

Properties of interaction networks underlying the Minority Game

Ines Caridi*

*Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires, Buenos Aires, Argentina*

(Dated: October 8, 2014)

The Minority Game is a well known agent-based model with no explicit interaction among its agents. However, it is known that they interact through the global magnitudes of the model and through their strategies. In this work we have attempted to formalize the implicit interactions among Minority Game agents as if they were links on a complex network. We have defined the link between two agents by quantifying the similarity between them. This link definition is based on the information of the instance of the game (the set of strategies assigned to each agent at the beginning) without any dynamic information on the game, and brings about a static, unweighed and undirected network. We have analyzed the structure of the resulting network for different parameters, such as the number of agents (N) and the agent's capacity to process information (m), always taking into account games with two strategies per agent. In the region of crowd-effects of the model, the resulting networks structure is a small world network, whereas in the region where the behavior of the Minority Game is the same as in a game of random decisions, networks become a random network of Erdos-Renyi. The transition between these two types of networks is slow, without any peculiar feature of the network in the region of the coordination among agents. Finally, we have studied the resulting static networks for the Full Strategy Minority Game model, a maximal instance of the Minority Game in which all possible agents take part in the game. We have explicitly calculated the degree distribution of the Full Strategy Minority Game network and, on the basis of this analytical result, we have estimated the degree distribution of the Minority Game network, which is in accordance with computational results.

PACS numbers: 89.65.Gh; 82.20.Wt

Keywords: Minority game; complex networks; social networks; Full Strategy Minority Game

I. INTRODUCTION

One of the first attempts to use tools of mathematics to study problems from the social world is originated by the work of J. Von Neumann and O. Morgenstern, who applied the game theory to economical phenomena [1]. Later, works of Axelrod gave an important impulse to the agent-based models [2, 3]. One of the questions followed by Axelrod is to understand if the cooperation can emerge in a system of selfish individuals. Recently, from statistical physics, a lot of works addressed the issues of cooperation under the framework of the social dilemmas [4, 5]. In some situations, individuals compete for resources which are limited, like in traffic problems. The Minority Game falls in this framework of competing agents with limited resources. It was introduced in 1997 by Challet and Zhang as an attempt to pick some essential characteristics of a competitive population, in which an individual achieves the best result when he manages to be in the minority group [6, 7]. They are inspired in El Farol Problem [8]. The Minority Game is an adaptive agent-based model, and presents emergent properties like coordination among agents under certain circumstances.

The original model was formulated as follows: there are N agents that in each t step of the game must choose one of two alternatives (0 or 1, for example). Once

agents have played, it is necessary to count how many of them went to each side, ($N_1(t)$ and $N_0(t)$) so that $N_1(t) + N_0(t) = N$ for all values of t and the winners are those who happen to be in the minority group (the minimum value between $N_1(t)$ and $N_0(t)$). The only information available for the agents is the state of the system μ , that is one of the \mathcal{H} possible patterns labeled for an integer value $\mu = 1, \dots, \mathcal{H}$. Agents play using strategies. A strategy is a function that assigns a prediction (0 or 1) to each of the possible states. In this way, there are $\mathcal{L} = 2^{\mathcal{H}}$ different strategies; the set of \mathcal{L} strategies defines the so-called Full Strategy Space (FSS). At the beginning of the game, each agent randomly chooses s strategies from the FSS. In this paper, we will work with $s = 2$. As strategies are chosen with repetition, it is possible for an agent to have two identical strategies, and for two agents to have the same pair of strategies. In the original formulation, μ is an endogenous variable determined by the sides which turn out to be the minority ones in the last m steps of the game. Therefore, the number of states is $\mathcal{H} = 2^m$. Every agent receives one point every time he manages to be in the minority group. On the other hand, agents also record whether their strategies were good or bad. They assign a virtual point (to distinguish it from the points assigned to the agents) to the strategies which correctly predicted the side that resulted minority for a given step, regardless of the fact that the strategy may or may not have been used. At every step, the agent chooses the best strategy to play (the one that accumulated the most virtual points up to that time). When two or more

*E-mail me at: ines@df.uba.ar

strategies have scored equally well, the agent randomly chooses one of them.

Many works have studied the effect which a definition of the state based on exogenous information has on the behavior of the model [9–11]. For example, Cavagna proposed a new *updating rule* for the state of the system [9] (here denoted MG_{rand}) consisting of choosing one of the \mathcal{H} states at random from a uniform distribution at every step of the game. The behavior of the MG_{rand} essentially turns out to be the same as that of the MG [11, 12]. Moreover, this model has allowed to carry out analytical approaches [13, 14].

An *instance* I of the MG with $s = 2$ is a particular assignment of two strategies to the agents, $I = \{(e_1^1, e_2^1), (e_1^2, e_2^2), \dots\}$. For $i = 1, \dots, N$, the pair (e_1^i, e_2^i) represents the set of strategies assigned to the agent i . We define a *realization* \mathcal{E} of the game as a pair $\mathcal{E} = \{\mathcal{S}_{\mathcal{E}}, I\}$, where $\mathcal{S}_{\mathcal{E}} = \{\tilde{\mu}^1, \tilde{\mu}^2, \dots\}$ is a sequence of states (generated by any updating rule), and I is an instance of the MG . We will follow this notation proposed in previous works ([15] for example).

The most studied variable of the MG is the so called reduced variance $\sigma^2/N = \langle (N_1(t) - N/2)^2/N \rangle_{\mathcal{E}}$, a measure of the population's squandering of resources [7]. It measures the population's waste of resources by averaging – over time and over different realizations \mathcal{E} – the quadratic deviation from $N/2$ of the number of agents that choose a fixed side (for example, N_1). If s is fixed for a game, σ could depend on two parameters, N and m . However, it has been found that the relevant control parameter of the model is $\alpha = \mathcal{H}/N = 2^m/N$ [16]. The magnitude σ^2/N shows scaling as a function of α , and the model presents a phase transition with symmetry breaking. Ref [17] shows that the transition occurs when σ^2/N reaches its lowest value, at α_c , and separates the symmetric ($\alpha < \alpha_c$) from the asymmetric ($\alpha > \alpha_c$) phase. The broken symmetry is the equivalence of the two sides: in the symmetric phase $\langle N_1(t)|\mu \rangle = 0$ for all μ state values, while in the asymmetric phase, $\langle N_1(t)|\mu \rangle \neq 0$ at least for one μ state. This means that the minority side becomes predictable. The behavior of the MG for $\alpha \gg 1$ is equivalent to that of a game in which agents take decisions at random. In this way, $\sigma^2/N \approx 0.25$. In the region where $\alpha \ll 1$, *crowd effects* arise at some steps of the game. These crowds of agents moving together to one of the two sides turn out to be the majority and are, of course, the losers; at this point where crowds emerge, the contribution to σ is very important. Then, crowd effects are the reason why σ is a large number in this region, a fact which shows that fewer resources are allocated to the population as a whole. *Period Two Dynamics* (PTD) in the sequence of the minority sides was first observed within this region by Savit et al. [16]. The dynamics established in this region enables the number of virtual points assigned to the strategies to be reduced to a limited set of cases [15, 18], which facilitates the analytical treatment of the model. One of the reasons why the Minority Game has attracted attention is that in certain region of the parameters, the

reduced variance is smaller than that obtained for a game in which each of the agents randomly chooses between the two sides (see Figure 1), meaning a better use of the resources by the population. But it is interesting to note that the reduced variance reflects that the population as a whole achieves more resources, but does not reveal how that wealth is distributed among the agents. Ho et al. redefined a Gini index for the MG [19] which showed that whenever the reduced variance takes its minimum value, the inequality among the agents is maximized.

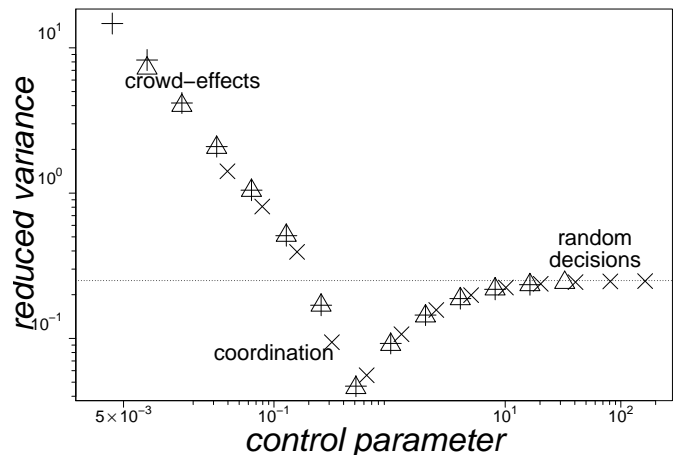


FIG. 1. Reduced variance (σ^2/N) as a function of the control parameter ($\alpha = \mathcal{H}/N = 2^m/N$) for the MG for different values of m (from 2 to 14) and N agents (plus symbols for 1001, triangle for 501 and X for 101 case). For each value of N and m , 100 runs have been performed, each one of $T = 100000$ time steps discarding the first 50000 steps. Dash line corresponds to the case of random decisions. Regions of crowd-effects, coordination and random decisions are highlighted.

In a recent work [15], we studied crowd region and presented a calculation of σ that matched the simulation results. This calculation was based on the FSMG (Full Strategy Minority Game), an imaginary instance of the MG where all the possible agents (let call \mathcal{N}) participate in the game. As a consequence, a lot of symmetries could be exploited to obtain an explicit analytical solution for the magnitude σ . Once the FSMG was solved, we obtained results for σ for the MG (in the region in which PTD is valid) by sampling N agents from the \mathcal{N} agents of the FSMG. Additionally, in a more recent work, we analyzed two well-known properties: quasiperiodicity of the sequence of outcomes of the game and Period Two Dynamics, by studying the sequence of minority sides through a network and by using the FSMG to shed light on these properties [20].

The MG has been studied using very different tools: numerical simulations [6][7][16], a generalization with a temperature-like variable [21], a mapping of the model to a spin glass [14], to mention just a few.

In recent years, many areas of science have become

highly interested in the properties of complex networks of very different complex systems, which include communication, biochemical, ecosystems, internet and social networks, among others [22] [23] [24]. If we think about individuals as nodes and their interactions as links, the social world around us can be represented as a social network.

Different extensions of the MG allow agents to access to information from their neighborhood, by including explicit interactions and the exchange of information among agents [25][26]. For example, in [27], μ_i is defined for the agent i by taking into account the output move of i - agent's neighbors, as if they were boolean agents. Recently, in [28], authors studied a system of boolean agents playing a generalized minority game. In [29], authors studied the evolution of the networked evolutionary minority game (NEMG), where each agent can evolve his strategies taking into account information from his neighbors. In [30], connections between agents are dynamically inserted or removed from the network. In [31], authors applied the crowd-anticrowd theory to the NEMG. In [32], we introduced some degree of local information, which is only available for some agents in an ordered network. In all these works, agents'neighborhood had been provided by a complex network before the game had started (althout it can evolve in some generalizations), partially changing the original rules of the game by introducing explicit interactions. These works studied how different network organizations affect the dynamics of the generalized MG, using ordered, random, small-world and scale-free networks.

In this work, we have proposed to characterize the implicit interactions among the MG agents which are established from the instance of the game (before the game started) on a static complex network. Agents play the original rules of the MG, without the context of the neighborhood or explicit interactions. We have not changed the rules of the game, but represented implicit interactions of the MG as links on a network. We were inspired by two strongly related questions which are usually asked in other complex systems: *What is the underlying complex network that connects individuals in this system like?* and, *Is the structure of this network related to different properties of the system behavior?*

We have defined a link between two agents by quantifying the similarity between their set of strategies. In section II, we present the chosen definition for the *link*, and the properties of the obtained networks. In section III, we will analyze the network of the FSMG, given the same link definition for the MG, and we will analytically calculate the degree distribution of these networks as a function of m . In the same section, we will describe how we have estimated the degree distribution of the MG networks from the analytical results for those of the FSMG. In section IV we will present our conclusions.

II. THE UNDERLYING NETWORK OF THE MINORITY GAME

Agents will be represented by nodes and the link between a pair of nodes is established whenever certain condition is given for both sets of the agents' strategies.

Before introducing the definition of the link, we need to define the Hamming distance between two strategies: *the Hamming distance between a pair of strategies e_1 and e_2 , $d(e_1, e_2)$, is the number of bits in which they differ, normalized by the length of the strategy (measured by the number of bits)*. For example, the Hamming distance between the two strategies $e_1 = (1, 1, 0, 0)$ and $e_2 = (1, 1, 1, 1)$ is $d(e_1, e_2) = 1/2$. Let us note that previous strategies correspond to a game with $m = 2$ and therefore there are $\mathcal{H} = 4$ states, that is why each strategy assigns four predictions, one for each possible state.

For games with two strategies per agent, we will establish a link $L_{ij} = 1$ between agents i and j , whose strategies are (e_1^i, e_2^i) and (e_1^j, e_2^j) , if the condition

$$d(e_k^i, e_l^j) < \frac{1}{2} \quad (1)$$

is met for all k and l values, such as $k=1,2$ and $l=1,2$. This means that there will be a link between two agents if, whichever strategy each agent chooses to play, the Hamming distance between those strategies is less than $1/2$. Let us note that this definition does not use any dynamic information, but the complete set of available strategies for the agents. In a context of *ergodicity* of the game (each state occurs with equal probability), the following is valid: $L_{ij} = 1$ if and only if agents i and j play the same side more than one half of the time steps. In a context of non-ergodicity, the previously-mentioned statement does not apply. $L_{ij} = 1$ does not imply anything about the moves of agents i and j , because they could be *frozen* agents who use only one of their strategies, and the system could concentrate on only a set of states. Conversely, two agents l and k could play the same side more than one half of the time steps and yet, L_{ij} may still be equal to 0.

In [20] we showed that a MG with a deterministic rule to play in case of a tie of the strategies for which the SPTD (Strict Period Two Dynamics, the *PTD* with probability equal to 1) for even occurrences of the states is met, is a periodic game. For example, it applies to the MG^{prior} where agents have an *a priori* favorite strategy to use in case of tie. We proved that if the MG is periodic and meets the SPTD for the even occurrences of the states, then all the states appear in the period the same even number of times, which implies the ergodicity of the game (see theorem 4 in section 4 of [20]). Then, $L_{ij} = 1$ ensures that agents i and j play the same side more than one half of the time steps. Moreover the period length P can be written as $P = 2n\mathcal{H}$, being n an integer number ≥ 1 . This deterministic version of the MG leaves the same behavior of the MG, a fact which is

evidenced in simulations of σ^2/N , which show the same curve as in the MG case. The only difference appears in the fluctuations of the reduced variance of the MG^{prior} , which are bigger than in the MG case.

This definition of link leads to an unweighed and undirected network. To find the network of connections for a particular game, we have generated an instance of the game (an allocation of strategies for both values of N and m), and looked at the resulting K connections complying with a previous definition of the link (1). These two sets, the nodes (set of agents) $\{n_1, n_2, \dots, n_N\}$ and the connections $\{l_1, l_2, \dots, l_K\}$, whose sizes are N and K respectively, define the associated network of the MG , which we note $G^{MG}(N, K(N, m))$, which depends both on the definition of the link, and on the instance of the game. We will study the properties of G^{MG} for different parameters of m (ranging from 2 to 14), and N (101, 501, 1001 and 5001), including the degree distribution, the degree correlation, the clustering coefficient, the average minimum path and some aspects of the clusters structure.

A. Degree distribution and degree correlation

Figure 1 shows the probability of finding links on the network $c = K/(N(N-1)) = \langle k \rangle / (N-1)$ where $\langle k \rangle = K/N$ is the mean degree of the network. Points for different values of N and the same value of m overlap. c grows with m up to a stabilization around a value which is close to but less than $1/16$ (in section III, we will return to this value). Standard deviation of c is shown in the inset plot of the figure on a different scale. The standard deviation becomes smaller while N increases; for the case of $N = 5001$ the error bar is smaller than the size of the circle symbol.

Figure 2 shows the degree distribution of G^{MG} for different values of m ($m = 3, 5, 8$ and 11) and the same value of $N = 1001$; the continuous curve in the Figure corresponds to the theoretical degree distribution for an Erdos-Renyi network, G^{ER} , with N nodes and mean degree $\langle k \rangle_I$, which is the value obtained by averaging the mean degree of the G^{MG} over 100 instances. For greater values of m , connections behave like on a random network, as reflected by the degree distribution for the case of $m = 11$ for $N = 1001$ which matches those of an Erdos-Renyi network. Hence, from $m = 9$ to $m = 14$, using the Pearson's Chi-square test, we cannot reject the hypothesis that *the observed degree distribution of G^{MG} is normally distributed with equal mean and variance* (see caption of Figure 2 for details). On the other hand, for small values of m , the obtained degree distributions are very different from those of a random network, as can be seen in the case of $m = 3$, where the histogram shows different peaks. This multimodal degree distribution could be understood using the FSMG in Section III.

Figure 3 shows the degree correlation as a function of m for different values of N . We have noticed that for small values of m , networks turn out to be dissortative,

while as m grows, the degree correlation of the network increases until there is no correlation.

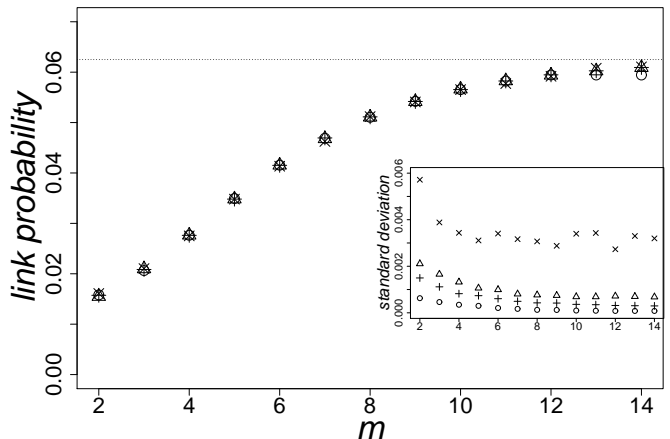


FIG. 2. Probability of finding links on the MG networks, $\langle k \rangle / (N-1)$, vs m for different values of N (circle symbols for $N = 5001$ case, plus for 1001, triangle for 501 and X for 101 case). Each point correspond to an average value over 100 instances. Inset of the figure show the standard deviations for the values of the probability of finding a link by using the same symbols as in the main figure. The case of $N = 5001$ presents an error bar which is lower than the symbol size of main figure.

We understand a *cluster* as the set of nodes V of the network for which for each pair of them $v, u \in V$, $u \neq v$, there exists a directed path from v to u throughout the network. For values of m greater than 4, for values of $N = 5001, 1001$ and 501 , and $m = 8$ for $N = 101$, G^{MG} turn out to be connected networks (with only one cluster).

B. Clustering and minimum mean path

We have also calculated the clustering coefficient and the minimum average path for G^{MG} . The clustering coefficient takes into account the biggest cluster of the network if it is not connected. We have studied the clustering in terms of m and N and we have found that MG networks for various values of N show roughly the same clustering in terms of m , as shown in black symbols of Figure 5. For example, for the case of $m = 6$, clustering coefficient of G^{MG} for different values of $N = 101, 501, 1001, 5001$ are very similar, as it is shown in the inset figure (error bars correspond to the standard deviation which is lower as N grows), and using a χ^2 statistical test, we cannot reject the null hypothesis H_0 : “the values of clustering coefficient obtained from G^{MG} for $N = 101, 501, 1001$ and 5001 are mutually consistent (i.e., correspond to variables with the same mean value)” for any value of m from 2 to 14 (see details in the cap-

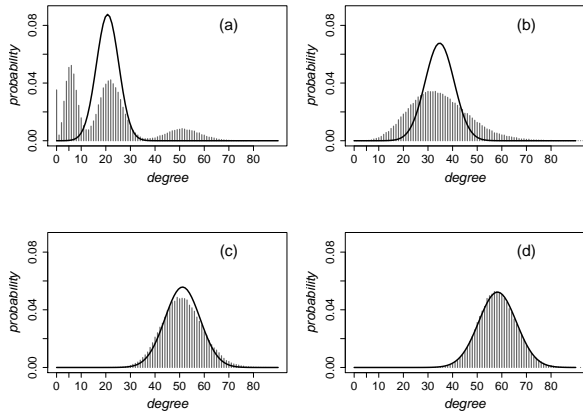


FIG. 3. In gray, degree distribution of MG networks for (a): $m = 3$, (b): $m = 5$, (c): $m = 8$ and (d): $m = 11$ for $N = 1001$, averaged over 100 different instances. The continuous black curves correspond to the degree distributions for equivalent Erdos Renyi networks, where we approximate the Poisson distribution with a Normal distribution $N(\mu, \sigma)$ with the parameters $\mu = \langle k \rangle$ and $\sigma = \sqrt{\langle k \rangle}$, where $\langle k \rangle$ is the value obtained by averaging the mean degree of the G^{MG} over 100 instances. We performed a Pearson's Chi-square statistical test in order to check the null hypothesis H_0 : "data observed of degree distribution of G^{MG} are distributed like in random networks, i.e. a normal distribution with equal mean and variance". We computed the statistical value X^2 as the weighed sum of the squared deviations from the number of events observed and expected in each class of the histogram, i.e., $X^2 = \sum_{i=1}^l (n_i - np_{oi})^2 / np_{oi}$ where n_i is the number of observed events in class i and np_{oi} is the predicted number of events in the class i , being n the total number of events, and p_{oi} the predicted probability of class i for the underlying distribution. In our case, we are testing that this is a Normal distribution with equal mean and variance μ which has been estimated from the same set of data. When H_0 is true, X^2 is approximately a χ_{l-2}^2 Chi-square distribution with $l-2$ degrees of freedom (because of the constraint of the normalization condition of p_{oi} and the parameter μ estimated from data) [33]. We can reject H_0 for $m = 8$ case with significance level smaller than 0.1%, in which case the normalized statistics ($X_n^2 = X^2 / (l-2)$ which is expected value 1) is: $X_n^2 = 2.43$ and $l = 67$. For $m = 9, \dots, 14$ we cannot reject the hypothesis H_0 because the values obtained for the normalized statistics are: $X_n^2 = 0.38$ and $l = 63$ for $m = 9$, $X_n^2 = 0.15$ and $l = 61$ for $m = 10$, $X_n^2 = 0.135$ and $l = 61$ for $m = 11$, $X_n^2 = 0.127$ and $l = 64$ for $m = 12$, $X_n^2 = 0.132$ and $l = 63$ for $m = 13$, finally $X_n^2 = 0.128$ and $l = 62$ for $m = 14$ case.

tion of the figure 5). On the other hand MG clustering is greater than the corresponding value for random networks for values of m up to $m = 9-10$. In Figure 5, gray symbols represent values of link probability previously reported in order to appreciate that for greater values of m , the clustering coefficient is very similar to them (although a little greater), as occurs in random networks where there is no correlation. The dashed line on the figure corresponds to the value $1/16$.

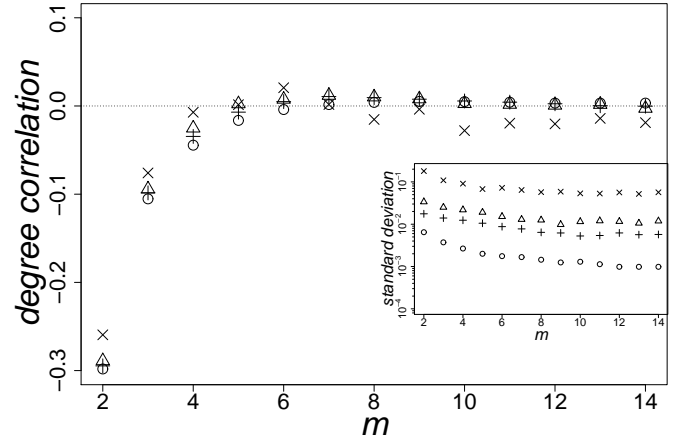


FIG. 4. Degree correlation of G^{MG} as a function of m for different values of N (circle symbols for $N = 5001$ case, plus for 1001, triangle for 501 and X for 101 case). Each point corresponds to an average value over 100 instances. Inset shows values of standard deviation of the degree correlation in a log scale by using the same symbols.

As observed in real networks, high clustering is typical of networks where the link represents a social relationship, such as in friendship networks (it is very likely somebody's friends are also friends to one another). In the case of MG, we have seen this feature for small values of m , more precisely, in the region where crowd effects emerge, and make the system inefficient in the use of resources. As m grows, although the number of connections increases, clustering reflects that these connections are allocated without transitivity; in other words, the probability that two neighbors of a node are connected to each other is the same as the probability of finding a link on a random network, where there is no correlation.

We have calculated the average shortest path length of G^{MG} and compared it with the corresponding values for G^{ER} , for all the values of N and m we have worked with. Figure 6 shows that the minimum average path of G^{MG} coincides with that of G^{ER} for all values of m greater than $m = 5$, and it is close to this value for lower values of m . Given the values of clustering, minimum average path, degree correlation and the degree distribution, we can say that for small values of m (when crowds emerge), the MG network is a small world network, while for greater values of m , G^{MG} behaves as a random network. It is remarkable that the instance of the game has sufficient information to define which of these two cases the game will fall in. The transition between these two types of networks is slow, without any peculiar feature of the network in the region of the coordination among agents. Perhaps the instance (i.e. static) network may not be sufficient to offer information about this region, and it might be necessary to consider the dynamic net-

work, which should be obtained using information related to the strategies which the agents actually used.

In order to understand aspects of the G^{MG} , we will analyze the network of a maximal instance of the MG, the Full Strategy Minority Game, in the next section.

III. THE UNDERLYING NETWORK OF THE FULL STRATEGY MINORITY GAME

The Full Strategy Minority Game (FSMG) model was introduced in [15, 18] as an instance of the MG where its \mathcal{N} agents were chosen in a particular way: they are all possible different agents that could exist by combining strategies from the set of $\mathcal{L} = 2^{\mathcal{H}}$ strategies of the FSS. Thus, for a game with $s = 2$ strategies per agent, there are as many agents as possible pairs of strategies (with repetition) from the FSS. Therefore, for $s = 2$, the number of players will depend on m in the form:

$$\mathcal{N}(m) = \mathcal{L} + \binom{\mathcal{L}}{2} = 2^{2^m} + \binom{2^{2^m}}{2} = \frac{2^{\mathcal{H}}}{2}(2^{\mathcal{H}} + 1) \quad (2)$$

The first term in the Eq. (2) represents the number of agents who have two identical strategies and the last term, the number of agents who have two different strategies. For the sake of simplicity, we will write \mathcal{N} instead of $\mathcal{N}(m)$ in the following part of this work. For example, in a game with $m = 2$, FSMG has $\mathcal{N} = 136$ agents. Ref [15] presents an extensive analysis of the FSMG and its application to obtain analytical results for the MG with different updating rules (standard MG , MG_{rand} just mentioned [9] and MG_{per} , a MG with a periodic updating rule introduced in [34]). Likewise, we have proved that the FSMG necessarily meets SPTD (Strict Period Two Dynamics), the PTD with probability equal to 1, and we have estimated the probability that a MG meets PTD from this analytical result.

By definition, the network of FSMG, G^{FSMG} , has \mathcal{N} nodes. Let us analyze the resulting network of FSMG, given the link definition proposed in (1). As \mathcal{N} is only a function of m , the set of connections, K , will also be a function of m . Then, the network and its properties will only depend on m , that is why we adopt the following notation for the FSMG network: $G^{FSMG}(\mathcal{N}(m), K(m))$ (although we use the simplified G^{FSMG}).

At a first stage, in subsection A we describe the analytical calculation of the degree distribution of G^{FSMG} as a function of m . At the second stage, in subsection B, we present the estimation of the degree distribution of MG networks from this exact result.

A. Calculation of the degree distribution of the network of the FSMG

Like before (see reference [15]), we will again benefit from the symmetry of the FSMG model, in which all pos-

sible strategy combinations in pairs (each of them representing one agent and therefore one node of the G^{FSMG} network) are present. As a result of this symmetry, the nodes of the FSMG network can take only a few values of degree, as we will proceed to prove. In fact, the number of different possible values of degree is, *at most*, $2^m + 1$. Actually, in the Appendix we will prove that the maximum is 2^m . For $m = 2$, for example, there exist nodes with only degree 0, 2, 3 and 14 on the FSMG network. To understand this, let us divide the set of agents (nodes) of the FSMG into different subsets, so that all nodes belonging to a given subset will be the same value of Hamming distance h between their pair of assigned strategies. We will show that nodes belonging to the same subset will have the same number of neighbors on the network though not the *same* neighbors and, as a consequence, they will be the same degree. The degree of each node from a given subset will be a function of both h and m .

We define the subset $\mathcal{N}_h \in \mathcal{N}$ as the set of agents whose pair of strategies show a Hamming distance equal to h between them. The number of subsets will be $2^m + 1$, because h can take on the values $h \in S_h = \{0, \frac{1}{2^m}, \frac{2}{2^m}, \dots, 1\}$. For example, in a FSMG of $m = 2$, there are five subsets, because h can take the values $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$, in general $h = \frac{i}{2^m}$ with i being an integer value in the range $[0, 2^m]$.

The symbol $\overline{\mathcal{N}_h}$ stands for the cardinality of a set, i.e. $\overline{\mathcal{N}_h}$ is the size of the subset \mathcal{N}_h . Let us count $\overline{\mathcal{N}_h}$ for each value of h , which is:

$$\overline{\mathcal{N}_h} = 2^{(2^m - 1)} \binom{2^m}{2^m h} \text{ for } h > 0 \text{ and } \overline{\mathcal{N}_0} = 2^{2^m} \quad (3)$$

The value of $\overline{\mathcal{N}_0}$ represents the number of agents whose two strategies are equal. There are exactly $\mathcal{L} = 2^{2^m}$ agents which meet this condition. Note that $\overline{\mathcal{N}_0}$ is the same value which appeared in the first term of the Eq. (2). In order to calculate the size $\overline{\mathcal{N}_h}$ we should count the number of agents whose two strategies differ in exactly $2^m h$ bits. This means that we need to count how many pairs of strategies there are so that the two strategies predict both different minority sides for $2^m h$ states, and the same minority sides for the $2^m(1 - h)$ remaining states. Then $\overline{\mathcal{N}_h}$ for $h > 0$ results from the product of three factors, where:

- factor 1 represents all the possible ways of choosing h states in which the pair of strategies make different predictions, thus: $\binom{2^m}{2^m h}$.
- factor 2 represents all the possible predictions for $(1 - h)2^m$ states, that is to say all those in which the pair of strategies coincide, thus: $2^{2^m(1-h)}$.
- factor 3 represents all the possible ways in which two strategies predict different minority sides for

the $2^m h$ states chosen in factor 1 (they can coincide because both strategies predict 1, or because both predict 0). Then, the dividing factor two is present because actually the pair of strategies is not an orderly pair, thus: $2^{2^m h}/2$.

For a particular m value and $h > 0$, the product of the three factors described results in expression (3), and the sum of all subset size will be the number of agents of the FSMG, thus:

$$\sum_{h \in S_h} \mathcal{N}_h = \mathcal{N} \quad (4)$$

as we will prove in the Appendix.

Now, let us suppose that we choose a node (that we call i) from a subset \mathcal{N}_h . Then, we know that the strategies of the i agent, e_1^i and e_2^i , are going to show a Hamming distance equal to h between them. In order to count the number of nodes of the G^{FSMG} that are connected to node i , let us count the number of strategies of the FSS meeting the condition of being separated from the two strategies of i agent (e_1^i and e_2^i) by a Hamming distance lower than $1/2$. That is to say, let us count the size of the set of strategies e that meet both conditions:

$$d(e_1^i, e) < \frac{1}{2} \quad \text{and} \quad d(e_2^i, e) < \frac{1}{2} \quad (5)$$

Now, we can imagine a box containing all strategies e meeting this condition. Although the composition of this box depends on the chosen node i , specifically of his pair of strategies, the size of the set of strategies in the box depends just on m and h . Therefore, we write $E(m, h)$ for the set of strategies in the box (for a particular i node from \mathcal{N}_h) and $\overline{E(m, h)}$ for the number of strategies in the box. Whenever we pick two strategies from this box (with repetition), we are *building* an agent (a network node, say j) who is going to have a connection to the chosen node i of the subset \mathcal{N}_h , thus $L_{ij} = 1$. Then, by combining the strategies from the set $E(m, h)$ in pairs (with repetition), we will get all the nodes that have a link with the chosen node i . The only consideration is that, sometimes, when the box of strategies $E(m, h)$ corresponding to a particular node from \mathcal{N}_h includes the particular node as one of the elements, we must then subtract a node (a pair of strategies), in order not to consider the link between a node and itself. This occurs for values of $h < 1/2$. Thus the degree of the particular node i , also applicable to any node belonging to \mathcal{N}_h , is going to be:

$$k(m, h) = \binom{\overline{E(m, h)}}{2} + \overline{E(m, h)} \quad \text{for } h \geq 1/2$$

$$k(m, h) = \binom{\overline{E(m, h)}}{2} + \overline{E(m, h)} - 1 \quad \text{for } h < 1/2 \quad (6)$$

All nodes belonging to the same subset \mathcal{N}_h will have the same degree, $k(m, h)$. However, the neighbors of each node will be different, in the same way as the composition of the box of strategies (with meet condition (5)) associated to each node is different.

As an example, in the current section, we will calculate $\overline{E(m, h)}$ for the particular case of $m = 2$, and we will leave the calculation for general values of m for the Appendix of this work. Table I can help us to understand how to compute the values of $\overline{E(m = 2, h)}$. We are interested in determining the size of those boxes of strategies associated with nodes whose pair of strategies show a Hamming distance of h (i.e. they differ in $2^m h$ bits). Thus, for each h value we choose one of the nodes belonging to the subset \mathcal{N}_h in order to show the particular set of strategies included in the subset $E(m, h)$ associated to this node. In panel $h = 0$ of Table I one node is shown as an example from the subset \mathcal{N}_0 . Then, both strategies of this node are identical ($e_1^i = e_2^i$) and the box of strategies associated to this node contains all strategies that differ in 0 or 1 bit from $e_1^i = e_2^i$, in order to meet condition (5). The size of this set is $\overline{E(m = 2, h = 0)} = 5$ strategies, that are shown in the box of strategies of the same panel. Then, from this box of strategies, we can build 15 different pairs of strategies, each of which represents a neighbor of the agent. But we will discount one of them (i.e. that with the same two strategies as the i node), in order not to consider a connection of the i node with itself. Thus, the degree of the i agent results in 14. And $k(m = 2, h = 0) = 14$ for all the nodes from the subset \mathcal{N}_0 . The cases $h = 1/2, 1/4, 3/4$ and $h = 1$ are represented in other panels and correspond to nodes that will be degree 2, 3, 0 and 0 in the FSMG with $m = 2$, and belong to the subsets $\mathcal{N}_{1/4}, \mathcal{N}_{1/2}, \mathcal{N}_{3/4}$ and \mathcal{N}_1 .

In order to obtain the degree distribution, it will be useful to define the probability of finding a node in the FSMG network whose strategies are separated by a Hamming distance given by h , thus $p(m, h) = \frac{\overline{N}_h}{\mathcal{N}}$. where \mathcal{N} and \overline{N}_h are actually functions of the m parameter. Therefore $p(m, h)$ is the probability of finding a node whose degree is $k(m, h)$ on the network. Then the degree distribution is going to show a set of peaks. There will be a peak for each of the possible values of $k(m, h)$. Hence, the maximum number of peaks is going to be the number of values that can take the variable h (i.e. the different subsets of nodes \mathcal{N}_h that could exist). Nevertheless, there could be fewer peaks of this maximum value because nodes belonging to two different subsets $\mathcal{N}_{h'}$ and $\mathcal{N}_{h''}$ may have the same value of degree $k(m, h') = k(m, h'')$, as it occurs in the example of $m = 2$, where $k(m, h = 3/4) = k(m, h = 1) = 0$. In the Appendix, we will show that in general $k(m, h = 1) = k(m, h = 1 - 1/2^m) = 0$.

As we will prove in the appendix, the general expression for $\overline{E(m, h)}$ is:

$$\overline{E(m, h)} = \sum_{i=\max(0, k-\delta')}^{\min(k, \delta')} \binom{k}{i} \sum_{j=0}^{\delta'-J(i)} \binom{2^m - k}{j} \quad (7)$$

where $J(i) = \max(i, k - i)$, $k = 2^m h$ and $\delta' = 2^m \delta$ with $\delta = 1/2 - 1/2^m$. Applying Eq. (6) in order to obtain $k(m, h)$ and using $p(m, h) = \overline{N}_h / \mathcal{N}$, the degree distribution of G^{FSMG} , that we call $P^{FSMG}(k)$, is parametrized in terms of h :

$$p(m, h) = \frac{2^{(\mathcal{H}-1)}}{\mathcal{N}} \begin{pmatrix} \mathcal{H} \\ \mathcal{H}h \end{pmatrix} \text{ for } h > 0$$

$$\text{and } p(m, h) = \frac{2}{2^{\mathcal{H}} + 1} \text{ for } h = 0 \quad (8)$$

Eq. (8) defines the probability of the values of the degree distribution in terms of h , and Eq. (6) describes the location of these values of degree also in terms of h . There is a discrete distribution with a maximum of 2^m possible values where, not only the maximum number of values of the degree but also their location and height depend on only parameter m .

Now we focus on the mode of the degree of G^{FSMG} which is reached when $h = 1/2$. In this case, \overline{N}_h and thus $p(m, h)$ reach the maximum value. Taking into account the limit case of $2^{\mathcal{H}} = 2^{2^m} \gg 1$ and using the well known approximation (which can be obtained straightforwardly from Stirling's formula) $\frac{1}{2^{\mathcal{H}}} \binom{\mathcal{H}}{\mathcal{H}/2} \sim \sqrt{\frac{2}{\mathcal{H}\pi}}$, we get

$$p(m, h = \frac{1}{2}) \simeq \sqrt{\frac{2}{\mathcal{H}\pi}}$$

In the Appendix, we will calculate the most probable value of k , k_{mod} , in the limit case of $2^{\mathcal{H}} = 2^{2^m} \gg 1$, which results in:

$$k_{mod} = k(m, h = \frac{1}{2}) \simeq \frac{\mathcal{N}}{16}$$

B. Estimating the degree distribution of MG networks from that of FSMG networks

We could estimate the degree distribution of G^{MG} from the degree distribution of G^{FSMG} calculated in the preceding section. The idea is to think about an instance of the MG of N agents as an statistical sample of size N from the set of \mathcal{N} agents of the FSMG. Then, the set of nodes of a particular network G^{MG} is a sample of N size of all the nodes of the FSMG network, and the G^{MG} is the induced subgraph from the G^{FSMG} by this set of N nodes. This sample is selected at random, because the assignment of strategies to agents of the MG is random. Thus, the new problem is well defined: we have to choose N nodes at random, with repetition (as in the MG there could be identical players) from the set of \mathcal{N} nodes of the G^{FSMG} . So, regarding this problem, the question is: how do we infer the degree distribution of G^{MG} ($P^{MG}(\tilde{k})$) from the known degree distribution of G^{FSMG} ($P^{FSMG}(k)$)? The relationship between N and \mathcal{N} tells us how representative the sample is. The

probability that a node of the FSMG be elected to form the network of MG is $q = N/\mathcal{N}$. As it was mentioned in subsection A, the degree distribution of FSMG networks, $P^{FSMG}(k)$, can take only a few values. The degree distribution of MG network will also be discrete (by definition of degree), but the variable may take any integer value, \tilde{k} (lower than the maximum value of degree of G^{FSMG}), with certain probability $P^{MG}(\tilde{k})$ as we will see in the following. Going from FSMG to MG for a particular value of m , we approximated $P^{MG}(\tilde{k})$ as:

$$P^{MG}(\tilde{k}) \simeq \sum_{k_i \geq \tilde{k}} P^{FSMG}(k_i) \binom{k_i}{\tilde{k}} q^{\tilde{k}} (1-q)^{k_i - \tilde{k}} \quad (9)$$

where k_i corresponds to the degree of the node belonging to a given subset \mathcal{N}_h , thus k_i is one of the possible values of $k(m, h)$ so that $\tilde{k} \leq k(m, h)$. To understand Eq. (9), let us consider a node belonging to the subset \mathcal{N}_h for which the degree is k_i . The probability that each of the neighboring nodes of this node is chosen is q . Therefore, the probability of this particular node which has degree k_i in the G^{FSMG} results with degree \tilde{k} in the G^{MG} is approximately a *Binomial*(k_i, q), which is the probability that \tilde{k} of his k_i neighbors being elected (by considering $\tilde{k} \leq k_i$). The same expression was obtained in [35] to study sampling process on networks. Here we use the approximation symbol in Eq. (9) because we are actually choosing exactly N nodes from \mathcal{N} rather than nodes with probability $q = N/\mathcal{N}$. The binomial factor has known variance and mean:

$$\langle \tilde{k} \rangle = k_i q \quad (10)$$

$$\sigma_k^2 = k_i q (1 - q) \quad (11)$$

Then the limit central theorem is used to approximate the binomial factor of the distribution for a normal distribution, with the same values of variance and mean:

$$P^{MG}(\tilde{k}) \simeq \sum_{k_i \geq \tilde{k}} P^{FSMG}(k_i) \frac{1}{\sqrt{2\pi}\sigma_k} e^{-(\tilde{k} - \langle \tilde{k} \rangle)^2 / 2\sigma_k^2} \quad (12)$$

For $q \ll 1$, we approximate $\sigma \simeq \sqrt{k_i q}$.

$$P^{MG}(\tilde{k}) \simeq \sum_{k_i \geq \tilde{k}} P^{FSMG}(k_i) \text{Normal}(k_i q, \sqrt{k_i q}) \quad (13)$$

where we note *Normal*(μ, σ) as the normal distribution with mean μ and variance σ^2 . The estimated degree distribution for MG network was compared with that obtained in the realization of networks of the MG. In Figure 7 it is possible to see the agreement when $N = 1001$

and $m = 3$. In this case, the degree distribution has five peaks. Note that the maximum number of peaks is $2^3 = 8$. The estimation of $P^{MG}(\tilde{k} = 0)$, i.e. the probability of finding a disconnected node is remarkable. In this case, $P^{MG}(\tilde{k} = 0) = P^{FSMG}(k = 0) = 0.035$.

Finally, by considering the limit case of $\mathcal{H} \gg 1$ and that all the agents of the FSMG have the value of degree k_{mod} mentioned in the previous section and calculated in Eq. (A20) in the Appendix, the expected value of \tilde{k} by sampling N nodes at random will be $\langle \tilde{k} \rangle \simeq qk_{mod} \simeq \frac{N}{N} \frac{1}{16} \simeq \frac{1}{16}$. Thus, the probability of finding links on the G^{MG} , in this limit case, is approximately $c \simeq \frac{1}{16}$.

IV. CONCLUSIONS

This work is an attempt to characterize the implicit interactions between MG agents as the links on a complex network. We have formalized an underlying network for the Minority Game, by quantifying the similarity of the strategies between a pair of agents. Given the resulting definition of the link, it can be said that in the MG region characterized by the presence of crowds, the underlying network can be identified as a small world network, whereas in the region where the system behaves like in a game of random decisions, the underlying network behaves as a random one, showing the same clustering coefficient, degree distribution and minimal path as a random Erdos-Renyi network. The transition between these two types of networks is gradual, that is why we cannot characterize the coordination region of the agents through a static network, which only contains information on the available strategies of the agents. In a context of non-ergodicity, similarities in the available set of strategies does not shed light on the actual moves of the agents. This fact poses a question on whether it would be possible to characterize the behavior of the game in the coordination region through a dynamic network, which also uses information regarding the strategies actually used during the game.

We have analytically calculated the degree distribution for the underlying network of the Full Strategy Minority Game (FSMG) model, and from this result, we have estimated the degree distribution of MG networks, with a very good agreement with the obtained results from simulations. This again shows how useful the FSMG is to understand the MG.

In the future it would be interesting to explore the effect of using weighed links in these networks. Additionally, in those cases where explicit interactions between some agents are introduced, as for example when some agents receive information from their neighbors, it could be helpful to consider the combined effects of explicit interactions network and the underlying interactions network formalized here. Let us remember that in the region of coordination of the agents, the population as a whole achieves more resources, but the inequality among

the agents is maximized in the coordination [19]. An interesting approach is to study if coevolutionary rules on the Minority Game can induce cooperation mechanism in this region as occurs in other social dilemmas [39].

ACKNOWLEDGMENTS

Ines Caridi would like to thank Horacio Ceva for his help on the onset of this work, and Gabriel Acosta for the useful discussions. Ines Caridi is a member of CONICET Argentina.

Appendix A

Firstly, we will prove Eq. (4), in which the sum of the number of agents of each subset \mathcal{N}_h over all the values of h results in \mathcal{N} , the total number of agents. Let us remember that h can take the values on the set $S_h = \{0, \frac{1}{2^m}, \frac{2}{2^m}, \dots, 1\}$. Thus h can be written as $h = f/2^m$, being f an integer number from 0 to 2^m . According to Eq. (3), the sum of \mathcal{N}_h for all values of h can be written as:

$$\sum_{h \in S_h} \mathcal{N}_h = \sum_{h \in S_h - \{0\}} 2^{(2^m-1)} \binom{2^m}{2^m h} + 2^{2^m} \quad (\text{A1})$$

we rewrite Eq. (A1) as:

$$\sum_{h \in S_h} \mathcal{N}_h = 2^{(2^m-1)} A + 2^{2^m} \quad (\text{A2})$$

where

$$A = \sum_{i \in \{0,1,\dots,2^m\}} \binom{2^m}{i} - \binom{2^m}{0} = 2^{2^m} - 1 \quad (\text{A3})$$

By replacing Eq. (A3) in Eq. (A2) we obtain the sought result:

$$\sum_{h \in S_h} \mathcal{N}_h = \frac{(2^{2^m} - 1)2^{2^m}}{2} + 2^{2^m} = \binom{2^{2^m}}{2} + 2^{2^m} = \mathcal{N} \quad (\text{A4})$$

Secondly, we will go on to detail how we carried out the calculation of $\overline{E(m,h)}$. As mentioned in III A, $E(m,h)$ is the set of strategies from which one can choose a pair of strategies to construct a node that will have a link with some particular node of the subset \mathcal{N}_h . Thus, the set $E(m,h)$ associated to the particular node i whose strategies are e_1^i and e_2^i is defined as the set of strategies that meet the condition of Eq. (5):

$$d(e_1^i, e) < \frac{1}{2} \quad \text{and} \quad d(e_2^i, e) < \frac{1}{2} \quad (\text{A5})$$

As we just mentioned, although the composition of $E(m, h)$ depends on the particular node selected from \mathcal{N}_h , the size of the set, $\overline{E(m, h)}$, is the same for all the nodes belonging to \mathcal{N}_h . Thus $\overline{E(m, h)}$ is only a function of m and h .

As h can take only the values of the set $\left\{0, \frac{1}{2^m}, \frac{2}{2^m}, \dots, \frac{2^{(m-1)}-1}{2^m}, \frac{1}{2}, \dots, 1\right\}$, by calling the threshold value

$$\delta = \frac{1}{2} - \frac{1}{2^m}, \quad (\text{A6})$$

condition (A5) becomes

$$d(e_1^i, e) \leq \delta \quad \text{and} \quad d(e_2^i, e) \leq \delta \quad (\text{A7})$$

Note that δ is actually a function of m , although we only write δ (and not $\delta(m)$) in order not to complicate the notation. For example, for the case of $m = 2$, $\delta = 1/4$, and for the case of $m = 3$, $\delta = 3/8$.

Before presenting the calculation for $\overline{E(m, h)}$, let us mention one aspect related to the transitivity of the Hamming distance between two strategies. Let us suppose that the Hamming distance between two strategies e_1 and e_2 is h , thus $d(e_1, e_2) = h$. And let us also consider a strategy e which differs from e_1 in exactly $2^m a$ bits. Thus, $d(e_1, e) = a$. What can we then said about the Hamming distance between e and e_2 , $d(e_2, e)$? We can say that $d(e_2, e)$ is limited by two boundaries:

$$|h - a| \leq d(e_2, e) \leq h + a \quad (\text{A8})$$

The high boundary ($h + a$) is reached when all the bits in which e differs from e_1 are chosen from those bits for which e_1 and e_2 coincide (and thus $a < 1 - h$). And the low boundary ($|h - a|$) is reached when all the bits in which e differs from e_1 are chosen from those bits for which e_1 and e_2 differ (when $a \leq h$); or when all the bits in which e differs from e_1 are all those bits for which e_1 and e_2 differ ($2^m h$) more other bits in which e_1 and e_2 coincide (when $a > h$). Once this property is described, we will focus on the calculation for $\overline{E(m, h)}$.

Let us consider a particular node belonging to the set \mathcal{N}_h , whose strategies are e_1 and e_2 , thus $d(e_1, e_2) = h$. In order to count the size of the set of strategies e that meet condition (A7), let us consider two sets of bits: the set \mathcal{B}_h of those $h2^m$ bits in which e_1 and e_2 differ, and the complementary set of $(1 - h)2^m$ bits in which e_1 and e_2 coincide, that we call $\overline{\mathcal{B}}_h$. Let us consider a strategy e so that $d(e_1, e) = a$; thus, e differs in $a2^m$ bits from e_1 . Additionally, let us suppose that $b2^m$ of these bits belong to the set \mathcal{B}_h , and $c2^m$ of these bits belong to the set $\overline{\mathcal{B}}_h$,

so that $b + c = a$. Then, the Hamming distances between strategies are:

$$d(e_1, e) = b + c \quad (\text{A9})$$

$$d(e_2, e) = h - b + c \quad (\text{A10})$$

Now we can ask what possible values b and c can take so that strategy e meets the condition (A7), knowing that $0 \leq b \leq h$ and $0 \leq c \leq 1 - h$. Hence, $b + c \leq \delta$ and $h - b + c \leq \delta$. As a consequence $0 \leq c \leq \delta - \max(b, h - b)$ and $\max(0, h - \delta) \leq b \leq \min(h, \delta)$. Now we can write

$$\overline{E(m, h)} = \sum_{i=\max(0, 2^m(h-\delta))}^{\min(2^m h, 2^m \delta)} \binom{2^m h}{i} \sum_{j=0}^{2^m \delta - J(i)} \binom{2^m(1-h)}{j} \quad (\text{A11})$$

being i an integer number between values 0 and $2^m h$ when $h \leq \delta$ and between values $2^m(h - \delta)$ and $2^m \delta$ when $h > \delta$; j is another integer number between values 0 and $2^m \delta - J(i)$ where $J(i) = \max(i, 2^m h - i)$. The first factor in Eq. (A11) represents all the ways in which we can choose i bits from the set \mathcal{B}_h , and the second factor, all the ways in which we can choose j bits from $\overline{\mathcal{B}}_h$ in order to construct a strategy e which differs in $i + j$ bits from e_1 and in $2^m h - i + j$ bits from e_2 .

A particular case included in Eq. (A11) is the case of $h = 0$. In this case, for a particular node belonging to the set $\mathcal{N}_{h=0}$, both of its strategies are equal, thus $d(e_1, e_2) = 0$. If we replace $h = 0$ in previous equation, we obtain

$$\overline{E(m, h = 0)} = \sum_{j=0}^{2^m \delta} \binom{2^m}{j} \quad (\text{A12})$$

which represents the total number of strategies e which differ in $0, 1, \dots, 2^m \delta$ bits from e_1 , i.e. the size of the set of strategies that meet the condition $d(e_1, e) = d(e_2, e) \leq \delta$.

Another particular case is the $h = 2\delta$ case. If we replace $h = 2\delta$ in Eq. (A11) we obtain

$$\overline{E(m, h = 2\delta)} = \binom{2\delta 2^m}{\delta 2^m} \quad (\text{A13})$$

which can be understood because, for a particular node of the set $\mathcal{N}_{h=2\delta}$, whose strategies are (e_1, e_2) , there is only one possibility for which strategy e differs from e_1 and from e_2 in less than $2^m \delta$ bits, or is equal to $2^m \delta$ bits. The possibility is that e differs exactly in $2^m \delta$ bits both from e_1 and from e_2 , and that those bits in which e differs from e_1 and e_2 should be chosen from those in which e_1 and e_2 differ. Because if we consider strategy e so that $d(e_1, e) = a = \delta$, then the low boundary for the distance $d(e_2, e) = 2\delta - a = \delta$ is reached only in this condition.

Lastly, another interesting case involves two values of h : $h = 1 - \frac{1}{2^m}$ and $h = 1$ (both strategies of a node are opposite), which correspond to the case $2\delta < h \leq 1$. In this case, using Eq. (A11) we obtain $\overline{E(m, 2\delta \leq h \leq 1)} = 0$. This fact can be understood because if we choose strategy e so that $d(e_1, e) \leq a$ where $a \leq 1/2 - 1/2^m$, then necessarily $d(e_2, e) \geq h - a$. Then the low boundary $h - a \geq \frac{1}{2}$ when $h = 1 - \frac{1}{2^m}$, and $h - a \geq \frac{1}{2} + \frac{1}{2^m} \geq \frac{1}{2}$ when $h = 1$. As a consequence, $d(e_2, e) \geq \frac{1}{2}$ and there is not any possible strategy e meeting condition (A7). That is why $\overline{E(m, 2\delta \leq h \leq 1)} = 0$. And for this fact we said that, as a maximum, the amount of peaks of degree for the G^{FSMG} is 2^m . Because there are always two cases of h ($h = 1$ and $h = 1 - 1/2^m$) for which the degree has the same value, 0.

Finally, the case of $m = 2$ described in the main part of the manuscript can be obtained by replacing $m = 2$ in previous analytical expressions:

$$\overline{E(m = 2, h = 0)} = 1 + 2^m = 5$$

$$\overline{E(m = 2, h = 1/4)} = 2^{2^m \delta} = 2$$

$$\overline{E(m = 2, h = 1/2)} = \binom{2}{1} = 2$$

$$\text{and } \overline{E(m = 2, h = 3/4)} = \overline{E(m = 2, h = 1)} = 0$$

Thirdly, we will calculate the value of the most probable degree for G^{FSMG} network, that we call k_{mod} , and whose probability is $P^{FSMG}(k_{mod})$. The maximum value of $p(m, h)$ is reached when $\overline{N_h}$ is maximum, which occurs when $h = 1/2$. In the main part of the manuscript, we have already discussed that in the limit case of $\mathcal{H} \gg 1$ this probability is approximately $\sqrt{\frac{2}{\mathcal{H}\pi}}$ (see the final part of section III A).

Let us calculate $\overline{E(m, h = 1/2)}$ in order to obtain the value of the most probable degree on the network, k_{mod} . We will now rewrite the expression of Eq. (A11) for the case of $h = 1/2$:

$$\overline{E(m, h = 1/2)} = \sum_{i=1}^{\mathcal{H}/2-1} \binom{\mathcal{H}/2}{i} \sum_{j=0}^{\mathcal{H}/2-1-J(i)} \binom{\mathcal{H}/2}{j} \quad (\text{A14})$$

with $J(i) = \max(i, \mathcal{H}/2 - i)$. Expression (A14) is the sum of some of the products of pairs of combinatorial coefficients in the form $\binom{\mathcal{H}/2}{i} \binom{\mathcal{H}/2}{j}$. By calling $C = \overline{E(m, h = 1/2)}$ the following expression is met (see [38] for a hint):

$$4C + 2 \sum_{i=0}^{\mathcal{H}/2} \binom{\mathcal{H}/2}{i}^2 - \left(\frac{\mathcal{H}/2}{\mathcal{H}/4}\right)^2 = \left[\sum_{i=0}^{\mathcal{H}/2} \binom{\mathcal{H}/2}{i} \right]^2 \quad (\text{A15})$$

By solving the two sums, Eq. (A15) is reduced to:

$$4C + 2 \binom{\mathcal{H}}{\mathcal{H}/2} - \left(\frac{\mathcal{H}/2}{\mathcal{H}/4}\right)^2 = 2^{\mathcal{H}} \quad (\text{A16})$$

In the limit case of $\mathcal{H} \gg 1$, we approximate $\binom{\mathcal{H}}{\mathcal{H}/2} \sim 2^{\mathcal{H}} \sqrt{\frac{2}{\mathcal{H}\pi}}$ (by using the Stirling's formula) and then:

$$C \simeq 2^{\mathcal{H}} \left[\frac{1}{4} + \frac{1}{\pi\mathcal{H}} - \frac{1}{\sqrt{2\pi\mathcal{H}}} \right] \quad (\text{A17})$$

In the limit case of $\overline{E(m, h = 1/2)} \gg 1$, thus $k_{mod} \simeq C^2/2$, we replaced Eq. (A17) in Eq. (6) to obtain the value of the k_{mod} :

$$k_{mod} \simeq \frac{2^{2\mathcal{H}}}{2} \left[\frac{1}{16} - \frac{1}{2\sqrt{2\pi\mathcal{H}}} + \frac{1}{\pi\mathcal{H}} - \frac{\sqrt{2}}{(\pi\mathcal{H})^{3/2}} + \frac{1}{(\pi\mathcal{H})^2} \right] \quad (\text{A18})$$

As in the limit case of $\mathcal{H} \gg 1$ we can approximate $\mathcal{N} \simeq \frac{2^{2\mathcal{H}}}{2}$, then

$$\frac{k_{mod}}{\mathcal{N}} \simeq \frac{1}{16} - \frac{1}{\sqrt{8}} a^{-1/2} + a^{-1} - \sqrt{2} a^{-3/2} + a^{-2} \quad (\text{A19})$$

being $a = \pi\mathcal{H}$. Finally, in the limit case of $a \gg 1$ ($\mathcal{H} \gg 1$) we obtain the expression:

$$\frac{k_{mod}}{\mathcal{N}} \simeq \frac{1}{16} \quad (\text{A20})$$

-
- [1] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press (1944).
 [2] R. Axelrod, *The evolution of cooperation*, Basic Book (1984).
 [3] R. Axelrod, *The complexity of cooperation*, Princeton

- University Press (1997).
 [4] M. Perc, A. Szolnoki, *Coevolutionary games A mini review*, BioSystems 99, 109125, 2010.
 [5] G. Szab, G. Fth, *Evolutionary games on graphs*, Physics Reports 446, 97216, 2007.
 [6] D. Challet and Y. C. Zhang, *Physica A*, **246**, 407, 1997.

- [7] D. Challet and Y. C. Zhang, *Physica A*, **256**, 514, 1998.
- [8] Arthur W. B., Am. Econ. Assoc. Papers and Proc 84, 406, 1994. (<http://www.santafe.edu/arthur/Papers/ElFarol.html>)
- [9] Cavagna A. Phys. Rev. E, **59**, R3783, 1999.
- [10] Marsili M., Challet D. and Zecchina R., *Physica A*, **280**, 522, 2000.
- [11] D. Challet, M. Marsili, Relevance of memory in minority games, *Phys Rev. E* 62 (2000) 1862.
- [12] K. H. Ho, W. C. Man, F. K. Chow, H. F. Chau, Memory Is Relevant In The Symmetric Phase Of The Minority Game, *Phys. Rev. E* 71 (2005) 066120.
- [13] D. Challet, M. Marsili, Y. C. Zhang, *Minority Games*, Oxford University Press (2005).
- [14] D. Challet, M. Marsili and R. Zecchina, *Physical Review Letters*, **84**, 1824, 2000.
- [15] G. Acosta, I. Caridi, S. Guala, J. Marengo, The Full Strategy Minority Game, *Physica A* 391 (2012) 217.
- [16] R. Manuca, Y. Li and R. Savit, *Physica A*, **339**, 574, 2000.
- [17] Challet D., Marsili M., *Phys. Rev. E*, **60**, 6271, 1999.
- [18] I. Caridi and H. Ceva, *Physica A*, **317**, 247, 2003.
- [19] K. H. Ho, F. K. Chow, H. F. Chau, Wealth inequality in the minority game, *Phys. Rev. E* 70, 066110, 2004.
- [20] G. Acosta, I. Caridi, S. Guala, J. Marengo, *Physica A* 391 (2013) 217.
- [21] A. Cavagna, J. P. Garrahan, I. Giardina and D. Sherrington, *Physical Review Letters*, **83**, 4429, 1999.
- [22] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D. H. Hwang, *Physics Reports*, **424**, 175, 2006.
- [23] A. L. Barabasi and R. Albert, *Review of Modern Physics*, **74**, 47, 2002.
- [24] S. H. Strogatz, *Nature*, **410**, 268, 2001.
- [25] S. Moelbert and P. De Los Rios, *Physica A*, **303**, 217, 2002.
- [26] T. Kalinowski, H. Schulz and M. Briese, *Physica A*, **277**, 502, 2000.
- [27] M. Paczuski, K. Bassler and A. Corral, *Physical Review Letters*, **84**, 3185, 2000.
- [28] A. Galstyan and K. Lerman, *Phys. Rev. E*, **66**, 015103, 2002.
- [29] L. Shang, X. F. Wang, *Physica A*, **377**, 616624, 2007.
- [30] B.A. Mello, V.M.C.S. Souza, D.O. Cajueiro, and R.F.S. Andrade, *Eur. Phys. J. B*, **76**, 147156, 2010.
- [31] T. S. Lo, H. Y. Chan, P. M. Hui, and N. F. Johnson, *Phys. Rev. E*, **70**, 056102, 2004.
- [32] I. Caridi and H. Ceva, *Physica A*, **339**, 574, 2004.
- [33] A. G. Frodesen, O. Skjeggstad, *Probability and Statistics in Particle Physics*, UNIVERSITETSFORLAGET (1979) pag. 415.
- [34] SS. Liaw, C.Liu, The quasi-periodic time sequence of the population in minority game, *Physica A* 351 (2005) 571.
- [35] Michael P. H. Stumpf, Carsten Wiuf, and Robert M. May, *PNAS*, **102**, 12, 4221-4224, 2005.
- [36] G. Csardi, T. Nepusz, The igraph software package for complex network research, *InterJournal, Complex Systems* (2006) 1695 (<http://igraph.sf.net>).
- [37] R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 (2008), URL <http://www.R-project.org>.
- [38] As a hint to obtain the expression, let us observe that the factors $\binom{\mathcal{H}/2}{i}$ from $i = 0$ to $i = \mathcal{H}/4$ can be arranged as both rows and columns of a matrix size $\mathcal{H}/4 + 1$. Let us think that each element of this matrix represents the product of the two factors associated to the column and the row of this element. Then let us tick off those elements on the matrix necessary to obtain C by adding them. By continuing ticking off those elements corresponding to the terms involved in expression (A15), it is possible for us to verify the relation.
- [39] A. Szolnoki and M. Perc, *EPL* 86, 30007, 2009.

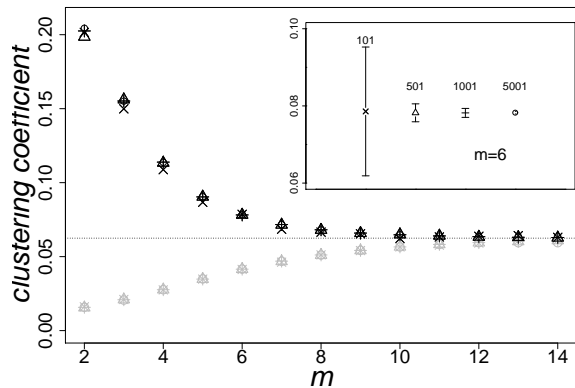


FIG. 5. Black symbols are the clustering coefficient as a function of m for games with different number of agents N (circle symbols for $N = 5001$ case, plus for 1001, triangle for 501 and X for 101 case). In all cases, the results shown are averages over 100 different instances. Gray symbols are the link probability just reported in Figure 1. The inset of the figure shows values of clustering coefficient for the case of $m = 6$ for different values of m ; error bars are the standard deviation obtained over the 100 instances. Error bars corresponding to the $N = 5001$ case are smaller than the size of the symbols. A χ^2 statistical test cannot reject the null hypothesis H_0 : “the values obtained from $N = 101, 501, 1001$ and 5001 are mutually consistent (i.e., correspond to variables with the same mean value)” for any value of m . The figure inset reports the values corresponding to $m = 6$ with error bars for different values of N . For each value of m we performed a χ^2 statistical test on means in order to check the null hypothesis H_0 : “the four values are mutually consistent within these error bars, i.e., the obtained values of clustering coefficient for $N = 101, N = 501, N = 1001$ and $N = 5001$ correspond to measurements of variables with the same mean value, and normal distribution” [33]. We computed the statistical value S as the weighed sum of the squared deviations from the weighed average value of the four measurements (\bar{x}), i.e., $S = \sum_{i=1}^4 (x_i - \bar{x})^2 / \Delta x_i^2$ where x_i and Δx_i are the variable and error corresponding to the i case (i.e. each one of the obtained clustering coefficients for G^{MG} with $N = 101, N = 501, N = 1001$ and with $N = 5001$ agents) and \bar{x} is the maximum likelihood estimator of the mean of the four values. When H_0 is true, S is approximately a χ_3^2 Chi-square distribution with 3 degrees of freedom (because the parameter \bar{x} is estimated from data) [33]. For $m = 2$, we obtained $S = 0.0013$; for $m = 3$, $S = 0.0009$; for $m = 4$, $S = 0.0009$; for $m = 5$, $S = 0.0006$; for $m = 6$, $S = 0.000008$, for $m = 7$, $S = 0.0007$; for $m = 8$, $S = 0.0002$; for $m = 9$, $S = 0.00003$; for $m = 10$, $S = 0.0006$; for $m = 11$, $S = 0.00005$; for $m = 12$, $S = 0.00002$; for $m = 13$, $S = 0.0004$ and for $m = 14$, $S = 0.0006$. For this reason H_0 cannot be rejected for any value of m .

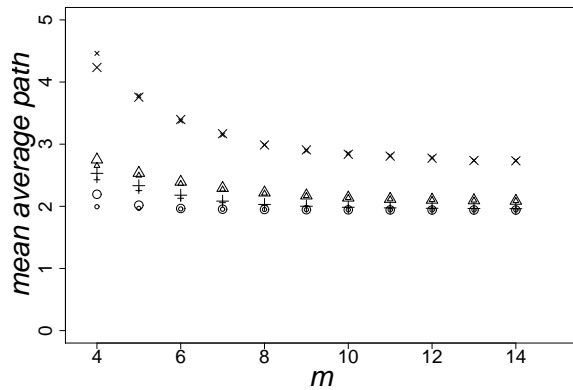


FIG. 6. Big symbols: minimum mean path as a function of m for games with different number of agents. In each case of MG network, data show the mean value and standard deviation of 100 instances (each instance consists in a different assignment of strategies to the agents). Small symbols: minimum mean path for Erdos-Renyi networks (G^{ER}) with the same value of N and $\langle k \rangle$ as G^{MG} , we used the *igraph* of R project package to simulate them and to calculate the minimum mean path [37], [36].

$h = 0$		$h = 1/4$	
chosen node i	box of strategies	chosen node i	box of strategies
$e_1^i \quad e_2^i$	$E(m = 2, h = 0)$	$e_1^i \quad e_2^i$	$E(m = 2, h = 1/4)$
1 1	1 0 1 1 1	1 0	1 0
1 1	1 1 0 1 1	1 1	1 1
1 1	1 1 1 0 1	1 1	1 1
1 1	1 1 1 1 0	1 1	1 1
	$\overline{E(m = 2, h = 0)} = 5$		$\overline{E(m = 2, h = 1/4)} = 2$
	$k(m = 2, h = 0) = \binom{5}{2} + 5 - 1 = 14$		$k(m = 2, h = 0) = \binom{2}{2} + 2 - 1 = 2$

$h = 1/2$		$h = 3/4$	$h = 1$
chosen node i	box of strategies	chosen node i	box of strategies
$e_1^i \quad e_2^i$	$E(m = 2, h = 1/2)$	$e_1^i \quad e_2^i$	$e_1^i \quad e_2^i$
1 0	1 0	1 0	1 0
0 1	0 1	1 0	1 0
1 1	1 1	1 0	1 0
1 1	1 1	1 1	1 0
	$\overline{E(2, h = 1/2)} = 2$		
	$k(m = 2, h = 1/2) = \binom{2}{2} + 2 = 3$	$k(h = 3/4) = 0$	$k(h = 1) = 0$

TABLE I. Each panel corresponds to a particular value of $h \in \{0, 1/4, 1/2, 3/4, 1\}$. In each panel we exemplify a particular node chosen from \mathcal{N}_h (*chosen node i*) and for this node, the set of strategies $E(m = 2, h)$ which meet the condition of Eq. (5) (*box of strategies*), and the size of this set $\overline{E(m, h)}$, which is applicable to every node from \mathcal{N}_h . Finally, we calculate the degree of the nodes from the \mathcal{N}_h set, $k(m, h)$.

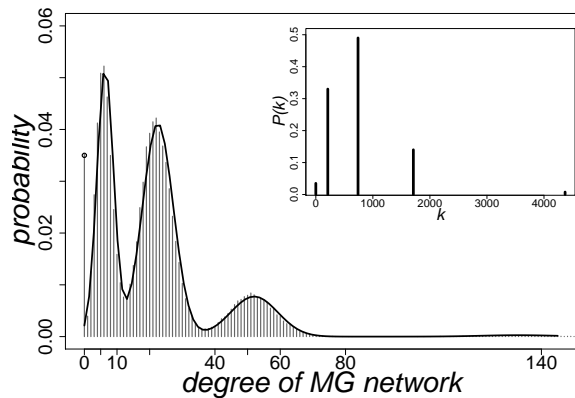


FIG. 7. In gray, degree distribution for MG networks, for $N = 1001$ and $m = 3$. Histogram is an average over 100 different instances. The continuous black line shows $P^{MG}(\tilde{k})$, the degree distribution estimated using the degree distribution of FSMG, $P^{FSMG}(k)$. The inset shows $P^{FSMG}(k)$ for $m = 3$: $P(k = 4370) = 0.0078$; $P(k = 1710) = 0.14$; $P(k = 740) = 0.49$ (actually there are two values of degree which were joined: 740, for the case of $h = 3/8$, and 741, for the case of $h = 1/2$); $P(k = 210) = 0.33$ and $P(k = 0) = 0.035$.