Crowd-Anticrowd Theory of Collective Dynamics in Competitive, Multi-Agent Populations and Networks

Neil F. Johnson

Physics Department, Oxford University, Oxford, OX1 3PU, UK

(Dated: October 21, 2003)

Abstract

We discuss a crowd-based theory for describing the collective behavior in a generic multi-agent population which is competing for a limited resource. These systems – whose binary versions we refer to as B-A-R (Binary Agent Resource) systems – have a dynamical evolution which is determined by the aggregate action of the heterogeneous, adaptive agent population. Accounting for the strong correlations between the choice of agents in a B-A-R system, we find that this theory can incorporate the effects of an underlying network within the population. Most importantly, its applicability is not just limited to the El Farol Problem and the Minority Game. Indeed, the Crowd-Anticrowd theory offers a powerful approach to tackling the dynamical behavior of a wide class of agent-based Complex Systems, across a range of disciplines. With this in mind, the present working paper is written for a general multi-disciplinary audience within the Agent-Based Complex Systems community.


+ Research collaboration with Pak Ming Hui, Sehyo Charley Choe, Sean Gourley and David Smith.

∗ Electronic address: n.johnson@physics.ox.ac.uk

1

I. INTRODUCTION

Complex Systems – together with their dynamical behavior known as Complexity – are thought to pervade much of the natural, informational, sociological, and economic world[1–5]. A unique, non-linear way of thinking about and understanding these systems provides a new viewpoint. In his paper, Casti[6] states that Complexity is a field of study where the interdependent actions of the elements of a system give rise to emergent structures and behaviors not present in the individual elements. As demonstrated by this Workshop, agent-based modeling and simulation provides a powerful approach to studying the dynamical behavior of a Complex System. Within this approach, agents interact with each other and with their environment using simple local rules. Thinking beyond the narrow confines of the study of a specific complex system, we might also uncover some fundamental theoretical principles concerning the behavior of Complex Systems as a whole.

Casti has argued that[1] '.... a decent mathematical formalism to describe and analyze the [so-called] El Farol Problem would go a long way toward the creation of a viable theory of complex, adaptive systems'. The rationale behind this statement is that the El Farol Problem is a simple model of human decision-making, a problem that arises when people must choose their behavior in a situation where others are doing the same. A good mathematical formalism to describe the El Farol Problem would provide a powerful tool for understanding how people make decisions in such situations. As a result, the El Farol Problem has received a great deal of attention from researchers in a variety of disciplines, including economics, psychology, and computer science.

In this paper, we present a crowd-based theory for describing the collective behavior of a multi-agent population. Our approach is based on the idea that the aggregate action of the agents in a population can be modeled using a simple mathematical formalism. This formalism allows us to describe the behavior of the population in terms of the choices made by the individual agents. In the case of the El Farol Problem, we can use this formalism to predict the behavior of the population and to understand how the agents make their decisions.

We hope that this work will be of interest to researchers in a variety of fields, including complex systems, decision-making, and computer science. We also hope that this work will help to advance our understanding of the behavior of Complex Systems as a whole.
The El Farol Problem, which was originally proposed by Brian Arthur [9] to demonstrate the essence of Complexity in financial markets involving many interacting agents, incorporates the key features of a Complex System in a way that is not easily replicated in other systems. For example, if all agents in the system made the same decision, then their common prediction scheme would become useless. This simple yet intriguing problem has inspired a huge amount of interest in the physics community over the years.

The Minority Game (MG) is discussed in detail in Refs. [12–39] and Ref. [40]. The Minority Game concerns a population of \( N \) heterogeneous agents with limited capabilities and information, who repeatedly compete to be in the minority group. The agents (e.g., people, cells, data-packets) are adaptive, but only have access to local information. In the natural world, most if not all biological and social systems have at least some degree of underlying connectivity between the agents, allowing for local exchange of information or physical goods [41]. In fact, it is this interplay of network structure, agent heterogeneity, and the resulting functionality of the overall system, which is likely to be of prime interest across disciplines in the Complex Systems field.

Three enormous challenges therefore face any potential 'Theory of Complex Systems': it must provide a quantitative explanation of the specific numerical results observed for the El Farol Problem 'test-case' and its variants; it should also be able to account for the presence of an arbitrary underlying network; yet it should be directly applicable to a much wider class of multi-agent 'game' with a much greater number of agents. These challenges are not easily solved, and the approach to solving them is likely to be different for each problem. It might then be possible to achieve the Holy Grail of predicting a priori how to engineer agents' reward structures, or which specific game rules to invoke, or what network communication scheme to introduce, in order to achieve some global objective [8].

In this paper, we attempt to take a step in this more general direction, by building on the success of the Crowd-Anticrowd theory in describing both the original El Farol Problem [10], and the Minority Game. The applicability of the Crowd-Anticrowd analysis is not limited to MG-like games, even though we focus on MG-like games in order to demonstrate the accuracy of the Crowd-Anticrowd theory in describing the Minority Game. The relevance of the Crowd-Anticrowd theory to a wider range of problems is depicted in Section II of this paper, which discusses the background to the Crowd-Anticrowd framework. The layout of the paper is as follows. Section III briefly discusses the background to the Crowd-Anticrowd framework. Section IV presents the results of the Crowd-Anticrowd theory applied to the El Farol Problem. Section V discusses the results of the Crowd-Anticrowd theory applied to the Minority Game. Section VI presents the results of the Crowd-Anticrowd theory applied to a wider range of problems.
provides a description of a wide class of Complex Systems which will be our focus: we call these B-A-R (Binary Agent Resource) problems in recognition of the stimulus provided by Arthur’s El Farol Problem. The problem is that of a complex multi-agent system, where the agents are connected in a network and interact with each other.

II. CROWD-ANTICROWD FRAMEWORK

The fundamental idea behind our crowd-based approach describing the dynamics of a complex multi-agent system, is to incorporate accurately the correlations in strategies followed by the agents. This is done by considering the system as a composite ‘super-particle’ which incorporates all the strong correlations in the system. A well-known example is an exciton gas: each exciton contains one negatively-charged electron and one positively-charged hole. In an analogous way, the Crowd-Anticrowd theory in multi-agent games forms groups containing like-minded agents (‘Crowd’) and opposite-minded agents (‘Anticrowd’). This is done in such a way that the strong strategy correlations are accounted for.

As indicated above, the Crowd-Anticrowd analysis breaks the N-agent population down into groups of agents, according to the correlations between these agents’ strategies. Each group G contains a crowd of agents using strategies which are positively-correlated, and a complementary anticrowd using strategies which are strongly negatively-correlated to the crowd. Hence a given group G might contain n_R[a] agents who are all using strategy R and hence act as a crowd (e.g. by attending the bar en masse in the El Farol Problem) together with n_R[b] agents who are all using the opposite strategy R and hence act as an anticrowd (e.g. by staying away from the bar en masse). Most importantly, the anticrowd n_R[b] will always take the opposite decisions to the crowd n_R[a] regardless of the current circumstances in the game, since the strategies R and R imply the opposite action in all situations. Note that this collective action may be entirely involuntary in the sense that typical games will be competitive, and there may be strategic interactions between agents. Since all the strong correlations have been accounted for within each group, these individual groups {G} = G_1, G_2, ..., G_n will then act in an uncorrelated way with respect to each other, and hence can be treated as n uncorrelated stochastic processes. The global dynamics of the system is then given by the sum of the n uncorrelated stochastic processes generated by the groups {G} = G_1, G_2, ..., G_n.

Regarding the special limiting case of the Minority Game (MG), we note that there have been various alternative theories proposed to describe the MG’s dynamics [13–17, 19, 35, 36, 38, 39]. Although elegant and sophisticated, such theories have however not been able to reproduce the original results of the correlations between agents’ strategies: in essence, these correlations produce a highly-correlated form of decision noise which cannot easily be averaged over or added in. By contrast these strong correlations take center-stage in the Crowd-Anticrowd theory, in a similar way to particle-particle correlations taking center-stage in many-body physics.
According to whether the strategy predicted the winning/losing action. One typical setup has agents adding/deducting one virtual point from strategies that would have been chosen. Strategies are then rewarded/penalized for each action. The global outcome is then announced. We note that in principle, the agents themselves do not actually need to know what game they are playing. Instead, they are fed with the global outcome: each of their actions is assigned a 0 (lost) or 1 (won). This binary representation of histories and strategies is due to Challet and Zhang [12].

The rules of the game determine the subsequent game dynamics. The particular rules chosen will depend on the practical system under consideration. It is an open question how appropriate 'games' can be constructed in order to capture the dynamics of real-world systems. The foraging mechanism adopted within such networks might well be active in other biological systems, and may also be of relevance in controlling arrays of imperfect nanostructures, microchips, nano-bio components, and other 'systems on a chip' [43].

III. B-A-R (BINARY AGENT RESOURCE) SYSTEMS

Figure 1 summarizes the generic form of the B-A-R (Binary Agent Resource) system. The rules of the game determine the subsequent game dynamics. The particular rules chosen will depend on the practical system under consideration. The participants must decide whether to enter a game where the choices are action +1 (e.g. attend the bar, take route A, or buy) and action −1 (e.g. go home, take route B, or sell). We consider the specific case of foraging fungal colonies as the market agent decides whether to enter a game where the choices are action +1 (e.g. attend the bar, take route A, or buy) and action −1 (e.g. go home, take route B, or sell). We use the notation for simplicity. We define a game as an arbitrary function of past values of the states of the agents. The foraging mechanism adopted within such networks might well be active in other biological systems, and may also be of relevance in controlling arrays of imperfect nanostructures, microchips, nano-bio components, and other 'systems on a chip' [43].
How do agents decide which strategy to use? Typically one might expect agents to play their highest-scoring strategy, as in the original El Farol Problem and Minority Game. However, agents may instead adopt a hedging strategy for choosing which strategy to use at each timestep. There may be 'dumb' agents, who use their worst strategy at each timestep. Of course, it is not obvious that such agents would then necessarily lose, i.e. whether such a method is actually 'dumb' will depend on the decisions of the other agents.

What happens in a strategy tie-break situation? Suppose agents are programmed to always use their highest-scoring strategy. If an agent were then to hold two or more strategies that are tied for the position of highest-scoring strategy, then a rule must be invoked for breaking this tie. One example would be for the agent to toss a fair coin in order to decide which of his tied strategies to use at that turn of the game. Alternatively, an agent may prefer to stick with the strategy which was highest on the previous timestep.

What are the rules governing the connections between agents? In terms of structure, the connections may be hard-wired or temporary, random or ordered or partially ordered (e.g., scale-free networks or small-world networks). In terms of functionality, there are also a number of possible choices of rules. For example, two connected agents might compare the scores of their highest-scoring strategies, and hence adopt the strategy of whichever agent has the higher score. The connections themselves may have directionality (e.g., perhaps agent a can influence agent b but not vice versa). Or maybe the connection (e.g. between b and a) is inactive unless a certain criterion is met. Hence it may turn out that agent b follows agent d even though agent b is actually following agent a.

Do agents have to play at each timestep? For simplicity in the present paper, we will consider that this is the case. This rule is easily generalized by introducing a confidence level: if the agent doesn't hold any strategies with sufficiently high confidence rate, then the agent does not participate at that timestep. This in turn implies that the number of agents participating at each timestep will vary, and hence the number of agents participating at each timestep will vary, and hence the volume of active agents [7, 33].

We emphasize that the setup in Figure 1 does not make any assumptions about the actual game being played, nor how the winning decision is to be decided. Instead, one hopes to obtain a fairly general description of the type of dynamics which are possible. For example, this extra property of confidence rate is a crucial ingredient for building a realistic market model, since it leads to fluctuations in the 'volume' of active agents [7, 33].

Figure 2 shows in more detail the $m=2$ example strategy space shown in Figure 1. A strategy is a set of instructions to describe what an agent should do in any given situation, i.e. given any particular history $\mu$, the strategy then decides what action the agent should take. The strategy space is the set of strategies from which agents are allocated their strategies. If this is a finite set, then the strategy allocation is fixed and can be described as a permutation of the strategies at the agent's disposal. The strategy space is then called a finite strategy space (FSS) and contains all possible permutations of the actions $-1$ and $+1$ for each history. As such there are $2^m$ strategies from the strategy space in this space. The $2^m$ dimensional hypercube shown contains all $2^m$ strategies from the strategy space in this space. The $2^m$ dimensional hypercube shown contains all $2^m$ strategies from the strategy space in this space.
D. History Space

Representations of the history

The history is the minimal set of strategies which span the FSS and form a Reduced Strategy Space (RSS) [12]. Since it contains the essential correlations of the Full Strategy Space (FSS), the RSS is not unique, i.e. within a given FSS there are many possible choices for a RSS. In particular, it is possible to choose strategies in the RSS which have a relative Hamming distance of 2. However, each strategy in the RSS has an anti-correlated strategy in the FSS. The advantage of using the FSS is that one can choose subsets of strategies such that any pair of strategies within this subset has a relative Hamming distance of 2.

For example, any two strategies within a Reduced Strategy Space have a Hamming distance of 2. Hence they will not contribute to fluctuations in the demand of the two agents will be uncorrelated on average. Consider, for example, the two groups of recent outcomes changes in time, i.e. it is a dynamical variable. The history dynamics can be represented on a directed graph (a so-called digraph). The particular form of directed graph is called a de Bruijn graph. Figure 3 shows some examples of the de Bruijn graph for $m = 2$. Although there are 2 strategies in the FSS but only 512 strategies in the RSS.

Any two strategies within a Reduced Strategy Space have a Hamming distance of 2. Hence they will not contribute to fluctuations in the demand of the two agents will be uncorrelated on average.

Consider, for example, the two groups of recent outcomes changes in time, i.e. it is a dynamical variable. The history dynamics can be represented on a directed graph (a so-called digraph). The particular form of directed graph is called a de Bruijn graph. Figure 3 shows some examples of the de Bruijn graph for $m = 2$. Although there are 2 strategies in the FSS but only 512 strategies in the RSS. We note that the choice of the RSS is not unique, i.e. within a given FSS there are many possible choices for a RSS. In particular, it is possible to choose strategies in the RSS which have a relative Hamming distance of 2. However, each strategy in the RSS has an anti-correlated strategy in the FSS. The advantage of using the FSS is that one can choose subsets of strategies such that any pair of strategies within this subset has a relative Hamming distance of 2.

For example, any two strategies within a Reduced Strategy Space have a Hamming distance of 2. Hence they will not contribute to fluctuations in the demand of the two agents will be uncorrelated on average. Consider, for example, the two groups of recent outcomes changes in time, i.e. it is a dynamical variable. The history dynamics can be represented on a directed graph (a so-called digraph). The particular form of directed graph is called a de Bruijn graph. Figure 3 shows some examples of the de Bruijn graph for $m = 2$. Although there are 2 strategies in the FSS but only 512 strategies in the RSS. We note that the choice of the RSS is not unique, i.e. within a given FSS there are many possible choices for a RSS. In particular, it is possible to choose strategies in the RSS which have a relative Hamming distance of 2. However, each strategy in the RSS has an anti-correlated strategy in the FSS. The advantage of using the FSS is that one can choose subsets of strategies such that any pair of strategies within this subset has a relative Hamming distance of 2.

For example, any two strategies within a Reduced Strategy Space have a Hamming distance of 2. Hence they will not contribute to fluctuations in the demand of the two agents will be uncorrelated on average. Consider, for example, the two groups of recent outcomes changes in time, i.e. it is a dynamical variable. The history dynamics can be represented on a directed graph (a so-called digraph). The particular form of directed graph is called a de Bruijn graph. Figure 3 shows some examples of the de Bruijn graph for $m = 2$. Although there are 2 strategies in the FSS but only 512 strategies in the RSS. We note that the choice of the RSS is not unique, i.e. within a given FSS there are many possible choices for a RSS. In particular, it is possible to choose strategies in the RSS which have a relative Hamming distance of 2. However, each strategy in the RSS has an anti-correlated strategy in the FSS. The advantage of using the FSS is that one can choose subsets of strategies such that any pair of strategies within this subset has a relative Hamming distance of 2.

For example, any two strategies within a Reduced Strategy Space have a Hamming distance of 2. Hence they will not contribute to fluctuations in the demand of the two agents will be uncorrelated on average. Consider, for example, the two groups of recent outcomes changes in time, i.e. it is a dynamical variable. The history dynamics can be represented on a directed graph (a so-called digraph). The particular form of directed graph is called a de Bruijn graph. Figure 3 shows some examples of the de Bruijn graph for $m = 2$. Although there are 2 strategies in the FSS but only 512 strategies in the RSS. We note that the choice of the RSS is not unique, i.e. within a given FSS there are many possible choices for a RSS. In particular, it is possible to choose strategies in the RSS which have a relative Hamming distance of 2. However, each strategy in the RSS has an anti-correlated strategy in the FSS. The advantage of using the FSS is that one can choose subsets of strategies such that any pair of strategies within this subset has a relative Hamming distance of 2.
E. Heterogeneity in strategy allocation, and initial conditions

The dynamics for a particular run of the B-A-R system will depend upon the strategies that the agents hold, and the random process used to decide tie-breaks. The particular dynamics which emerge also depend on the initial conditions of the system. The initial conditions can be described by a tensor Ω,[32] which represents the distribution of strategies among the N individual agents.

If the strategy allocation is fixed from the beginning of the game, then it acts as a quenched disorder in the system. The rank of the tensor Ω is given by the number of strategies S that each agent holds. For example, for S = 3, the element Ω_{i,j,k} gives the number of agents assigned strategy i, then strategy j, and then strategy k, in that order.

The initial choice of history is not considered to be an important effect. It is assumed that any transient effects resulting from the particular history seed will have disappeared, i.e. the initial history seed does not introduce any long-term bias.

The strategy allocation among agents can be described in terms of a tensor Ω. This tensor Ω describes the distribution of strategies among the N individual agents. If this strategy allocation is fixed from the beginning of the game, then it acts as a quenched disorder in the system. The rank of the tensor Ω is given by the number of strategies S that each agent holds. For example, for S = 3, the element Ω_{i,j,k} gives the number of agents assigned strategy i, then strategy j, and then strategy k, in that order. Hence

\[ \sum_{i,j,k,...} \Omega_{i,j,k,...} = N, \]

where the value of \( \sum \) represents the number of distinct strategies that exist within the strategy space chosen:

\[ \sum^X = 2^m \text{ in the FSS, and } \sum^Y = 2^m \text{ in the RSS}. \]

Figure 4 shows an example distribution Ω for N = 101 agents in the case of m = 2 and S = 2, in the reduced strategy space RSS. We note that a single Ω 'macrostate' corresponds to many possible 'microstates' describing the specific partitions of strategies among the agents. For example, consider an N = 2 agent system with S = 1: the microstate (R, R') in which agent 1 has strategy R while agent 2 has strategy R', belongs to the same macrostate Ω as (R', R) in which agent 1 has strategy R' while agent 2 has strategy R. Hence the present Crowd-Anticrowd theory retained at the level of a given Ω, describes the set of all games which belong to that same Ω. For example, consider a two-agent system with S = 2. The microstate (R, R') in which agent 1 has strategy R while agent 2 has strategy R', belongs to the same macrostate Ω as (R', R) in which agent 1 has strategy R' while agent 2 has strategy R. Hence the present Crowd-Anticrowd theory retained at the level of a given Ω, describes the set of all games which belong to that same Ω.

In general, \( \Omega \) would be allowed to change in time, possibly evolving under some pre-defined selection criteria. Elsewhere we will show how changes in \( \Omega \) can be invoked in order to control the future evolution of the system, or to avoid certain future scenarios or trajectories for the system as a whole.[47]

F. Design Criteria for Collective

In addition to the excess demand \( D[t] \) in such B-A-R systems, one is typically also interested in higher order-moments of this quantity – for example, the standard deviation of \( D[t] \) (or 'volatility' in a financial context). This gives a measure of the fluctuations in the system, and hence can be used as a measure of 'risk' in the system. In particular, the standard deviation gives an idea of the size of typical fluctuations in the system. However, if \( D[t] \) has a power-law distribution, the standard deviation may be a misleading representation of such risk because of the 'heavy tails' associated with the power-law.

Practical risk is arguably better associated with the probability of hitting a critical value \( D_{crit} \) in a similar way to the methodology of financial Value-at-Risk.[7] However, a note of caution is worthwhile: if there are high-order correlations in the system and hence in the time-series \( D[t] \) itself, any risk assessment based on Probability Distribution Functions over a fixed time-scale (e.g. single timestep) may be misleading. Instead, it may be the cumulative effects of a string of large negative values of \( D[t] \) which constitute the true risk, since these may take the system into dangerous territory. The possibility of designing a Collective in order to reduce such large cumulative changes or endogenous 'extreme events', or alternatively, the possibility of performing 'on-line' soft control in an evolving system, is a fascinating topic.[7, 47] For simplicity, we will focus here on developing a description of the excess demand \( D[t] \) and its standard deviation, noting that the same analytic approach would work equally well for other statistical functions of \( D[t] \).
special case in which all histories are visited equally on average: this may arise as the result of uncorrelated. We now consider the property that

\[ t \equiv 1 \]

where we have used the property that \( n \equiv 1 \). The calculation of the average demand \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]

will now be shown, where the average \( D \) is a reminder that this number of agents will depend on the strategy score \( t \). The quantity \( D \) is given by Equation 1. The standard deviation

\[ \langle |\mu(\Psi)| - u \rangle \]
\[
\sum_{P=1}^{R} \left( \begin{array}{c}
\frac{\sqrt{\sigma_n}}{\sqrt{S}} \langle \mu \rangle \\
\frac{\sigma_n}{S}
\end{array} \right) \left( \begin{array}{c}
\frac{\sqrt{\sigma_n}}{\sqrt{S}} \langle \mu \rangle \\
\frac{\sigma_n}{S}
\end{array} \right)
\]

\[
\sum_{P=1}^{R} \left( \frac{\sqrt{\sigma_n}}{\sqrt{S}} \langle \mu \rangle \right) \left( \frac{\sigma_n}{S} \right) = \frac{\sigma_n^2}{S}
\]

\[
\sum_{P=1}^{R} \left( \frac{\sqrt{\sigma_n}}{\sqrt{S}} \langle \mu \rangle \right) \left( \frac{\sigma_n}{S} \right) = \frac{\sigma_n^2}{S}
\]

\[
\sum_{P=1}^{R} \left( \frac{\sqrt{\sigma_n}}{\sqrt{S}} \langle \mu \rangle \right) \left( \frac{\sigma_n}{S} \right) = \frac{\sigma_n^2}{S}
\]

Equation 12 is an important intermediary.
and R that any agent holding strategies \( R \) or \( R' \) will play at this timestep. This means that being assigned the rank \( K \) \( t \) now becomes degenerate at timestep \( R \). Suppose that the two strategies become degenerate at timestep \( R \) and \( R' \) of agents playing strategy \( K \) \( t \). Hence

\[
\Psi_{t+1} = \Psi_t + \frac{1}{2}
\]

Equation 1 can hence be rewritten exactly as

\[
\Psi_{t+1} = \Psi_t + \frac{1}{2}
\]

Consider a given timestep \( t \). In a game where each agent resolves strategy tie-breaks using a coin-toss, then both strategies might vary, using the new virtual-point ranking scheme. Also shown are the lowest-scoring two strategies, which in this case, any two agents holding the same pair of degenerate strategies may then disagree as to the ranking of the two strategies. Depending on the exact microscopic rule adopted, the global dynamics on the rules of the game, such degeneracy may imply one of the following: (i) an agent must throw a coin to break a tie between any two degenerate strategies in his possession. In some situations (ii) the game evolution may also be influenced by the strategy scores. For example, in the case that the strategies all start off with zero points, these anticorrelated strategies appear as the mirror-image, i.e. all strategy scores start off at zero, then we know that \( S_1 = S_2 \)

\[
S_1 - S_2 = 0
\]

\[
S_1 > S_2
\]

\[
S_1 < S_2
\]

\[
S_1 = S_2
\]

which is used to denote the rank in terms of strategy score, i.e. the \( n \)th highest scoring strategy. The label \( t \) to the time-evolution of the virtual points of the strategy would imply one of the following: (i) an agent must throw a coin to break a tie between any two degenerate strategies in his possession. In some situations (ii) the game evolution may also be influenced by the strategy scores. For example, in the case that the strategies all start off with zero points, these anticorrelated strategies appear as the mirror-image, i.e. all strategy scores start off at zero, then we know that \( S_1 = S_2 \)

\[
S_1 - S_2 = 0
\]

\[
S_1 > S_2
\]

\[
S_1 < S_2
\]

\[
S_1 = S_2
\]

which is used to denote the rank in terms of strategy score, i.e. the \( n \)th highest scoring strategy. The label \( t \) to the time-evolution of the virtual points of the strategy would imply one of the following: (i) an agent must throw a coin to break a tie between any two degenerate strategies in his possession. In some situations (ii) the game evolution may also be influenced by the strategy scores. For example, in the case that the strategies all start off with zero points, these anticorrelated strategies appear as the mirror-image, i.e. all strategy scores start off at zero, then we know that \( S_1 = S_2 \)

\[
S_1 - S_2 = 0
\]

\[
S_1 > S_2
\]

\[
S_1 < S_2
\]

\[
S_1 = S_2
\]

which is used to denote the rank in terms of strategy score, i.e. the \( n \)th highest scoring strategy. The label \( t \) to the time-evolution of the virtual points of the strategy would imply one of the following: (i) an agent must throw a coin to break a tie between any two degenerate strategies in his possession. In some situations (ii) the game evolution may also be influenced by the strategy scores. For example, in the case that the strategies all start off with zero points, these anticorrelated strategies appear as the mirror-image, i.e. all strategy scores start off at zero, then we know that \( S_1 = S_2 \)

\[
S_1 - S_2 = 0
\]

\[
S_1 > S_2
\]

\[
S_1 < S_2
\]

\[
S_1 = S_2
\]

which is used to denote the rank in terms of strategy score, i.e. the \( n \)th highest scoring strategy. The label \( t \) to the time-evolution of the virtual points of the strategy would imply one of the following: (i) an agent must throw a coin to break a tie between any two degenerate strategies in his possession. In some situations (ii) the game evolution may also be influenced by the strategy scores. For example, in the case that the strategies all start off with zero points, these anticorrelated strategies appear as the mirror-image, i.e. all strategy scores start off at zero, then we know that \( S_1 = S_2 \)

\[
S_1 - S_2 = 0
\]

\[
S_1 > S_2
\]

\[
S_1 < S_2
\]

\[
S_1 = S_2
\]

which is used to denote the rank in terms of strategy score, i.e. the \( n \)th highest scoring strategy. The label \( t \) to the time-evolution of the virtual points of the strategy would imply one of the following: (i) an agent must throw a coin to break a tie between any two degenerate strategies in his possession. In some situations (ii) the game evolution may also be influenced by the strategy scores. For example, in the case that the strategies all start off with zero points, these anticorrelated strategies appear as the mirror-image, i.e. all strategy scores start off at zero, then we know that \( S_1 = S_2 \)

\[
S_1 - S_2 = 0
\]

\[
S_1 > S_2
\]

\[
S_1 < S_2
\]

\[
S_1 = S_2
\]
The mean square deviation of an initial function \( f(x) \) as long as we neglect tie-breaks, we will also have

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]

Hence, the rankings in terms of highest virtual-points and popularity are identical. By choosing a suitable constant \( \sigma \) which in turn will depend on the standard probability result involving functions of two variables has been used.

In the event that Equation 17 is a reasonable approximation for the system under study, the question then arises as to how to evaluate it. In general, its value will depend on the detailed form of the joint probability function \( P(x, y) \) which in turn will depend on the detailed form of the joint probability distribution for the system under study. Each term in the sum \( \langle n \rangle \) will be sequential, i.e. it is not possible to evaluate the individual terms in the sum, i.e.

\[
\langle n \rangle = \sum_{n} P(n) n = \langle n \rangle.
\]
\[ N \left( \frac{dE}{dP} - 1 \right) - \left( N - \frac{dE}{dP} - 1 \right) N = \frac{M}{2} \]

In the case where each agent holds two strategies, \( m = 2 \), and the number of strategies \( S \) is large compared to the number of agents \( N \), the number of strategies \( S \) can be simplified to \( S = 2 \). In the case where each agent holds two strategies, \( m = 2 \), the strategy allocation matrix \( \Psi \) can be filled sparsely.

In this section, we will consider the application of the Crowd-Anticrowd theory in some limiting cases. As suggested by the previous discussion, we will break the implementation of the Crowd-Anticrowd theory down into two separate regimes: small and large. Hence these two regimes are defined by the ratio of the number of strategies to agents being much less/greater than unity, and hence the strategy allocation matrix \( \Psi \) being densely/sparsely filled.

A. Implementation of Crowd-Anticrowd Theory

The Crowd-Anticrowd theory is implemented under the presence of network connections which naturally lead to many more agents than available strategies, and large limiting cases. In the presence of network connections - this is because very small values this becomes

\[ \sum_{i=1}^{S} \left( P_{\text{score} - 1} \right) \]

Formore general cases, this becomes

\[ \sum_{i=1}^{S} \left( P_{\text{score} - 1} \right) \]

Each element of \( \Psi \) has a mean of \( \frac{1}{N} \) and hence the effective group-size for the Crowd-Anticrowd pair in a random-walk contribution to the total variance. Hence, the net contribution by this Crowd-Anticrowd pair to the variance is given by

\[ \frac{(\sum_{i=1}^{S} \left( P_{\text{score} - 1} \right))}{N} = \frac{1}{N} \]
As a switch is now made to the popularity-labels the general analysis is much more complicated, and should really appeal to the dynamics. However, an approximate theory can still be developed. The features of the case of ensembles will overestimate the upper bound.

As we will see later in Figure 7, the disorder in the matrix \( \Psi \) may distort the number of agents playing a given strategy to such an extent, that it is no longer true that

\[
\sum_{n=0}^{N} f_n - n\sigma_f = 0
\]

The appearance of a significant number of non-Gaussian disorder matrices in the structure of the matrix \( \Psi \) does indeed act as an approximate upper bound. (Here the term \( \delta \) is valid for small \( \frac{n}{N} \))

\[
\frac{1}{N} \left( \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \left( \frac{1}{N} \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \right) \right) = \sigma_f
\]

\[
\frac{1}{N} \left( \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \left( \frac{1}{N} \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \right) \right) = \sigma_f
\]

\[
\frac{1}{N} \left( \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \left( \frac{1}{N} \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \right) \right) = \sigma_f
\]

\[
\frac{1}{N} \left( \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \left( \frac{1}{N} \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \right) \right) = \sigma_f
\]

\[
\frac{1}{N} \left( \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \left( \frac{1}{N} \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \right) \right) = \sigma_f
\]

\[
\frac{1}{N} \left( \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \left( \frac{1}{N} \sum_{\delta=1}^{M} \sum_{\delta=1}^{N} \right) \right) = \sigma_f
\]
\[ \mathcal{O} = 1 + d \mathcal{D} = \mathcal{D} \] 

for higher moments. The proof of this result will be given in a separate section. 

\[ \mathcal{D}' \] 

where the factor \( d \mathcal{D} \) in Equation (27) has been dropped. Hence \( \mathcal{D}' = (d \mathcal{D} + d \mathcal{D}) \) and is now labelled by \( \mathcal{D}' \). In the limit of the product matrix, it will be a product function, and at small values of \( \mathcal{D}' \) or in the limit of the product matrix, the function \( \mathcal{D}' \) will be sharply peaked around the zeroth-order approximation the values of \( \mathcal{D}' \) are anticorrelated to \( \mathcal{D}_0 \) and \( \mathcal{D}_0 \) is described by the bin-counting method. Substituting \( \mathcal{D}' = 1 \) into Equation (27) we have

\[ \mathcal{D}' = \frac{(1 + d \mathcal{D}) N^\lambda}{N} = \frac{d \mathcal{D} N^\lambda}{N} = \mathcal{D} \]
in picking their most successful strategy. For general S, this then yields
\[ \sigma_{\text{flat}}^f, \text{high}_m = \sqrt{N \left(1 - \frac{\text{NS} - 1}{2} m + 1\right)} \]  
\[ (36) \]

[N.B. This latter expression gives an even better fit to the numerical results shown later in Figure 7.]

D. Reduced vs. Full Strategy Space

The general dynamics for a typical game are likely to be similar in character when played in both the FSS and the RSS [12]. Hence the RSS-based Crowd-Anticrowd theory would generally be expected to provide a good approximation. However, for a game of interest, there is a theoretical justification which goes as follows. For a game played in the FSS there are \( 2^m \) distinct subsets of strategies. Each subset can be considered as a separate RSS. Note that the strategies that belong to a given RSS are optimally spread out across the corresponding FSS hypercube. Figure 2 shows the distribution of the 16, \( m = 2 \) strategies across the 4 dimensional hypercube. The positions of the strategies belonging to the RSS are such that no two strategies have a Hamming separation less than \( \frac{2^m}{2} \). The same can be said for any other choice of RSS. Due to the nature of a RSS, and given similar strings of past outcomes over which to score strategies, each strategy within the RSS attains a score in an uncorrelated or independent manner. Any other RSS within the FSS will score its strategies in a similar way, although slightly out of phase. For example for \( m = 3 \), the first RSS to be considered could contain the strategy 00001111, and the second RSS considered could contain the strategy 01001111. It is easy to see that the RSS could contain the strategies 10000000 and 11101111. For example, in \( m = 3 \), the RSS to be considered could contain the strategies 10000000 and 11101111 in an uncorrelated or independent manner. Any other RSS in the FSS can be obtained by adding or subtracting the Hamming separation less than \( \frac{2^m}{2} \) from the above. The same can be said for any other choice of RSS. The resulting dynamics of a game of a RSS and full strategy space are similar. The detailed derivation of the above formula is given in Appendix B. For future work.

VI. CROWD-ANTICROWD THEORY APPLIED TO GENERALIZED MINORITY GAMES

In this section we show that the present Crowd-Anticrowd theory, even within its RSS formulation, provides a quantitative theory for both the Minority Game and several variants. There is currently no other competitive model that can provide a good quantitative description over such a range of game variants. This justifies our belief that the Crowd-Anticrowd theory is an important milestone in such multi-agent systems, and possibly even for Complex Systems in general. It is also pleasing from the point of view of an economic interpretation of the RSS the model is not only able to provide a good quantitative description of the system, but is also in agreement with the empirical results of the economic interpretation of the RSS the model is not only able to provide a good quantitative description of the system, but is also in agreement with the empirical results of the model.
(Y)^d (Y - 1) - \theta_{1+\omega=\varepsilon} = (Y)^d \theta

\sum_{i=1}^{M_{\text{L}}} = m^L

\text{Agents, this analytic result has been confirmed in numerical}

\text{probability that he uses the best. The probability that the agent plays}

\text{contrast, let } S \text{thetab be the probability that a given agent possesses}

\text{As mentioned earlier, in the small}

\text{We will now show that the Crowd-Anticrowd theory provides a quantitative explanation}

\text{The common underlying phenomenon}

\text{In the 'Thermal Minority Game' (TMG) \cite{38}, agents choose between their}

\text{In the mixed population, we assume non-adjacent bit-strings}

\text{are pushed below the}
The Crowd-Anticrowd theory can be extended to account for the effect of fluctuations in strategy-use for agents playing their worst strategy with increasing probability. An increase in the parameter $\sigma$ increases the standard deviation of the resulting distribution describing the strategy-use, as in other words, the standard deviation decreases as $\theta$ increases (recall $0 \leq \theta < 1$). The mean value $\sigma_0$ is obtained from Equations 40 and 41 for $\theta = 0$.

The quantities $\mu$ and $\sigma$ are now functions of $\theta$ (see Equation 42). Reducing that only the first term in Equation 42 actually a function of $\theta$ and $K$ and hence substituting Equations 40 and 42 for $K$ and $\sigma$ into Equation 19 yields

$$\sigma^2 = \sum p_i (x_i - \mu)^2.$$

where $\sigma^2$ is the standard deviation, $\mu$ is the mean value, and $x_i$ is the number of agents playing strategy $i$. The observed higher-than-random to smaller-than-random as $\theta$ increases (recall $0 \leq \theta < 1$) in other words, the standard deviation decreases as $\theta$ increases.

In the TMG, each agent is equipped at each timestep with a random number of strategies, each of which it can choose to use at random probability $\theta$ as well as at a particular moment in time. The number of agents using each strategy is typically distributed around the mean value $\mu$ and hence $\sigma^2$ is the standard deviation when $\theta = 0$. The resulting distribution describing the strategy-use is therefore non-flat. It is these fluctuations about the mean values $\mu$ and $\sigma$ which give rise to the random courtesies result as $\theta = 0$. The Crowd-Anticrowd theory can be explained to account for the effect of fluctuations in strategy-use for agents playing their worst strategy with increasing probability. An increase in $\theta$ decreases $\sigma$ and hence $\sigma_0$.

Note that $p_i (K) = (1 - 2\gamma) N(K) + \gamma N(K) - 1$ where $\gamma = 0.03$, $N(K) = 2^{K - 1}$. The numerical data for different runs has a significant natural spread. Most of these data points do lie in the region in between the two analytic curves, which act as approximate upper and lower-bounds. Above $\theta = 0.15$, the numerical data tend to flatten off while the present theory is accurate to within $0.3\%$. There is a wide range of strategies picked at the start of the game, and at a particular moment in time, the number of agents using each strategy is typically distributed around the mean value $\mu$ and hence $\sigma^2$ is the standard deviation when $\theta = 0$. The resulting distribution describing the strategy-use is therefore non-flat. It is these fluctuations about the mean values $\mu$ and $\sigma$ which give rise to the random courtesies result as $\theta = 0$. The Crowd-Anticrowd theory can be explained to account for the effect of fluctuations in strategy-use for agents playing their worst strategy with increasing probability. An increase in $\theta$ decreases $\sigma$ and hence $\sigma_0$.

Note that $p_i (K) = (1 - 2\gamma) N(K) + \gamma N(K) - 1$ where $\gamma = 0.03$, $N(K) = 2^{K - 1}$. The numerical data for different runs has a significant natural spread. Most of these data points do lie in the region in between the two analytic curves, which act as approximate upper and lower-bounds. Above $\theta = 0.15$, the numerical data tend to flatten off while the present theory is accurate to within $0.3\%$. There is a wide range of strategies picked at the start of the game, and at a particular moment in time, the number of agents using each strategy is typically distributed around the mean value $\mu$ and hence $\sigma^2$ is the standard deviation when $\theta = 0$. The resulting distribution describing the strategy-use is therefore non-flat. It is these fluctuations about the mean values $\mu$ and $\sigma$ which give rise to the random courtesies result as $\theta = 0$. The Crowd-Anticrowd theory can be explained to account for the effect of fluctuations in strategy-use for agents playing their worst strategy with increasing probability. An increase in $\theta$ decreases $\sigma$ and hence $\sigma_0$.
MG is in the worse-than-random regime, and therefore $0 < \sigma_1^2 - \sigma_2^2 < 0$.

Since the $T_{\theta}^{(i)}$th highest scoring $K$-Recall from Equation 39, the probability that the agent plays the $q, \theta$ action is given by $\sigma_1^2 - \sigma_2^2 < 0$.

The mean number of agents in the mixed-population game containing $K_{\text{TMG}}$ agents and $N_{\text{MG}}$ agents is now considered. This is given by

$$\langle K \rangle = \sum_{K_{\text{TMG}}} K_{\text{TMG}} \left[ \left( \frac{N_{\text{TMG}}}{N_{\text{MG}}} \right)^{K_{\text{TMG}}} \times \left( \frac{N_{\text{MG}}}{N_{\text{TMG}}} \right)^{K_{\text{MG}}} \right]$$

where $N_{\text{TMG}} + N_{\text{MG}} = N$. Provided that the basic MG is in the crowded regime as discussed earlier, Equation 52 predicts that the effect on the standard deviation caused by a change in population composition and/or 'temperature' can be described by a simple prefactor. Hence the theoretical connection with the Thermal Minority Game discussed above is obtained. The remaining $(1 - \theta)N_{\text{TMG}}$ agents will be called 'TMG agents' because of the direct connection with the Thermal Minority Game. A variation of the Thermal Minority Game is now considered in which a concentration $q < q_0$ is chosen.

The mean number of agents in the mixed-population game containing $K_{\text{TMG}}$ agents and $N_{\text{MG}}$ agents is now considered. This is given by

$$\langle K \rangle = \sum_{K_{\text{TMG}}} K_{\text{TMG}} \left[ \left( \frac{N_{\text{TMG}}}{N_{\text{MG}}} \right)^{K_{\text{TMG}}} \times \left( \frac{N_{\text{MG}}}{N_{\text{TMG}}} \right)^{K_{\text{MG}}} \right]$$

where $N_{\text{TMG}} + N_{\text{MG}} = N$. Provided that the basic MG is in the crowded regime as discussed earlier, Equation 52 predicts that the effect on the standard deviation caused by a change in population composition and/or 'temperature' can be described by a simple prefactor. Hence the theoretical connection with the Thermal Minority Game discussed above is obtained. The remaining $(1 - \theta)N_{\text{TMG}}$ agents will be called 'TMG agents' because of the direct connection with the Thermal Minority Game. A variation of the Thermal Minority Game is now considered in which a concentration $q < q_0$ is chosen.
and the first factor sums over all agents whose highest-\(K\) is higher-ranked than \(G\). In other words, if \(j\) therefore have access to the highest-scoring strategy of agent \(i\), the second factor accounts for the contribution to an agent whose highest-scoring strategy is \(G\) and hence \(K < G\) where \(K\) is lower-ranked than \(G\) but that the highest-ranking strategy of agent \(J\) is \(K\)→\(n\) (57).

\[
\sum_{i=1}^{n} (d-1) \sum_{j=1}^{n} (d-1) = n^d - n
\]

The resulting expression is given by summing over all agents whose own highest-scoring strategy is \(K\)→\(n\) and who themselves have a highest-ranked strategy \(\hat{S}\) in a game where each agent holds \(S\) strategies (e.g. \(S = 2\)). We also suppose for the moment, that the ranking of strategies is unique (i.e. there are no strategies which are tied in virtual-points). Suppose that the highest-ranking strategy of agent \(i\) has higher \(K\) than \(i\) connected to another agent with an even higher-scoring strategy. The competition between these two effects will determine what then happens to the standard deviation in the presence of connections.

We can consider the following rule governing functionality of connections. Consider agent \(j\) who is connected to another agent holding strategy \(G\)→\(j\) (i.e. \(G < j\)). Hence these agents use the strategy ranked \(j\)→\(G\) in order to incorporate this effect. In particular, the number of agents \(n\) who are connected to an agent with strategy \(G\)→\(j\) is \(\sum_{i=1}^{n} K \rightarrow n\) and hence \(K < G\).

\[
\sum_{i=1}^{n} (d-1) \sum_{j=1}^{n} (d-1) = n^d - n
\]

In order to implement the renormalization of \(\theta\) we consider the appropriate modification of the strategy ranked \(j\)→\(G\) in the case of a network in which agent \(i\) is connected to another agent with strategy \(G\)→\(j\). Hence these agents use the strategy ranked \(j\)→\(G\) in order to incorporate this effect. In particular, the number of agents \(n\) who are connected to an agent with strategy \(G\)→\(j\) is \(\sum_{i=1}^{n} K \rightarrow n\) and hence \(K < G\).
In the limit of high connectivity $p$, there is a substantial change in the network B-A-R system and hence degeneracy of strategy scores. Accounting for the correct frequency of degenerate/nondegenerate timesteps yields excellent quantitative agreement with the numerical simulations \[44\].

These expressions can then be used to evaluate $\sigma_D$, which is the number of degenerate/nondegenerate timesteps for a given strategy in the absence of the network. As before, we will consider the case of the small-world limit with a flat strategy allocation matrix which, from Equation 23, gives a suitable expression $\sigma_D$, which is the number of degenerate/nondegenerate timesteps for a given strategy in the absence of the network. As before, we will consider the case of the small-world limit with a flat strategy allocation matrix which, from Equation 23, gives a suitable expression $\sigma_D$, which is the number of degenerate/nondegenerate timesteps for a given strategy in the absence of the network.

The population of agents therefore will contain, in general, a certain number of strategies at their disposal, thereby increasing the chances of uncovering a good strategy. It is this competition which interests us here. We consider the interesting case in which the B-A-R system has very few connections. This might correspond to a design problem in which initially there are no connections, and where the agents themselves are already designed and built. Hence possible control of the system is limited to introducing communication links between the agents. It is likely that such links will cost the designer something, hence a cost-benefit question arises as to the extent to which the system can be further modified or controlled by introducing a few connections into the system. In particular, how many links should be added, and how should they be introduced? Here we will focus on a simple model for the system which we will use as a proof-in-principle that the addition of connections can lead to a substantial reduction in the standard deviation of demand $\sigma_D$. In the limit of high connectivity $p$, there is a substantial change in the network B-A-R system and hence degeneracy of strategy scores. Accounting for the correct frequency of degenerate/nondegenerate timesteps yields excellent quantitative agreement with the numerical simulations \[44\].
VIII. CONCLUSION

We have given an in-depth presentation of the Crowd-Anticrowd theory in order to understand the fluctuations in competitive multi-agent systems, in particular those based on an underlying binary structure. We have shown that the partial contributions are given as follows:

\[ \sigma_1 \cdot (S' \cdot v_m)^{1-p} + \left[ (1-N) \right] = \text{min} \]

The expression can then be minimized with respect to \( S' \) under a constraint in an \( \mathcal{D} \) that represents the effective domain of the function. Note that we have overestimated the coordination within a dimer, in addition to the fact that the fluctuations in excess demand \( \text{min} \) is likely to be smaller than \( \text{min} \). Interestingly, such a result is larger than \( \text{min} \). Since the expression is smaller than \( \text{min} \), we have that the Crowd-Anticrowd concept might serve as a fundamental theoretical concept in spirit to the approximation used for the case of the alloy of mixed-memory agents in Section VI. Hence the contribution to the isolated monomer agents who only have the monomer agents forms a stochastic process.

\[ \sum_{i=1}^{N} \left[ \sigma_1 \cdot (S' \cdot v_m)^{1-p} + \left[ (1-N) \right] \right] \]

The partial contributions are given as follows:

\[ \cdot \left( \frac{1}{S' \cdot M - 1 + \alpha \cdot \beta} - 1 \right) + \left( \frac{1}{S' \cdot (1 - M) - 1} \right) \]

For the case of no network (i.e. \( S' = S' \)).
averaged theories. In particular, an observation of a real-world Complex System which is thought to resemble a multi-agent game, may correspond to a single run which evolves from a specific initial configuration of agents’ strategies. This implies a particular $\Psi$, and hence the time-averaging within the Crowd-AntiCrowd approach (47) can be carried out for that particular choice of $\Psi$. However this problem can still be cast in terms of the Crowd-AntiCrowd approach, since the time-averagings are then just carried out over some sub-set of paths in history space, which is conditional on the particular strategy configuration of the System. We also emphasize that a single $\Psi$ ‘macrostate’ corresponds to many possible ‘microstates’, where each microstate corresponds to one particular partition of strategy allocation among the agents. Hence the Crowd-AntiCrowd theory retained at the level of a given specified $\Psi$, is equally valid for the entire set of games which share this same ‘macrostate’ (47). [See Refs. [30, 31] for the simpler case of the Minority Game.]

We have been discussing a Complex System based on multi-agent dynamics, in which both deterministic and stochastic processes co-exist, and are indeed intertwined. Depending on the particular rules of the game, the stochastic element may be associated with any of five areas: (i) disorder in the fixed rules of the game, (ii) might typically be fixed from the outset (i.e., quenched disorder) hence it is interesting to see the interplay of (i) and (ii) in terms of the overall performance of the system. The extent to which these two ‘hard-wired’ disorders compensate for each other, as for example in the Parrondo effect or stochastic resonance, is an interesting question. Such a compensation effect might be engineered, for example, by altering the rules-of-the-game concerning inter-agent communication on the particular strategy that the agent follows (e.g. power level in the case of mechanical agents) hence breaking the tie. In this way, the coin-toss mimics the effect of this additional microscopic characteristic of the individual agents, which had been left out of the original game.

Finally, we make a more general comment concerning the relevance of the present work. Although at first sight, these Binary Agent Resource (B-A-R) models may sound esoteric with little relation to general problems and phenomena, they are in fact prototypes of socio-economic systems. As an example of this, the interplay of (i) and (ii) for a given level of the global resource $L$, and the global resource $L$ can emerge from these multi-agent numerical simulations. Just as important, the Crowd-AntiCrowd theory provides a method for assessing the effects of variants of the basic B-A-R setup without having to run the simulations themselves. In short, if the multi-agent simulations are seen as ‘experiments’, then the Crowd-AntiCrowd approach of this paper provides a complementary theory to describe the simulation outcomes. In short, the Crowd-AntiCrowd approach of this paper provides a complementary theory to describe the simulation outcomes.
We presented the Grand Canonical Minority Game (GCMG) model at the APFA1 conference (2000).


See http://www.unifr.ch/econophysics/minority


S. Gourley, C. Choe, P.M. Hui and N.F. Johnson (submitted for publication).

The Thermal Minority Game discussed in Ref. [38] depends on a parameter \( T \) (or equivalently \( 1/\beta \)) called a ‘temperature’. We could similarly define \( T \) by setting the probability of playing the worst strategy \( \theta = e^{-\beta} \). Hence \( T = 2[\ln(\theta^{-1} - 1)]^{-1} \). \( T = 0 \) corresponds to \( \theta = 0 \) while \( T \rightarrow \infty \) corresponds to \( \theta \rightarrow 1/2 \), hence we will only consider \( 0 \leq \theta \leq 1/2 \).


D. Smith and N.F. Johnson (in preparation).


[8] See the works of D. Wolpert and K. Tumer, at http://ic.arc.nasa.gov


[21] R. Savit, R. Manuca and R. Riolo, Phys. Rev. Lett. 82, 2233 (1999). Interestingly, we had found a similarly shaped curve for the volatility in the El Farol problem (see Ref. [10]).


[29] We presented the Grand Canonical Minority Game (GCMG) model at the APFA1 conference in Dublin (1999) and at the first International Workshop of Applications of Statistical Physics to Financial Analysis in Palermo (September 1998). See also Ref. [33].


[42] The Thermal Minority Game discussed in Ref. [36] depends on a parameter \( T \) (or equivalently \( 1/\beta \)) called a ‘temperature’. We could similarly define \( T \) by setting the probability of playing the worst strategy \( \theta = e^{-\beta} \). Hence \( T = 2[\ln(\theta^{-1} - 1)]^{-1} \). \( T = 0 \) corresponds to \( \theta = 0 \) while \( T \rightarrow \infty \) corresponds to \( \theta \rightarrow 1/2 \), hence we will only consider \( 0 \leq \theta \leq 1/2 \).


FIG. 1: Schematic representation of B-A-R (Binary Agent Resource) system. At timestep $t$, each agent decides between action $-1$ and action $+1$ based on the predictions of the strategies that he possesses. A total of $n-1$ agents choose $-1$, and $n+1$ choose $+1$. In the simplified case that each agent's confidence threshold for entry into the game is very small, then $n-1[t]+n+1[t]=N$ (i.e. all agents play at every timestep). Agents may be subject to some underlying network structure which may be static. The strategy space, which may be equated to a hypercube, shows all possible strategies of length $m$.

FIG. 2: Strategy Space for $m=2$, together with some example strategies (left). The strategy space shown is known as the Full Strategy Space (FSS), and contains all possible permutations of the actions $-1$ and $+1$. The strategy space is shown as the red shaded area in the RSS. The shaded strategies form a Reduced Strategy Space (RSS). There are $2^2m$ strategies in the RSS. The red shaded line connects two strategies with a Hamming distance separation of 4.

FIG. 3: Schematic representation of B-A-R. The agent remembers the previous action (red). Agent memory $m = 2$.
FIG. 3: History Space. Examples of the de Bruijn graph for $m = 1$, $2$, and $3$. Red transitions correspond to the most recent global outcome $0$. Blue transitions correspond to the most recent global outcome $1$.

FIG. 4: Example distribution for the tensor $\Omega$ describing the strategy allocation for $N = 101$ agents in the case of $m = 2$ and $S = 2$. Between each transition to the next strategy, the strategy allocation for $N = 101$ is determined. Service times $	au$ and $S$ are fixed and are equal to $1$ and $2$, respectively.
FIG. 5: Schematic diagram of a fairly typical variation in strategy scores, as a function of time, for a competitive game. This behavior is particularly relevant for the low regime where there are many more agents than strategies, and hence strategy rankings change in time due to being overplayed. The strategies at any given timestep can be ranked in terms of virtual-point ranking $K$, with $K = 1$ being the highest-scoring and $K = m$ the lowest-scoring. The actual identity of the strategy in rank $K$ changes as time progresses, as can be seen. Ignoring accidental ties in score, there is a well-defined ranking of strategies at each timestep in terms of their $K$ values.

FIG. 6: Schematic representation of the strategy allocation matrix $\Psi$ with $m = 2$ and $s = 2$, in the RSS. The strategies are ranked according to strategy score, and are labelled by the rank $K$. In the limit that $\Psi$ is essentially flat, then the number of agents playing the $K$th highest-scoring strategy is just proportional to the number of shaded bins at that $K$. In the limit that $\Psi$ is essentially flat, then the number of agents playing the $K$th highest-scoring strategy is just proportional to the number of shaded bins at that $K$.
FIG. 7: Crowd-Anticrowd theory vs. numerical simulation results for Minority Game as a function of memory size $m$, for $N=10^1$ agents, at $S=2, 4, 8$. At each $S$ value, analytic forms of standard deviation in excess demand $D^t$, are shown corresponding to $\sigma_{\delta f}$ (upper solid line), $\sigma_{f}$ (lower dashed line) and $\sigma_{f, high}$ (monotonically-increasing solid line which is independent of $S$). The numerical values were obtained from different simulation runs (triangles, crosses and circles).

FIG. 8: Crowd-Anticrowd theory vs. numerical simulation results for Thermal Minority Game as a function of stochastic probability $\theta$, or 'temperature'. The analytic results (lines) correspond to the $\sigma_{\delta f}$ (solid upper line) and $\sigma_{f}$ (solid lower line) limiting-case approximations.

---

$\text{Minority Game}$

$\text{Thermal}$
FIG. 9: Schematic representation of B-A-R (Binary Agent Resource) network system in limit of few connections.