowed transitions between paths $\mu_t$ of the last $\tau$ time steps used to score the strategies is:

$$\vec{s}_{\mu_t} = \sum_{i=0}^{\tau-1} \delta\vec{S}_{\mu(t-i),\mu(t-i-1)}. \tag{A.7}$$

For example, for $m=2$ and $\tau=1$, the strategy scores are kept for only a single time-step. There is no summation so (A.7) in matrix form reduces to the score:

$$\vec{s}_{\mu_t} = \delta\vec{S}_{\mu(t),\mu(t-1)} \tag{A.8}$$

or, listing the results for all 8 path histories:

$$\hat{\mathbf{s}}_\mu = \delta\hat{\mathbf{S}}. \tag{A.9}$$

$\delta\hat{\mathbf{S}}$ is an $8\times 8$ matrix that can be read as a lookup table. It denotes the change in points accumulated over $\tau=1$ time steps for each of the 8 strategies over each of the 8 path-histories.

Instead of computing $A(t)$, we compute $A(\mu_t)$. Then for each of the $2^{m+\tau}=8$ possible $\mu_t$, $A(\mu_t)$ is composed of a subset of wholly determined agent votes and a subset of undetermined agents whose votes must be determined by a coin toss:

$$A(\mu_t) = A_D(\mu_t) + A_U(\mu_t). \tag{A.10}$$

Some agents are undetermined at time t because their strategies have the same score and the tie has to be broken with a coin toss. $A_U(\mu_t)$ is a random variable characterized by the binomial distribution. Its actual value varies with the number of undetermined agents. This number can be explicated (using an extension to the method employed in [9]) as:

$$N_U(\mu_t) = \left\{\left(1-\left[\hat{\mathbf{a}}^T_{(Mod[\mu_t-1,4]+1)} \otimes_\delta \hat{\mathbf{a}}^T_{(Mod[\mu_t-1,4]+1)}\right]\right) \circ (\vec{s}_{\mu_t} \otimes_\delta \vec{s}_{\mu_t}) \circ \hat{\mathbf{\Omega}}\right\}_{(Mod[\mu_t-1,2^m]+1)} \tag{A.11}$$

"$\otimes_\delta$" is a generalized outer product, with the product being the Kronecker delta. $\vec{N}_U$ constitutes a vector of such values. The summed value of all undetermined decisions for a given $\mu_t$ is distributed binomially. Similarly:

$$A_D(\mu_t) = \left(\sum_{r=1}^{8}\left\{\left[(1-\mathrm{sgn}\left[\vec{s}_{\mu_t} \ominus \vec{s}_{\mu_t}\right]) \circ \hat{\mathbf{\Psi}}\right] \bullet \hat{\mathbf{a}}\right\}_r\right)_{(Mod[\mu_t-1,2^m]+1)} \tag{A.12}$$

Details may also be found in reference [12]. We define $\vec{A}_D$ as a vector of the determined contributions to $A(t)$ for each path $\mu_t$. In expressions (A.11) and (A.12) $\mu_t$ numbers paths from 1 to 8 and is therefore here an index. $\vec{s}_{\mu_t}$ is the "$\mu_t^{th}$" vector of net point gains or losses for each strategy when at $t$ the system has traversed the path $\mu_t$ (i.e., it is the "$\mu_t^{th}$" element of the matrix $\hat{\mathbf{s}}_\mu = \delta\hat{\mathbf{S}}$ in (A.9)). "$\ominus$" is a generalized outer product of two vectors with subtraction as the product. The two vectors in this instance are the same, i.e., $\vec{s}_{\mu_t}$. "$\circ$" is Hadamard (element-by-element) multiplication and "$\bullet$" the standard inner product. The index $r$ refers to strategies in the RSS. Summation over $r$ transforms the base-ten code for $\mu_t$ into $\{1,2,3,4,1,2,3,4\}$. Selection of the proper number is indicated by the subscript expression on the entire right-hand side of (A.11). This expression yields an index number, i.e., selection takes place 1 + Modulo 4 with respect to the value of $(\mu_t - 1)$.

To obtain the transition matrix for the system as a whole, we require the $2^{m+\tau}\times 2^{m+\tau}$ adjacency matrix that filters out disallowed transitions. Its elements are $\Gamma_{\mu_t,\mu_{t-1}}$:

$$\hat{\mathbf{\Gamma}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{A.13}$$

Equations (A.11), (A.12) and (A.13) yield the history-dependent $(m+\tau)(m+\tau)$ matrix $\hat{\mathbf{T}}$ with elements $T_{\mu_t,\mu_{t-1}}$, representing the 16 allowed probabilities of transitions between the two sets of 8 path-histories $\mu_t$ and $\mu_{t-1}$:

$$T_{\mu_t,\mu_{t-1}} = \Gamma_{\mu_t,\mu_{t-1}} \sum_{x=0}^{N_U(\mu_t)} \left\{ \underset{x}{\overset{N_U(\mu_t)}{\mathbb{C}}} \left(\frac{1}{2}\right)^{N_U(\mu_t)} \times \delta\Big[\mathrm{sgn}\left(A_D(\mu_t) + 2x - N_U(\mu_t)\right) + \left(2Mod\{\mu_{t-1}, 2\} - 1\right)\Big]\right\}. \quad (A.14)$$

The $\mathbb{C}$ expression $\underset{x}{\overset{N_U(\mu_t)}{\mathbb{C}}} \left(\frac{1}{2}\right)^{N_U(\mu_t)}$ in (A.14) represents the binomial distribution of undetermined outcomes under a fair coin-toss with mean $= A_D(\mu_t)$. Given a specific $\hat{\mathbf{\Omega}}$,

$$\langle A(\mu_t)\rangle = A_D(\mu_t) \ \forall \ \mu_t. \quad (A.15)$$

We now tabulate the number of times each strategy is represented in $\hat{\mathbf{\Omega}}$, regardless of coupling (i.e., of which strategies are associated in forming agent $S$-tuples):

$$\vec{\kappa} \equiv \sum_{k=1}^{2^{m+\tau}} \left(\hat{\mathbf{\Omega}} + \mathbf{\Omega}^T\right)_k = 2\sum_{k=1}^{2^{m+\tau}} \hat{\mathbf{\Psi}}_k = \{n(\sigma_1), n(\sigma_2), \ldots n(\sigma_{2^{m+\tau}})\} \quad (A.16)$$

where $\sigma_k$ is the $k$th strategy in the RSS, $\hat{\mathbf{\Omega}}_k$, $\hat{\mathbf{\Omega}}_k^T$ and $\hat{\mathbf{\Psi}}_k$ are the $k$th element (vector) in each tensor and $n(\sigma_k)$ represents the number of times $\sigma_k$ is present across all strategy tuples. Therefore

$$\langle \Delta W_{Agent}\rangle = -\frac{1}{N} Abs\left(\vec{A}_D\right) \cdot \vec{\mu} \quad (A.17)$$

and

$$\langle \Delta W_{Strategy}\rangle = \frac{1}{2N}\left(\hat{\mathbf{s}}_\mu \cdot \vec{\kappa}\right) \cdot \vec{\mu} \quad (A.18)$$

with $\vec{\mu}$ the normalized steady-state probability vector for $\hat{\mathbf{T}}$. Expression (A.17) states that the mean per-step change in wealth for agents equals $-1$ times the probability-weighted sum of the (absolute value of the) *determined* vote imbalance associated with a given history. Expression (A.18) states that the mean per-step change in wealth for individual strategies equals the probability-weighted sum of the representation of each strategy (in a given $\hat{\mathbf{\Omega}}$) times the sum over the per-step wealth change associated with every history. The $-1$ in (A.17) reflects the minority rule. I.e., the awarding of points is the negative of the direction of the vote imbalance. No minus sign is required in (A.18) as it is already accounted for in (A.5).

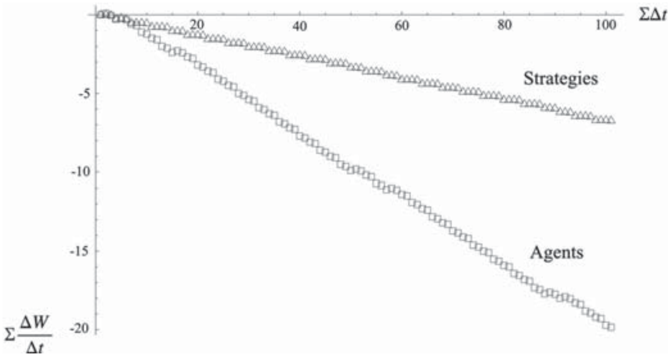


**Fig. A.1.** Mean strategy versus agent cumulative change in wealth in the THMG. $\{m, S, N\} = \{2, 2, 31\}$; 100 time steps.

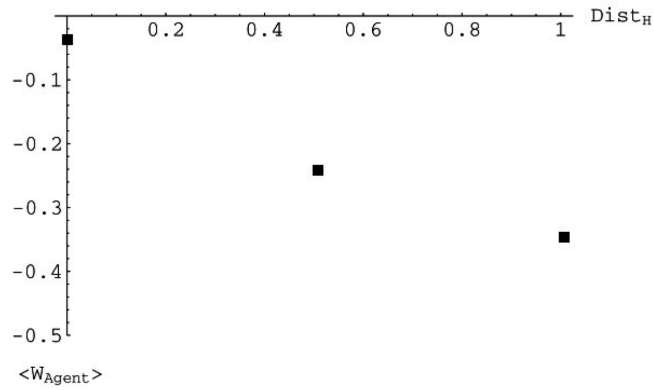


**Fig. A.2.** Agent wealth as a function of Hamming distance between strategy pairs in agents for the example simulation.

Figure A.1 shows the cumulative mean change in wealth for strategies versus agents over time, given (A.4).

As first studied in [13,14], and discussed in the body of the manuscript, agent performance is inversely proportional to the Hamming distance between strategies within agents. With the variation expected of a single example, our sample $\hat{\mathbf{\Omega}}$ given by (A.4) reproduces this relation as shown in Figure A.2. The mean over many $\hat{\mathbf{\Omega}}$ corresponds to a "flat" $\hat{\mathbf{\Omega}}$.

## A.2 The Minority Game: choosing the worst strategy

First, we re-cast the initial quenched disorder on the set of strategies attributed to the $N$ agents in a given game realization as a two-component tensor $\hat{\mathbf{\Omega}} = \{\hat{\mathbf{\Omega}}^+, \hat{\mathbf{\Omega}}^-\}$. $\hat{\mathbf{\Omega}}^+$ represents standard (S) agents that adapt as before; $\hat{\mathbf{\Omega}}^-$ represents "counteradaptive" (C) agents that instead select their worst-performing strategies. In our example (A.4) then, suppose we select at random 3 agents

**Table A.1.** Numerical/analytic results of THMG with and without 3 C Agents 28 S Agents (=left value/right value).

| | $\langle \Delta W_{Agent} \rangle$ | $\langle \Delta W_{Strategy} \rangle$ |
|---|---|---|
| With | –0.14/–0.14 | –0.05/–0.05 |
| Without | –0.26/–0.26 | –0.05/–0.05 |

to use the C rule, one each at $\Omega_{1,2}$, $\Omega_{2,6}$ and $\Omega_{7,8}$:

$$\hat{\mathbf{\Omega}} = \left\{ \hat{\mathbf{\Omega}}^+, \hat{\mathbf{\Omega}}^- \right\}$$
$$= \left\{ \begin{pmatrix} 1\,1\,0\,0\,1\,1\,0\,0 \\ 0\,0\,0\,0\,3\,2\,1\,1 \\ 0\,0\,2\,0\,1\,0\,0\,0 \\ 0\,0\,0\,1\,1\,0\,0\,1 \\ 0\,0\,0\,0\,1\,0\,2\,1 \\ 0\,0\,0\,0\,0\,2\,2\,1 \\ 0\,0\,0\,0\,0\,0\,2\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \end{pmatrix}, \begin{pmatrix} 0\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,1\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0\,0\,0\,0 \end{pmatrix} \right\}. \tag{A.19}$$

For any number of C agents in $\hat{\mathbf{\Omega}}$ thus redefined, the analytic expression for $\hat{\mathbf{T}}$ need only be modified by decomposing $A_D(\mu_t)$ accordingly. The new term in $A_D(\mu_t)$ makes evident the symmetry of the C rule with respect to the S rule, and the lack of privilege of C agents. Thus:

$$A_D(\mu_t) = \left( \sum_{r=1}^{8} \left\{ \left[ (1 - \mathrm{sgn}[\vec{s}_{\mu_t} \ominus \vec{s}_{\mu_t}]) \circ \hat{\mathbf{\Psi}}^+ + (1 + \mathrm{sgn}[\vec{s}_{\mu_t} \ominus \vec{s}_{\mu_t}]) \circ \hat{\mathbf{\Psi}}^- \right] \bullet \hat{\mathbf{a}}_1 \right\}_r \right)_{(Mod[\mu_t - 1, 2^m]+1)} \tag{A.20}$$

with

$$\hat{\mathbf{\Psi}}^+ = \frac{1}{2}\left( \hat{\mathbf{\Omega}}^+ + \hat{\mathbf{\Omega}}^{+^T} \right); \quad \hat{\mathbf{\Psi}}^- = \frac{1}{2}\left( \hat{\mathbf{\Omega}}^- + \hat{\mathbf{\Omega}}^{-^T} \right). \tag{A.21}$$

The number of undetermined agent votes remains unchanged. In (A.11), $\hat{\mathbf{\Omega}}$ need only be replaced with ($\hat{\mathbf{\Omega}}^+ + \hat{\mathbf{\Omega}}^-$):

$$N_U(\mu_t) = \left\{ \left( 1 - \left[ \left( \hat{\mathbf{a}}_1^T \right)_{(Mod[\mu_t - 1,4]+1)} \otimes_\delta \left( \hat{\mathbf{a}}_1^T \right)_{(Mod[\mu_t - 1,4]+1)} \right] \right) \circ (\vec{s}_{\mu_t} \otimes_\delta \vec{s}_{\mu_t}) \circ \left( \hat{\mathbf{\Omega}}^+ + \hat{\mathbf{\Omega}}^- \right) \right\}_{(Mod[\mu_t - 1, 2^m]+1)}. \tag{A.22}$$

Results for numerical simulation and analytic calculation are in close agreement even for a single short simulation, as illustrated in Table A.1.

The 3 C agents of 31 now perform so well that they significantly raise the overall performance of the system as detailed in Figure A.3. They not only outperform both their own strategies and the other $S$ agents on average, they generate net positive gain. The *hypothetical* outperformance of unused relative to used strategies in the MG was first observed in [14]. But the explicit generation of positive results, by agents simply deploying their unused strategies (without privileging), has not been tested. (In the case of $S = 2$, "unused" are by definition the "worst-performing".)

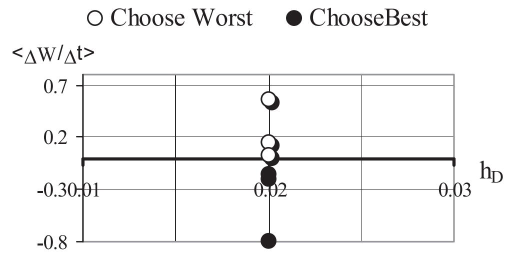


**Fig. A.3.** Average wealth variation per time step for different agents. White circles represent the wealth variations of the three among the 31 agents which use counteradaptive ("C", choose worst) strategy selection. The usual underperformance of agents compared to individual strategies when using standard selection rule ("S", choose best) is shown in black.

We discuss in the manuscript and elsewhere the phenomena that arise as the proportion of $S$ and $C$ agents are varied for different parameters of the MG. We emphasize here only the fact that extensive numerical studies confirm that the phenomenon here illustrated persist over a very wide range of parameters for the MG and quite generally in the THMG.

### A.3 The Minority Game: random agents

We provide in Figure A.4 some numerical results for a THMG game with $N$ agents total, $N_R$ of which employ a fully random symmetric choice. The remaining $N_S = N - N_R$ "special" agents (with $N_R \gg N_S$) will all be one of two possible types: (i) anti-optimizing agents with $S$ fixed strategies that choose their worst performing strategy to make the decision at the next step (referred to above as counteradaptive); (ii) agents with a single fixed strategy. We use the simplest example, that of $N_S = 1$ (with $\tau = 1$), to illustrate the fact that, in the THMG, agents allowed/restricted to a fully symmetric random choice outperform agents that attempt to optimize. (Note that the outperformance and absolute positive returns, associated with a small proportion of anti-optimizing agents, requires *the remaining agents to optimize*, as described above. Here the small proportion of optimizing and anti-optimizing agents compete with fully random agents.)

## Appendix B: Analytic methods for the general Parrondo effect

Consider $N > 1$ $s$-state Markov games $G_i$, $i \in \{1, 2, \ldots, N\}$, and their $N$ $s \times s$ transition matrices, $\hat{\mathbf{M}}^{(i)}$.

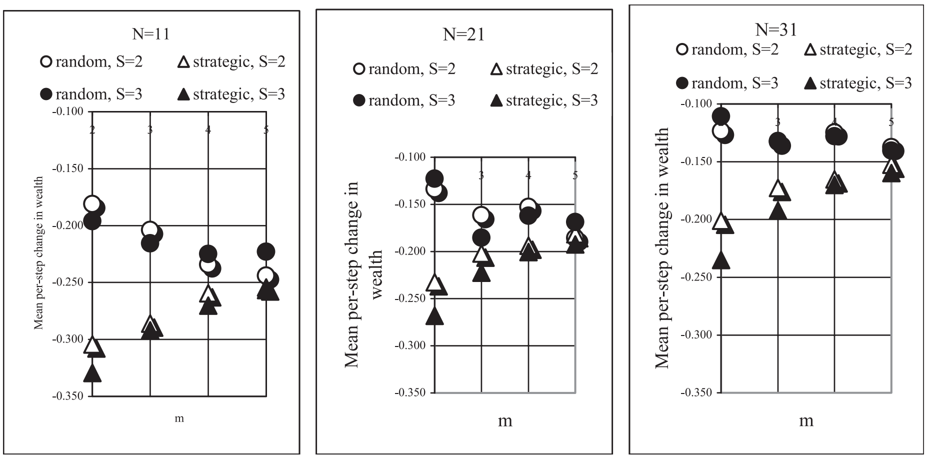


**Fig. A.4.** Performance (mean change in wealth per step) of a single optimizing agent versus all other agents making a symmetric random choice in a MG-like game. From left to right $n = 11$, 21, 31. $S = 2$, 3 $m = 2$, 3, 4, 5 and $\tau = 1$. Random agents always outperform optimizing agents. Similar results are found for other values of $n$, $m$, $S$ and $\tau$. Within statistical fluctuations typical for the number of runs/random selection of strategies comprising the optimizing agent (100 runs), results for anti-optimizing agents are identical.

For every $\hat{\mathbf{M}}^{(i)}$, denote its vector of $s$ winning probabilities conditional on each of the $s$-states as $\vec{p}^{(i)} = \{p_1^{(i)}, p_2^{(i)}, \ldots p_s^{(i)}\}$ and its steady-state equilibrium distribution vector as $\vec{\Pi}^{(i)} = \{\pi_1^{(i)}, \pi_2^{(i)}, \ldots, \pi_s^{(i)}\}$. For each game, the steady-state probability of winning is therefore $P_{win}^{(i)} = \vec{p}^{(i)} \cdot \vec{\Pi}^{(i)}$. Consider also a sequence of randomly alternating $G_i$ with individual time-averaged proportion of play $\gamma_i \in [0,1]$, $\sum_{i=1}^{N} \gamma_i = 1$. The transition matrix for the combined sequence of games is the convex linear combination $\hat{\mathbf{M}}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)} \equiv \sum_{i=1}^{N} \gamma_i \hat{\mathbf{M}}^{(i)}$ with conditional winning probability vector $\vec{p}^{(\gamma_1,\gamma_2,\ldots,\gamma_n)} = \sum_{i=1}^{n} \gamma_i \vec{p}^{(i)}$ and steady-state probability vector $\vec{\Pi}^{(\gamma_1,\gamma_2,\ldots,\gamma_n)}$ (which is a complex nonlinear mixture of the $\vec{\Pi}^{(i)}$'s). The steady-state probability of winning for the combined game is therefore

$$P_{win}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)} = \vec{p}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)} \cdot \vec{\Pi}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)} \qquad \text{(B.1)}$$

A PE occurs whenever (and in general it is the case that)

$$\sum_{i=1}^{N} \gamma_i P_{win}^{(i)} \neq P_{win}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)}, \text{ i.e., } \sum_{i=1}^{N} \gamma_i \vec{p}^{(i)} \cdot \vec{\Pi}^{(i)} \neq \vec{p}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)} \cdot \vec{\Pi}^{(\gamma_1,\gamma_2,\ldots,\gamma_N)} \qquad \text{(B.2)}$$

hence the PE, or "paradox", when the left hand sides of (B.2) are less than zero and the right-hand sides greater.

## References


1. J.D. Taylor, S.E. Brown, Psych. Bull. **103**, 193 (1988)
2. A. Bandura, *Self-efficacy: the exercise of control* (WH Freeman, New York, 1997)
3. E. Langer, J. Pers. Soc. Psych. **7**, 185 (1975)
4. D. Challet, Y.C. Zhang, Phys. A **246**, 407 (1997)
5. D. Challet, Y.C. Zhang, Phys. A **256**, 514 (1998)
6. D. Challet, M. Marsili, Y.-C. Zhang, Phys. A **276**, 284 (2000)
7. J.S. Doran, C. Wright, *What really matters when buying and selling stocks?* Working paper of Florida State University, 2007) (`http://ssrn.com/abstract=980291`)
8. J.M.R. Parrondo, *How to cheat a bad mathematician*, in EEC HC&M Network on Complexity and Chaos (#ERBCHRX-CT940546) (ISI, Torino, Italy, 1996), Unpublished
9. G.P. Harmer, D. Abbott, Stat. Sci. **14**, 206 (1999)
10. W.B. Arthur, *Out-of-Equilibrium Economics and Agent-Based Modeling*, in the *Handbook of Computational Economics*, Vol. 2: *Agent-Based Computational Economics*, edited by K. Judd, L. Tesfatsion (Elsevier/North-Holland, 2005)
11. M. Hart, P. Jefferies, P.M. Hui, N.F. Johnson, Eur. Phys. J. B **20**, 547 (2001)
12. M.L. Hart, P. Jefferies, N.F. Johnson, Phys. A **311**, 275 (2002)
13. (The mathematical derivation is given in Appendix A)
14. R. D'Hulst, G.J. Rodgers, Phys. A **270**, 514 (1999)
15. Y. Li, R. Riolo, R. Savit, Phys. A **276**, 265 (2000)
16. M. Andrecut, M.K. Ali, Phys. Rev. E **64**, 67103 (2001)
17. R.M. Araujo, L.C. Lamb, *Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence* (ICTAI'04)-Volume 00, 727 (2004)
18. W.C. Man, H.F. Chau, Phys. Rev. E **73**, 36106 (2006)
19. M. Sysi-Aho, A. Chakraborti, K. Kaski, Eur. Phys. J. B **34**, 373 (2003)
20. M. Sysi-Aho, A. Chakraborti, K. Kaski, Phys. Rev. E **69**, 36125 (2004)
21. W.S. Yang, B.H. Wang, Y.L. Wu, Y.B. Xie, Phys. A **339**, 583 (2004)
22. J. Menche, J.R.L. de Almeida, *It is worth thinking twice or Improving the performance of minority games*, e-print `arXiv:preprint cond-mat/0308181` (2003)

23. M. Marsili, D. Challet, Adv. Complex Systems. **3-I**, 3 (2001)
24. D. Challet, M. Marsili, R. Zecchina, Phys. Rev. Lett. **84**, 1824 (2000)
25. J. Duffy, E. Hopkins, Games Econ. Behav. **51**, 31 (2005)
26. M. Smoluchowski, Z **13**, 1069 (1912); R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman lectures on physics* (Addison-Wesley Redwood City, Calif., 1963)
27. A. Adjari, J. Prost, C. R. Acad. Sci. Paris II **315**, 1635 (1992); M.O. Magnasco, Phys. Rev. Lett. **71**, 1477 (1993)
28. J. Prost, J.-F. Chawin, L. Peliti, A. Adjari, Phys. Rev. Lett. **72**, 2652 (1994)
29. R. Pyke, *On Random Walks and Diffusions Related to Parrondo's Games*, e-print `arXiv:math.PR/0206150` (2002)
30. D. Kinderlehrer, M. Kowalczyk, Arch. Rat. Mech. Anal. **161**, 149 (2002)
31. H. Moraal, J. Phys. A **33**, L203 (2000)
32. G.P. Harmer, D. Abbott, P.G. Taylor, J.M. Parrondo, Chaos **11**, 705 (2001)
33. S. Maslov, Y.C. Zhang, Int. J. Theor. Appl. Finan. **1**, 377 (1998)
34. M. Boman, S.J. Johansson, D. Lyback, *Parrondo strategies for artificial traders*, in *Proceedings of the 2nd International Conference on Intelligent Agent Technology* (World Scientific, 2001)
35. W.S. Almberg, M. Boman, *in Artificial Intelligence and Computer Science*, edited by S. Shannon (Nova Science Publishers, 2005), p. 123
36. J.B. Satinover, D. Sornette, Phys. A **386**, 339 (2007)
37. J.M.R. Parrondo, G.P. Harmer, D. Abbott, Phys. Rev. Lett. **85**, 5226 (2000)
38. R.J. Kay, N.F. Johnson, Phys. Rev. E **67**, 56128 (2003)
39. L. Dinis, Europhys. Lett. **63**, 319 (2003)
40. J.M.R. Parrondo, L. Dinis, J. Buceta, K. Lindenberg, *Paradoxical games, ratchets, and related phenomena, In Advances in Condensed Matter and Statistical Mechanics*, edited by E. Korutcheva, R. Cuerno (Nova Science Publishers, 2003)
41. J. Huber, M. Kirchler, M. Sutter, *Is more information always better? Experimental financial markets with cumulative information*, Journal of Economic Behavior and Organization, forthcoming (2007)
42. C.W. Eurich, K. Pawelzik, in *Artificial Neural Networks: Formal Models and Their Applications* - ICANN 2005 (Springer, Berlin, 2005), p. 0302
43. S.P. Berczuk, B. Appleton, *in Software Configuration Management Patterns: Effective Teamwork, Practical Integration* (Addison-Wesley, 2002)
44. B.G. Malkiel, *A Random Walk Down Wall Street:: the Time-tested Strategy for Successful Investing* (WW Norton & Company, 2003)
45. R.J. Shiller, *Market Volatility* (MIT Press, 1992)
46. T. Grandin, C. Johnson, *Animals in Translation: Using the Mysteries of Autism to Decode Animal Behavior* (Scribner, 2005)