

Strategy switches and co-action equilibria in a minority game

V. Sasidevan^{1,*} and Deepak Dhar^{1,†}

¹*Department of Theoretical Physics, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India*

(Dated: September 15, 2013)

We propose an analytically tractable variation of the minority game in which rational agents use probabilistic strategies. In our model, N agents choose between two alternatives repeatedly, and those who are in the minority get a pay-off 1, others zero. The agents optimize the expectation value of their discounted future pay-off, the discount parameter being λ . We propose an alternative to the standard Nash equilibrium, called co-action equilibrium, which gives higher expected pay-off for all agents. The optimal choice of probabilities of different actions are determined exactly in terms of simple self-consistent equations. The optimal strategy is characterized by N real parameters, which are non-analytic functions of λ , even for a finite number of agents. The solution for $N \leq 7$ is worked out explicitly indicating the structure of the solution for larger N . For large enough future time horizon, the optimal strategy switches from random choice to a win-stay lose-shift strategy, with the shift probability depending on the current state and λ .

PACS numbers:

I. INTRODUCTION

There has been a lot of interest in applying techniques of statistical physics to economics in the past two decades, in particular for a better understanding of the behaviour of fluctuations in systems with many interacting agents, as in a market. A prototypical model is the El Farol bar problem [1] in which agents optimize their personal pay-offs by guessing what other agents would be doing. A particular realization of this is the Minority Game (MG) introduced in 1997 by Challet and Zhang [2]. It has been described by Arthur as a classic ‘model problem that is simple to describe but offers a wealth of lessons’ [3]. In this model, an odd number of agents repeatedly make choices between two alternatives, and at each step the agents who belong to the minority group are considered as winners. In MG, the agents cannot communicate with each other, and base their decision on the common-information which is the history of the winning choices in the last few days.

Each agent has a small number of predictors termed as strategies available with her, and at any time uses her best-performing strategy to decide her immediate future action. The agents are adaptive, and if they find that the strategy they are using is not working well, they will change it. This in turn affects the performance of other agents, who may then change their strategies and so on. Thus, this provides a very simple and instructive model of learning, adaptation, self-organization and co-evolution in a group of interacting agents.

Simulation studies of this model showed that the agents self-organize into a rather efficient state where there are more winners per day than would be expected if agents made the choice by a random throw of a coin, for a range of the parameters of the model. This has led to a lot of interest in the model, and soon after the original paper of Challet and Zhang, a large number of papers appeared, discussing several aspects of the model, or variations. Several good reviews are available in the literature [10, 11], and there are excellent monographs that describe the known results [4, 9]. It is one of the few non-trivial models of interacting agents that is also theoretically tractable.

However, one would like to understand how well a particular strategy for learning and adaptation works, and compare it with alternate strategies. This strategy to select strategies may be called a meta-strategy. Clearly, the meta-strategy that gives better pay-off to its user will be considered better. In this respect, the meta-strategy used in MG does not work so well. While in some range of parameters, the agents are found to self-organize into a globally efficient state, in other regions of the parameter space (for large number of agents), its overall efficiency is worse than if the agents simply chose the restaurants at random. This is related to the fact that in MG, agents use deterministic strategies, and each agent has only a limited number of deterministic strategies available to her. Also, the rule to select the strategy to use, in terms of performance scores of strategies, is known to be not very effective. In fact, simulation studies have shown [5] that if an agent always selects from her strategies the one with worst performance

*Electronic address: sasi@theory.tifr.res.in

†Electronic address: ddhar@theory.tifr.res.in

score, she does better than other agents who are using the usual MG rule of selecting the strategy with the best performance score!

It seems worthwhile, if only to set a point of reference, to determine how well agents in a Minority-like game could do, if they use some other meta-strategy. In this paper, we will study a model where the agents use mixed strategies. In our formulation, each agent, on each day, selects a probability p , and then generates a new random number uniformly between 0 and 1, and switches her choice from the previous day if it is $\leq p$. The choice of p depends on the history of the game, and her own history of pay-offs in the last m days, and constitutes the strategy of the agent [The deterministic strategies are special cases when p is 0 or 1]. We discuss how the optimal value of this parameter depends on the history of the game.

For this purpose, we propose a new solution concept, as an alternative to the usual notion of Nash equilibrium [8]. We show that the Nash equilibrium states are not very satisfactory for our model, giving rise to ‘trapping states’ (discussed below), and our proposed alternative, to be called co-action equilibrium, avoids this problem. To distinguish it from the original Minority game, we will refer to the new game as the Co-action Minority game (CAMG), and refer to the original MG as Challet-Zhang Minority Game (CZMG).

In CZMG, on any day, each agent selects one strategy from a small set of deterministic strategies given to her at the beginning of the game. We make the basket of strategies given to the agents much bigger, and make all strategies, within a specified infinite class, available to all agents. The use of stochastic, instead of the deterministic, strategies makes the CAMG more efficient than CZMG. Also, the absence of quenched disorder - in the form of assigning strategies to agents in the beginning of the game - in our model makes it much more tractable. One can determine the behaviour of many quantities of interest in more detail, using only elementary algebra. The theoretical analysis of CZMG requires more sophisticated mathematical techniques such as functional integrals, and the limit of large number of agents, large backward time horizon, and large times (explained later in the paper).

We find that the optimal strategies of agents can be determined by a mean-field theory like self-consistency requirement. For the N -agents case, we get coupled algebraic equations in N variables [7]. Also, the agents organize themselves into a globally efficient state quickly, in time of order $\log N$ days, compared to time of order N days in CZMG. The simplicity of our analysis makes this model an interesting and instructive analytically tractable model of interacting agents. Interestingly, this also provides us with a non-trivial example of a non-equilibrium steady state which shows a non-analytic dependence on a control parameter *even for finite number of agents*.

The plan of the paper is as follows: In Sec. II, we recapitulate the main features of the CZMG, and what is known about its behaviour. In Sec. III, we introduce the CAMG game. In Sec. IV, we show that the model has Nash equilibrium states that are trapping states, where all agents stay with the same choice next day, and the system gets into a frozen state. In Sec. V, we introduce the solution concept of co-action equilibrium to avoid these trapping states. Sec. VI develops the general theoretical framework of Markov chains to calculate the expected pay-off functions of agents in CAMG, which is used to determine the optimal strategies by agents. In Sec VII, we work out explicitly, the optimal strategies when the number of agents $N = 3, 5$ and 7. The behaviour for larger N is indicated. Sec. VIII contains a summary of our results, and some concluding remarks.

II. THE CHALLET-ZHANG MINORITY GAME

In CZMG [2], each of the N agents, with N odd, has to choose between two alternatives, say two restaurants A and B, on each day and those in the restaurant with fewer people get a pay-off 1, and others 0. The agents cannot communicate with each other, and make their choice based only on the information of which was the minority restaurant for each of the last m days. A strategy gives which one of the two choices (A or B) is preferred, for each of the 2^m possible histories of the game. The total number of possible strategies is 2^{2^m} . Each agent has a small fixed number k of strategies randomly picked out of all possible strategies at the beginning of the game. For each of the strategies assigned to an agent, she keeps a performance score which tells how often in the past the strategy correctly predicted the winning choice. On each day, she decides which restaurant to go to, using the strategy that performed best in the recent past.

We write $N = 2M + 1$. Clearly, on any day, the number of people that are happy (i.e. have a positive pay-off) is $\leq M$. The amount by which the average number of happy people per day differs from the maximum possible value M is a measure of the social inefficiency of the system. For a system of agents in a steady state \mathbb{S} , we will characterize the inefficiency of the system in terms of a parameter η , called the inefficiency parameter, defined as,

$$\eta_{\mathbb{S}} = \frac{W_{max} - \langle W \rangle_{\mathbb{S}}}{W_{max} - W_{rand}} \quad (1)$$

where $W_{max} = M/N$ is the maximum possible pay-off per agent, $\langle W \rangle_{\mathbb{S}}$ is the average pay-off per agent in the steady state \mathbb{S} , and W_{rand} is the average pay-off per agent when agents select randomly between A and B.

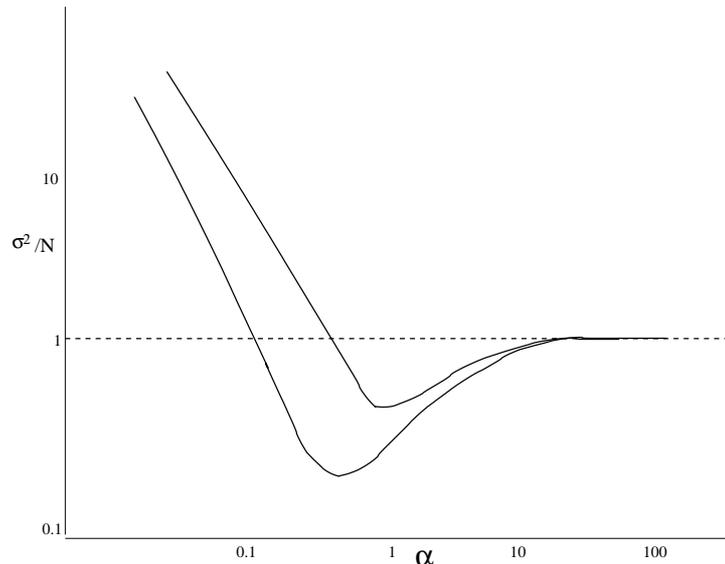


FIG. 1: Schematic representation of the variation of the fluctuation in the attendance difference between the two restaurants σ^2/N with the parameter α for two different values of k . Upper curve corresponds to the larger k value. The dashed horizontal line shows the value of η when agents choose randomly between the two restaurants.

The general qualitative behaviour of MG is quite well understood from simulations. Fig. 1 shows the schematic behaviour of the system as seen in simulations. The theoretical analysis is rather complicated, and involves several limits: large N and m , with $2^m/N = \alpha$ held fixed. Also, one has to rescale time with N , and the exact theoretical results are possible for fixed $\tau = t/N$. The asymptotic behaviour in the steady state can be determined exactly only in these limits, and only for α greater than a critical value α_c , using concepts and formalism developed originally for the spin-glass problem. For a more detailed discussion, see [4, 9].

III. THE CO-ACTION MINORITY GAME

We would like to construct a game which preserves the basic simplicity of the Challet-Zhang minority game, but changes it in several important ways, to make it more tractable. We will keep the allowed actions, and pay-off function the same, but consider different strategies used by agents. We discuss these changes one by one.

A. Stochastic versus deterministic strategies

It is well-known that in repeated games where agents have to make their choices simultaneously, probabilistic strategies are much more effective than deterministic ones. In fact, Arthur, in the forward of [4], recalls that when he introduced the El-Farol Bar problem at a conference, the session chair Paul Krugman had objected that the problem has a simple efficient strategy, where each agent uses a mixed strategy, and decides between the two options by tossing a coin. In the CAMG, we allow the agents to use mixed strategies. We will show that this results in a different emergent behaviour of the system.

In the CZMG, the agents are assigned a small number of strategies at random, and different agents have different basket of strategies. In CAMG, we allow each agent to choose his shift probability to be any value p , with the only constraint being $0 \leq p \leq 1$. Since the choice of p constitutes the strategy of an agent, each agent is allowed to choose from an infinite set of strategies. Also, the same set of strategies is available to all the agents. Thus we do not have any quenched disorder in the model, and this simplifies the analytical treatment of the model considerably.

Mixed strategies in the context of minority game have been discussed before. An example is the thermal minority game [6]. A somewhat similar model to ours, involving a minority game, with probabilistic strategies was studied earlier by Reents et al [13]. However, there is an important difference between these earlier studies, and ours. In the earlier studies, the probabilities of different actions was thought to be due to a kind of noise in the system, not under the control of the agents. The probability of ‘non-optimal’ choice is externally prescribed in the beginning. In our

model, the agent is free to choose the value of p , and chooses it to maximize her pay-off. Also, the agent's choice can vary from day to day.

B. Rational versus adaptive agents

Another important difference from the CZMG is that we will treat our agents as intelligent and rational who also expect other agents to act rationally. The agents are selfish, and choose their actions only to maximize their personal expected gain. This is an important difference from CZMG, where one of the motivations for introducing the game was to model the behaviour of agents with only bounded rationality who resort to inductive reasoning. In CZMG, the agents follow rather simple rules to decide when to switch their strategies, based on the performance scores of strategies. One may imagine that these agents are unthinking machines, following some pre-programmed instructions.

We will also assume that all agents are equally intelligent. Thus, there is no built-in heterogeneity in the behaviour of agents. When the agents follow deterministic strategies as in CZMG, they are forced to differentiate amongst themselves in the strategies they use, as many agents following the same strategy is clearly not good. When agents use mixed strategies, the strategies need not be different, as the differentiation is naturally provided by the random number generators used by the agents. In fact, this differentiation is rather efficient. We will see that even without any assumed heterogeneity of agents, the system shows non-trivial emergent behaviour, and reaches a more efficient state quicker.

We need not discuss here the question whether full rationality or bounded rationality can describe the behaviour of real-life agents. Clearly, there would be situations where one or the other is a better model. We only note that the general probabilistic 'win-stay lose-shift' strategy has been seen to be used in many learning situations [14], which is the strategy found to be optimal by rational agents in CAMG.

C. Agents' optimization aim

The next issue is deciding the pay-off function that is optimized by the rational agents. Clearly, maximizing the probability of winning next day is a possible goal. But agents in repeated games need not be concerned only about their immediate payoffs.

We note that if all agents had the same optimization goal of minimizing the system inefficiency, there is no competition between them, and we get a trivial game. There is a simple strategy which will give the best possible result of long-term average of the inefficiency parameter being zero. In this strategy, each day, agents choose some shift probability strictly between 0 and 1, until one finds M people in one restaurant, and $M + 1$ in the other. Once this state is reached, each agent goes to the same restaurant on all subsequent days. In this simple strategy, the long-time average gain per agent per day is $M/(2M + 1)$, as each agent is equally likely to end up being in the winning set.

Clearly, this optimization goal makes the game trivial. A more reasonable goal for selfish agents would be to try to optimize their personal long-term average pay-off. Selfish agents in the majority restaurant would not be interested in pursuing the strategy outlined above. But the game with selfish agents who aim to maximize personal long-term average pay-off is also easily solved. If we allow agents in the state with $(M, M + 1)$ decomposition to shift with a small probability ϵ , then in the long-time steady state, the system jumps between different decompositions of the type $(M, M + 1)$. When ϵ is small, in the steady state, the state $(M, M + 1)$ still occurs with a large weight, and this weight will tend to 1 as ϵ tends to zero. However, for any non-zero value of ϵ , for times $T \gg 1/\epsilon$, the fraction of time spent in the minority by each agent will be nearly the same. Thus, for any specified history, this strategy still gives expected future long-term pay-off that tends to the highest possible, as ϵ tends to zero.

The problem with the the above game is that, for small ϵ , an agent in the unhappy situation in the $(M, M + 1)$ breakup may have to wait very long before his pay-off changes. This suggests that a reasonable model of the agent behaviour would be that he/she does not want to wait for too long for the next winning day. We therefore consider agents who have a finite future time-horizon. Clearly, we can think of agents who try to maximize their net pay-off over the next H days.

It is more convenient to introduce a real parameter λ , lying between 0 and 1, and assume that any agent X only wants to optimize her weighted expected future pay-off,

$$\text{ExpPayoff}(X) = \sum_{\tau=0}^{\infty} [(1 - \lambda)\lambda^{\tau}] \langle W_X(\tau + 1) \rangle \quad (2)$$

where $\langle W_X(\tau) \rangle$ is the expected pay-off of the agent X on the τ -th day ahead, and λ is a parameter $0 \leq \lambda < 1$. It is called as the discount parameter in the literature and is easy to deal with than the discrete parameter H . The factor

$(1 - \lambda)$ has been introduced in the definition of $\text{ExpPayoff}(X)$ so that the maximum possible value of the expected payoff is 1.

The advantage of using the real parameter λ , instead of the discrete parameter H , is that we can study changes in the steady state of the system as we change the parameter λ continuously. We will find that there are strategy switches in the optimal strategies of agents as λ is varied, which leads to discontinuous changes in several properties of the steady state. These are of interest, as they are analogous to dynamical phase transitions in steady states in non-equilibrium statistical mechanics.

Consistent with our assumption of homogeneity of agents, we assume that all agents use the same value of λ .

D. Common information

In the CZMG, on each day, which restaurant was the minority is announced publicly, and this information is available to all the agents. In CAMG, we assume that the common information is more detailed: each agent knows the time-series $\{n(t)\}$ of how many people went to A on different days in the past. We note that in the El Farol bar problem from which MG was motivated has the same information as in our variation. On any day, the attendance in restaurant A can take $(N + 1)$ possible values. Then history of m previous days can take $(N + 1)^m$ values, and each strategy is specified by $(N + 1)^m$ real numbers.

We restrict our discussion here to the simplest case, where $m = 1$ for all agents. Then, an agent's strategy is solely determined by the number of people who were in the restaurant she went to the previous day. Since this number cannot be zero, the number of possible histories here is N , not $(N + 1)$.

IV. THE PROBLEM OF TRAPPING STATES

This model was first defined in [12]. In that paper, we tried to determine the optimal choice of the shift probabilities using the standard ideas of Nash equilibrium. However, we realized that in this problem, there are special states such that none of the agents in that state would prefer to shift to a different restaurant the next day, following a Nash-like analysis. This frozen steady state may be called the trapping state.

The existence of such a trapping state is paradoxical, as rational agents in the majority restaurant have no reason to pursue a strategy that makes them stay in a losing position for ever. The resolution of this paradox requires a new solution concept, that we discuss now.

The most commonly used notion in deciding optimal strategies in N -person games is that of Nash equilibrium: A state of the system in which agent i uses a strategy \mathcal{S}_i is a Nash equilibrium, if for all i , \mathcal{S}_i is the best response of i , assuming that all agents $j \neq i$ use the strategy \mathcal{S}_j . There may be more than one Nash equilibria in a given game, and they may not be very efficient. For CAMG also, the Nash equilibrium is not very satisfactory: it gives rise to a trapping state.

Consider, for simplicity, the case $\lambda = 0$, where agents optimize only next day's pay-off. Now, during the evolution of the game, at some time or the other, the system will reach a state with M agents in one restaurant (assume A), and $M + 1$ agents in the other restaurant B. What is the best strategy of these agents who want to maximize their expected pay-off for the next day?

We imagine that each agent hires a consultant to advise them. To an agent in A (we will call her Alice), the advise would be to stay put, if the probability that no person switches from the restaurant B is greater than $1/2$. If the agents in the restaurant B switch with probability p_{M+1} , the probability that no one switches is $(1 - p_{M+1})^{M+1}$. In this case, the expected pay-off of Alice would be $(1 - p_{M+1})^{M+1}$. So long as this p_{M+1} is small enough that this pay-off is $> 1/2$, Alice's best strategy would be to choose $p_M = 0$.

If an agent in the restaurant B (let us call him Bob) expects that agents in A would not switch, what is his best response? The consultant argues that if Bob switches, he would be in the majority, and his pay-off would be zero. Hence his best strategy is to set his switching probability p_{M+1} to zero. Then, there is some possibility that he will be in the winning set the next day, if some other agent from B shifts. In fact, with agents in A staying put ($p_M = 0$), the probability that he wins is a linear function of $(1 - p_{M+1})$, and is maximized for $p_{M+1} = 0$.

This value $p_{M+1} = p_M = 0$, is then a self-consistent choice corresponding to the the fact that the choice $p_M = 0$ is an individual agent's best response to opposite restaurant's people choosing $p_{M+1} = 0$, and vice versa. It is a Nash equilibrium.

This advice is given to all agents in restaurant B, and then no one shifts, and the situation next day is the same as before. Thus, the system gets trapped into a state where all agents stick to their previous day's choice. In this state, the total number of happy people is the best possible, and the state has the best possible global efficiency.

However, this situation is very unsatisfactory for *the majority of agents* (they are on the losing side for all future days). Setting p_{M+1} equal to zero by agents in B, is clearly not an optimal choice.

Let us denote the state of an agent who is in a restaurant with total of i people in it as \mathcal{C}_i . In the Nash equilibrium concept, an agent in the state \mathcal{C}_{M+1} , who believes that agents in the opposite restaurant would be setting their switch probability $p_M = 0$, is advised that his best response is to set $p_{M+1} = 0$. If other agents are using $p_{M+1} = 0$, one cannot do better by moving. If other agents in the restaurant switch with probability $p_{M+1} \neq 0$, this is the ‘optimal’ solution. *This does not take into account the fact that if all agents follow this advice, their expected future gain is zero*, which is clearly unsatisfactory: No other advice could do worse!

In our previous paper [12], we realized this problem, but could see no way out within the Nash solution concept. We adopted an ad hoc solution, where the agents in state \mathcal{C}_{M+1} were required by external fiat to shift with a non-zero probability ϵ . The alternate solution concept of co-action equilibrium provides a natural solution to this problem. This is explained in the next section.

V. AN ALTERNATE SOLUTION CONCEPT: CO-ACTION EQUILIBRIUM

The problem with the consultant’s reasoning lies in the Nash-analysis assumption of optimizing over strategies of one agent, *assuming that other agents would do as before*. Let the marked agent be denoted by X . All agents in the same restaurant, who are not X denoted by X' . Then, the agent X determines his jump probability p_X to optimize his expected payoff $\text{ExpPayoff}(X) = (1 - p(X))(1 - \prod_{X'}(1 - p(X')))$. In this case, by varying with respect to $p(X)$, keeping all $p(X')$ constant, the payoff is clearly maximized at $p(X) = 0$.

In the alternate co-action equilibrium concept proposed here, an agent in state \mathcal{C}_i realizes that she can choose her switching probability p_i , but all the other fully rational $(i - 1)$ agents in the same restaurant, with the same information available, would argue similarly, *and choose the same value of p_i* . Determining the optimal value of p_i that maximizes the pay-off of agents in state \mathcal{C}_i does not need communication between the agents.

If $p_M = 0$, then the expected pay-off W_{M+1} of an agent in restaurant B is clearly given by the probability that he does not shift, but at least one of the other agents in his restaurant does. This is easily seen to be $q_{M+1}(1 - q_{M+1}^M)$, where $q_{M+1} = 1 - p_{M+1}$. This is zero for $q_{M+1} = 0$ or 1, and becomes maximum when $q_{M+1} = (M + 1)^{-1/M}$. In particular, q_{M+1} equal to 1 is *no longer the optimal response*.

One may argue that this solution concept is not so different from the usual Nash equilibrium, if one thinks of this as a two-person game each day, where the two persons are the majority and the minority groups, and they select the optimal values of their strategy parameters p_i and p_{N-i} . The important point is that these groupings are temporary, and change with time. For non-zero λ , one cannot think of this game as a series of two-person games.

In our model, the complete symmetry between the agents, and the assumption of their being fully rational, ensures that they will reach the co-action equilibrium.

Note that an agent in B may want to ‘cheat’ by deciding not to shift, assuming that other agents would shift with a nonzero probability. But this is equivalent to setting his strategy parameter $p = 0$. Our assumption of rationality then implies that all other agents in the same situation would do the same.

To understand how the co-action concept leads to a much more efficient state, consider the classical game of Prisoner’s Dilemma. We recall that in this game, there are two prisoners, who have to decide if they should ‘cooperate’ or ‘defect’. Assume that the pay-off to each, if both cooperate is 5, and if both defect, the pay-off is 1. If one cooperates, and the other defects, let the pay-off be 0 for the one who cooperates, and 6 for the defector. In the Nash analysis, A argues that his best choice is to defect if he knows that B is cooperating. Also if B is defecting, it is better for A to defect. Thus, in either case, he is better off defecting. And B argues similarly. The game has a stable equilibrium which is that both defect. Of course, they could have obtained much better pay-off, if they both cooperate.

In the co-action concept, A will argue that B will decide similarly as her. Then, it is clear that both cooperating is a better choice than both defecting. The Prisoners’ Dilemma has been discussed for so long, as the argument that the best strategy for both prisoners is to defect seems so unconvincing, and contrary to common sense. In the co-action equilibrium concept, the paradox disappears [15].

VI. DETERMINING THE OPTIMAL MIXED STRATEGY

For a given N , a person’s full strategy \mathbb{P} is defined by the set of N numbers $\mathbb{P} \equiv \{p_1, p_2, \dots, p_N\}$. In CAMG, all rational agents would end up selecting the same optimal values of strategy parameters $\{p_1^*, p_2^*, \dots\}$. It would have been very inefficient for all agents to use the same strategy, if they were using deterministic rules. This is not so in CAMG. We now discuss the equilibrium choice $\{p_1^*, p_2^*, \dots, p_N^*\}$. The co-action equilibrium condition that p_i^* is chosen

to maximize the expected pay-off of agent in state C_i , implies N conditions on the N parameters $\{p_i^*\}$. There can be more than one self-consistent solution to the equations, and each solution corresponds to a possible steady state.

Clearly, as all agents in the restaurant with i agents switch independently with probability p_i , the system undergoes a Markovian evolution, described by a master equation. As each agent can be in one of the two states, the state space of the Markov chain is 2^N dimensional. However, we use the symmetry under permutation of agents to reduce the Markov transition matrix to $N \times N$ dimensional. Let $|Prob(t)\rangle$ be an N -dimensional vector, whose j -th element is $Prob_j(t)$, the probability that a marked agent X finds herself in the state C_j on the t -th day. On the next day, each agent will switch according to the probabilities given by \mathbb{P} , and we get

$$|Prob(t+1)\rangle = \mathbb{T}|Prob(t)\rangle \quad (3)$$

where \mathbb{T} is the $N \times N$ Markov transition matrix.

Explicit matrix elements are easy to write down. For example, \mathbb{T}_{11} is the conditional probability that the marked agent will be in state C_1 on the next day, given that she is in C_1 today. This is the sum of two terms: one corresponding to everybody staying with the current choice [the probability of this is $(1-p_1)(1-p_{N-1})^{N-1}$], and another corresponding to all agents switching their choice [the probability is $p_1 p_{N-1}^{N-1}$].

The total expected pay-off of X , given that she is in the state C_j at time $t=0$ is easily seen to be

$$W_j = (1-\lambda) \left\langle L \left| \frac{\mathbb{T}}{1-\lambda\mathbb{T}} \right| j \right\rangle \quad (4)$$

where $|j\rangle$ is the column vector with only the j -th element 1, and rest zero; and $\langle L|$ is the left-vector $\langle 1, 1, 1, 1, \dots, 0, 0, \dots |$, with first $M = (N-1)/2$ elements 1 and rest zero.

In fact, we can use the permutation symmetry between the agents to block-diagonalize the matrix \mathbb{T} into a two blocks of size $(M+1)$ and M . This is achieved by a change of basis, from vectors $|i\rangle$ and $|N-i\rangle$ to the basis vectors $|s_i\rangle$ and $|a_i\rangle$, where

$$\begin{aligned} |s_i\rangle &= |i\rangle + |N-i\rangle \\ |a_i\rangle &= (N-i)|i\rangle - i|N-i\rangle. \end{aligned} \quad (5)$$

This choice is suggested by the fact that in the steady state

$$Prob(\text{being in } |i\rangle)/i = Prob(\text{being in } |N-i\rangle)/(N-i)$$

It is easily verified that using the basis vectors $|s_i\rangle$ and $|a_i\rangle$, the matrix \mathbb{T} is block-diagonalized.

One simple choice is that $p_i^* = 1/2$ for all i , which is the random choice strategy, where each agent just picks a restaurant totally randomly each day, independent of history. We will denote this strategy by \mathbb{P}_{rand} . In the corresponding steady state, it is easy to see that W_j is independent of j , and is given by

$$W_j = W_{rand} = 1/2 - \binom{N-1}{M} 2^{-N}, \quad \text{for all } j. \quad (6)$$

By the symmetry of the problem, it is clear that $p_N^* = 1/2$ for all λ . Now consider the strategy $\{p_i^*\} = \{p_1^*, 1/2, 1/2, 1/2, \dots\}$. If X is in the state C_1 , and next day all other agents would switch with probability $1/2$, it does not matter if X switches or not: payoffs W_1 and W_{N-1} are independent of p_1^* . Hence p_1^* can be chosen to be of any value. It is easy to see that the strategy \mathbb{P}'_{rand} in which $p_1^* = 0$ and $p_{N-1}^* < 1/2$, chosen to maximize W_{N-1} , is better for all agents and is stable, and hence is always preferred over \mathbb{P}_{rand} .

VII. EXACT SOLUTION FOR SMALL N

A. $N = 3$

We consider first the simplest case $N = 3$. Since $p_1^* = 0$, $p_3^* = 1/2$, the only free parameter is p_2^* . The value of p_2^* is decided by the agents in state C_2 , and they do it by maximizing W_2 .

In this case, the transfer matrix is easily seen to be

$$\mathbb{T} = \begin{bmatrix} q_2^2 & p_2 q_2 & 1/4 \\ 2p_2 q_2 & q_2 & 1/2 \\ p_2^2 & p_2^2 & 1/4 \end{bmatrix} \quad (7)$$

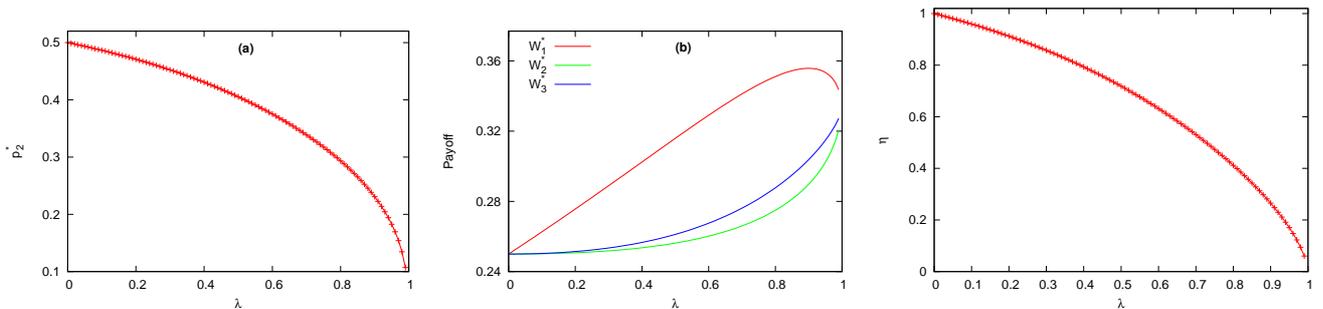


FIG. 2: $N = 3$: (a) Variation of p_2^* with λ , (b) The optimum payoffs W_i^* , ($i = 1$ to 3), as functions of λ and (c) Inefficiency η as a function of λ .

where $q_2 = 1 - p_2$. The pay-off W_2 is given by

$$W_2 = (1 - \lambda) [1 \ 0 \ 0] \frac{\mathbb{T}}{(1 - \lambda \mathbb{T})} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{4p_2q_2 - \lambda p_2(q_2 - p_2)}{(1 - \lambda q_2(q_2 - p_2))(4 + \lambda(4p_2^2 - 1))} \quad (8)$$

The eigenvalues of the transfer matrix \mathbb{T} are easily seen to be $\left(1, \frac{1}{4}(1 - 4p_2^2), q_2(q_2 - p_2)\right)$. The eigen vectors are easily written down. The average gain in the steady state W_{avg} is seen to be

$$W_{avg} = \frac{1}{3 + 4p_2^2} \quad (9)$$

From Eq. 8, the value of p_2 that maximizes W_2 is easily seen to be root of the following cubic equation in λ .

$$16 - 32p_2^* - (24 - 56p_2^* + 32p_2^{*2})\lambda + (9 - 28p_2^* + 40p_2^{*2} - 96p_2^{*3} + 144p_2^{*4} - 64p_2^{*5})\lambda^2 - (1 - 4p_2^* + 8p_2^{*2} - 24p_2^{*3} + 48p_2^{*4} - 32p_2^{*5})\lambda^3 = 0$$

The variation of p_2^* with λ is shown in Fig 2a. p_2^* monotonically decreases with λ from its value $1/2$ at $\lambda = 0$, and tends to 0 as λ tends to 1. The pay-off of agents in various states with this optimum strategy is shown in Fig. 2b and the variation of the inefficiency parameter η with λ is shown in Fig. 2c.

It is easily seen that W_{avg} is a monotonically increasing function of λ , and tends to the maximum possible value $W_{max} = 1/3$ as $\lambda \rightarrow 1$. The variation of the inefficiency parameter η with λ is shown in Fig. 2c.

B. $N = 5$

We can similarly determine the optimal strategy for $N = 5$. This is characterized by the five parameters $(p_1^*, p_2^*, p_3^*, p_4^*, p_5^*)$. The simplest strategy is \mathbb{P}_{rand} , which corresponds to $p_i^* = 1/2$, for all i . As explained above, the strategy $\mathbb{P}'_{rand} = (0, 1/2, 1/2, p_4^*(\lambda), 1/2)$, gives higher pay-off than \mathbb{P}_{rand} for all agents, for all λ .

Now consider agents in the states C_2 and C_3 . What values of p_2 and p_3 they would select, given their expectation/belief about the selected values of p_1, p_4 and p_5 ? We can determine these by analyzing the variation of payoffs W_2 and W_3 as functions of p_2 and p_3 for fixed values of p_1, p_4, p_5 and λ as discussed below.

Let us denote the best response of agents in state C_2 (that maximizes W_2), if the agents in the opposite restaurant jump with probability p_3 by $r_2^{opt}(p_3)$. Similarly, $r_3^{opt}(p_2)$ denotes the best response of agents in state C_3 , when those in the opposite restaurant jump with probability p_2 .

In Fig. 3, we plot the functions $r_2^{opt}(p_3)$ (OAP) and $r_3^{opt}(p_2)$ (BP), in the (p_3, p_2) plane, for three representative values of λ . For small p_3 , $r_2^{opt}(p_3)$ remains zero, and its graph sticks to x-axis initially, (segment OA in figure), and then increases monotonically with p_3 . The strategy \mathbb{P}'_{rand} is the point $(1/2, 1/2)$, denoted by P. We also show the lines PC corresponding to $W_3 = W'$, and PD, corresponding to $W_2 = W'$, where W' is the expected gain of agents in

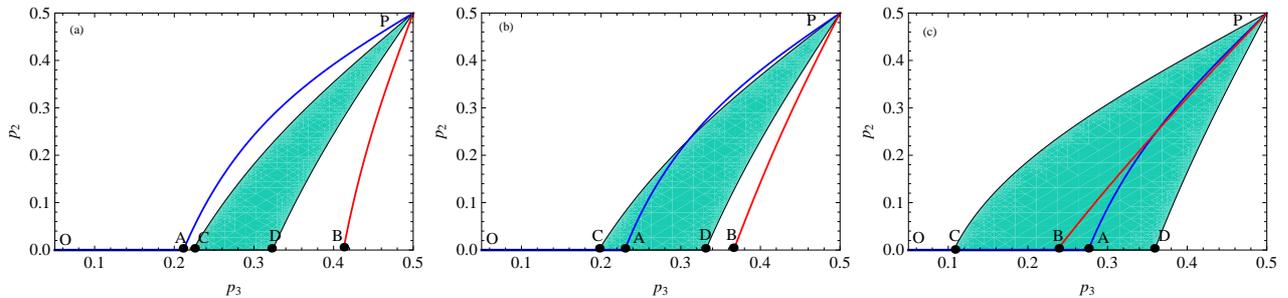


FIG. 3: Region in the p_2 - p_3 plane showing the best responses $r_2^{opt}(p_3)$ (blue) and $r_3^{opt}(p_2)$ (red) for agents in state $|2\rangle$ and $|3\rangle$ respectively, for (a) $\lambda = .1$, (b) $\lambda = .4$ and (c) $\lambda = .8$. The line PC and PD show the curves $w_3 = W'$ and $W_2 = W'$ respectively. In the curvilinear triangle PCD , all agents do at least as well as at P .

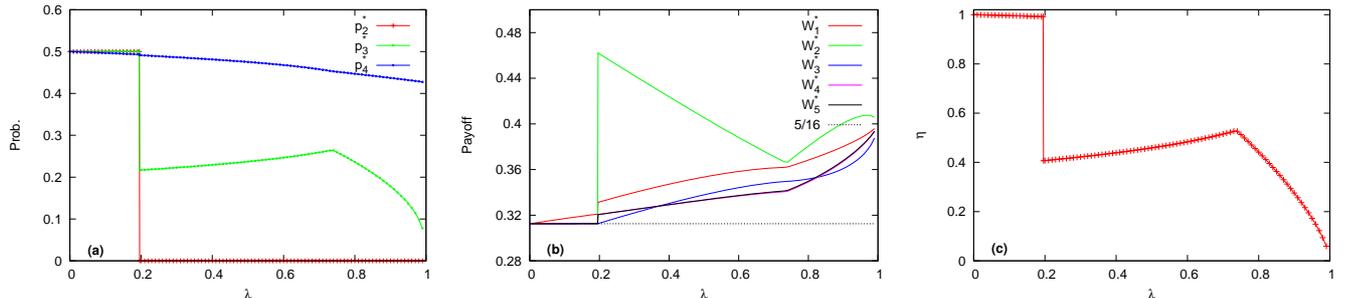


FIG. 4: $N = 5$: (a) Variation of p_2^* , p_3^* and p_4^* with λ , (b) Optimum payoffs as functions of λ , (c) Inefficiency η as a function of λ .

state C_2 or C_3 under \mathbb{P}'_{rand} . For all points in the curvilinear triangle PCD , both W_2 and $W_3 \geq W'$. Clearly, possible equilibrium points are the points lying on the lines $r_2^{opt}(p_3)$, or $r_3^{opt}(p_2)$ that lie within the curvilinear triangle PCD . However, along the blue curve OAP representing $r_2^{opt}(p_3)$, maximum value for W_2 is achieved when $p_2 = 0$. Therefore we can restrict the discussion of possible equilibrium points to the line segment CD in Fig. 3.

For small λ (shown in Fig. 3a for $\lambda = 0.1$), The point A is to the left of C , and the only possible self-consistent equilibrium point is P . For example, if agents in the state C_3 (Bob) assumes that agents in the minority restaurant (Alice) is going to set $p_2^* = 0$, Bob can get better pay-off than \mathbb{P}'_{rand} , by choosing his probability parameter p_3^* in the range CD in Fig 3a. But a rational Alice would not choose $p_2^* = 0$, if she expects p_3^* to be in the range CD . Similar argument rules out all points in the colored curvilinear triangle PCD as unstable. This implies that the agents would choose $p_2^* = p_3^* = 1/2$. This situation continues for all $\lambda < \lambda_{c1} = 0.195 \pm 0.001$.

For $\lambda > \lambda_{c1}$, the point A is to the right of C . This is shown in Fig. 3b, for $\lambda = 0.4$. In this case, possible equilibrium points lie on the line-segment CA , and out of these, A will be chosen by agents in state C_3 . At A , both W_2 and W_3 are greater than W' , and hence this would be preferred by all. Further optimization of p_4 changes p_3 and p_4 only slightly.

As we increase λ further, for $\lambda > \lambda_{c2}$ [numerically, $\lambda_{c2} = 0.737 \pm 0.001$], the point B comes to the left of A . Out of possible equilibria lying on the line-segment CA , the point preferred by agents in state C_3 is no longer A , but B . The self-consistent values of p_2^* , p_3^* , and p_4^* satisfying these conditions and the corresponding payoffs are shown in Fig. 4a and Fig. 4b respectively.

In Fig. 4c, we have plotted the inefficiency parameter η as a function of λ . For $\lambda < \lambda_{c1}$, there are possible values of p_2^* and p_3^* , that would increase the expected pay-off for everybody. However, Alice and Bob can not be sure that the other party would not take advantage of them, and hence stick to the default sub-optimal-for-both choice $p_2^* = p_3^* = 1/2$.

Also, in the range $\lambda_{c1} < \lambda < \lambda_{c2}$, the inefficiency rises as the agents optimize for farther into the future. This may appear paradoxical at first, as certainly, the agents could have used strategies corresponding to lower λ . This happens because though the game for larger λ is slightly less efficient overall, in it the majority benefits more, as the difference between the optimum payoffs W_2^* and W_3^* is decreased substantially (Fig. 4b).

We note that the optimal strategies, and hence the (non-equilibrium) steady state of the system shows a non-analytic dependence on λ , even for finite N . This is in contrast to the case of systems in thermal equilibrium, where

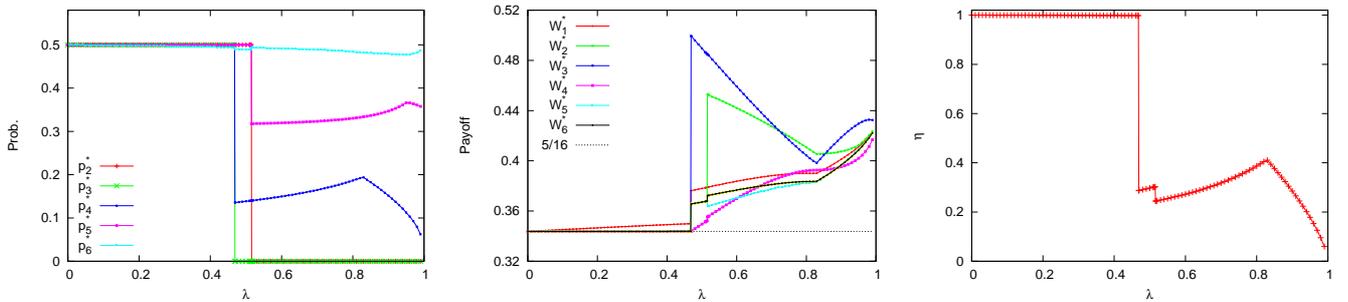


FIG. 5: $N = 7$: (a) Variation of optimum switch probabilities with λ , (b) Optimum payoffs as functions of λ , (c) Inefficiency η as a function of λ .

mathematically sharp phase transitions can occur only in the limit of infinite number of degrees of freedom N . This may be understood by noting that the fully optimizing agents in CAMG make it more like an equilibrium system at zero-temperature. However note that unlike the latter, here the system shows a lot of fluctuations in the steady state.

C. Higher N

For higher values of N , the analysis is similar. For the case $N = 7$, we find that there are four thresholds λ_{ci} , with $i = 1$ to 4. For $\lambda < \lambda_{c1}$, the optimal strategy has the form $(0, 1/2, 1/2, 1/2, 1/2, p_6^*, 1/2)$. For $\lambda_{c1} \leq \lambda \leq \lambda_{c2}$, we get $p_3^* = 0$, and $p_4^* < 1/2$. For still higher values $\lambda_{c2} < \lambda \leq \lambda_{c3}$, agents in the states C_2 and C_5 also find it better to switch to a win-stay lose-shift strategy, and we get $p_2^* = 0$, $p_5^* < 1/2$. The transitions at λ_{c3} and λ_{c4} are similar to the second transition for $N = 5$, in the (p_4, p_3) and (p_5, p_2) planes respectively. Numerically, we find $\lambda_{c1} \approx 0.47$, $\lambda_{c2} \approx 0.52$, $\lambda_{c3} \approx 0.83$ and $\lambda_{c4} \approx 0.95$. We present some graphs for the solution for $N = 7$. Fig. 5a shows variation of the optimum switch probabilities in various states and Fig. 5b shows the variation of the optimum payoffs. Fig. 5c shows the variation of inefficiency with λ . The general structure of the optimum strategy is thus clear. As λ is increased, it changes from random switching to a complete win-stay lose-shift strategy in stages.

An interesting consequence of the symmetry between the two restaurants is the following: If there is a solution $\{p_i^*\}$ of the self-consistent equations, another solution with all payoffs unchanged can be obtained by choosing for any j , a solution $\{p_i^{*'}\}$, given by $p_j^{*'} = 1 - p_j^*$, and $p_{N-j}^{*'} = 1 - p_{N-j}^*$, and $p_i^{*'} = p_i^*$, for all $i \neq j$ and $i \neq N - j$. How agents choose between these symmetry related 2^M equilibria can only be decided by local conventions. For example, if all agents follow the ‘Win-stay lose-shift’ convention, this would select a unique equilibrium point.

VIII. SUMMARY AND CONCLUDING REMARKS

In this paper, we have analyzed a stochastic variant of the minority game, where the N agents are equal (no quenched randomness in strategies given to agents). This permits an exact solution in terms of N self-consistently determined parameters. The solution shows multiple sharp transitions as a function of the discount parameter λ , even for finite N . The optimal strategy is found to be more efficient than possible in CZMG.

Our treatment of the model here differs from that in [12]: in that paper, the game was discussed only for $\lambda = 0$ (corresponds to agents optimizing only next day’s payoff), and in terms of Nash equilibrium. Within the Nash solution concept, it was not clear how to avoid the problem of trapping states, and we had made an ad hoc assumption that whenever the system reaches a trapping state, a major resetting event occurs where all agents switch restaurant with some largish probability. In the co-action equilibrium concept proposed here, the decision to switch to $p = 1/2$ or not is made rationally by the agents themselves depending upon their future time horizon.

Generalizations of the model for $m > 1$, or when all agents do not use the same value for the discount parameter etc. are easy to define, and appear to be interesting subjects for further study. The technique may be used to study other games with different pay-off functions, e.g. agents win when their restaurant has attendance exactly r .

Acknowledgments

We thank D Challet, K. Damle, P. Grassberger, M. Marsili and J. Stilck for their comments on an earlier version of this paper.

-
- [1] W. Brian Arthur, Amer. Econ. Rev. (Papers and Proc.) **84** 406, 1994.
 - [2] D. Challet and Y. C. Zhang, Physica A, **246** 407 (1997).
 - [3] W. Brian arthur, in the foreword of [4].
 - [4] D. Challet, M. Marsili and Y. C. Zhang, *Minority Games: Interacting agents in financial markets* (Oxford Univ. Press, Oxford, 2005), page 7.
 - [5] J. B. Satinover and D. Sornette, Eur. Phys. J. B **60** 369 (2007).
 - [6] A. Cavagna, A. P. Garrahan, I. Giardina and D. Sherrington, Phys. Rev. Lett. **83** 4429 (1999).
 - [7] For other applications of mean-field methods to game theory, see O. Gueant, J.-M. Lasry and P. -L. Lions, Lecture Notes in Mathematics, Vol. 2003, Eds. A. R. Carmona et al., 205 (Springer, Berlin, 2010).
 - [8] For a discussion of other alternates to Nash Equilibrium, see John. C. Harsanyi and R. Selten, *A general theory of equilibrium selection in games* (MIT Press, Cambridge, 1988).
 - [9] A. C. C. Coolen *The mathematical theory of minority games: Statistical mechanics of interacting agents* (Oxford Univ. Press, Oxford, 2005).
 - [10] E. Moro, in *Advances in Condensed Matter and Statistical Mechanics*, Eds. E. Korutcheva and R. Cuerno (Nova Science Publishers, New York, 2004).
arXiv:cond-mat/0402651v1 [cond-mat.dis-nn].
 - [11] W. Kets, Journal of Economic Surveys, Vol. 26, Issue 5, pp. 865-878, 2012.
 - [12] D. Dhar, V. Sasidevan and Bikas K. Chakrabarti, Physica A, **390** 3477 (2011).
 - [13] G. Reents, R. Metzler and W. Kinzel, Physica A, **299** 253 (2001).
 - [14] M. Nowak and K. Sigmund, Nature, **364** (1993) 56.
 - [15] V. Sasidevan and D. Dhar, in preparation.