The Simplex Game: Can Selfish Users Learn to Operate Efficiently in Wireless Networks?

Panayotis Mertikopoulos
Dept. of Physics, University of Athens
157 84 Athens, Greece
pmertik@phys.uoa.gr

Aris L. Moustakas
Dept. of Physics, University of Athens
157 84 Athens, Greece
arislm@phys.uoa.gr

ABSTRACT

We introduce and analyse the simplex game, a non–cooperative game between selfish heterogeneous players with bounded rationality that compete for limited resources. In this game, players are asked to place their bet among a set of B choices and the game rewards those in the minority. Players start out completely uneducated and naive but, through a selfish learning scheme that seeks to maximise their own gain, they become more experienced and quickly learn to adapt and perform with an unexpected efficiency. Employing methods of Statistical Physics (namely the theory of replicas) we establish explicit analytic estimates of the game's performance that clearly reflect the users' emergent efficiency. We further map the general simplex game to the minority game, a simple model introduced in the context of econophysics. This mapping allows us to study the effect that the number of choices has on the game's performance. For concreteness, our analysis has focused on a system of WLAN access points, but it can be customised to other networks with non-cooperative players, such as OFDMA.

Keywords

Evolutionary Games, Nash Equilibria, Minority Game, Statistical Physics

1. INTRODUCTION

A common feature in telecommunications networks is the existence of entities competing for the usage of limited resources (e.g. wireless users competing for power and bandwidth). This phenomenon is bound to become even more prominent in the context of wireless networks, where ad-hoc and deregulated networks have already appeared. As a result, a substantial amount of research has recently emerged, applying the paradigm of non-cooperative games to various aspects of networks.

One direction pursued is in the context of uncoordinated radio access in a common radio channel. For example, the slotted and unslotted ALOHA protocols have been analysed by optimizing their transmission probabilities [11] or their power control [2]. Yet another application is in the realm of CDMA systems, where game-theoretic techniques have been used in order to choose the optimal transmitting power [3, 13] and also the optimal carrier in the multi-carrier CDMA setting [13]. Finally, game theory has also been applied in finding strategies to relay packets in ad hoc networks [15].

In all of the above, the analysis has been focused on finding a stable operating point for the network, assuming that the network is static and the users are identical. Indeed, since all entities are assumed identical, they can all be programmed to react in the same way to external stimuli and thus, after a certain number of game iterations, the network may converge to its predetermined symmetric socially optimal state. As a concrete example, in the case of slotted ALOHA [11], one may calculate the probability of transmission corresponding to the Nash equilibrium, given the number of users that are waiting to transmit. This amounts to giving a single (mixed) strategy to each player which, assuming everyone plays it, is guaranteed to take the network to a socially optimal state. Therefore, one can *pre-program* users to play a single strategy. In fact, this is similar with previous optimisation approaches; for example, in OFDMA, users are equipped with a random (or quasi-random) frequency hopping pattern, which is designed to minimise the multi-user interference to everyone.

Although this is convenient in a setting of identical preprogrammed users, it is not clear if this approach will work properly in a more general context. For one, it is known that more efficient non-symmetric equilibria do exist, even in the simplest situation of identical players [11]. Moreover, even in the existing centrally controlled networks, there exist different classes of users, each with vastly different quality of service criteria, and each opting for different strategies to maximise their utility functions. To make things worse, these differences in strategies can be exacerbated when the external conditions (and, hence, the state of the network itself) are not constant. For example, data users would be willing to wait longer times compared to voice users, in exchange for higher rates when the channel quality is better and/or the interference is smaller. As a result, in a deregulated, non-centrally controlled wireless network, heterogenous players entering the game may play along different strategies to maximise their utility function. It is therefore important for them to be able to have to choose among *several* pre-programmed strategies, in order to be able to adapt to the diversity of the environment they may face.

In this paper we will relax these assumptions, working

GameComm '07, October 22, 2007, Nantes, France

with a “worst case scenario”: a multitude of different heterogeneous users, constantly learning how to maximise their own payoff but only with a limited ability to process information. Specifically, we will assume that each player has a set of strategies, randomly drawn at the outset of the game and they are to use these strategies in order to adapt to the network setting and maximise their utility. Also, the players are endowed with a limited-complexity, inductive reasoning scheme [1], which is what allows them to learn along the way: thanks to a centrally transmitted training signal (for example, broadcasting the best performing access point in the previous instance of the game), the users use this scheme to rank their strategies and use the one that has the highest potential for personal gain.

Needless to say, this approach is general in scope but, for the sake of concreteness, we will focus at a specific example, namely the situation of several WLAN access points and a multitude of users trying to connect to the one with the least interference level. In section 2 we will define the details of the model and discuss its similarities with other situations and, as a metric of its performance, we will analytically calculate the variance of the steady state solution of the iterated game in the case of users having B different choices to make (B access points). This variance will actually turn out to be a Lyapunov function for the evolutionary dynamics of the game.

To make analytic progress, we will borrow analytic tools from statistical physics (namely the replica approach) in order to average over the dynamics of the game, all the while keeping the underlying randomness of the strategy space fixed; this approach has already been successfully applied to communication theory in e.g. [16, 17]. The remarkable outcome of this analysis is that the global efficiency is dramatically better compared to the case when users make uneducated (effectively random) choices and, in fact, rivals that of centrally controlled systems. This then is a clear indication that, despite the apparent chaotic set-up of the problem, players are indeed able to learn to quickly adapt to each others’ actions and converge to a socially optimal state.

Our model is essentially an application of the minority game of [12], a simple model developed in the context of econophycics where players are trained to buy or sell and are rewarded when they are in the minority. There, the training and the corresponding valuation of their randomly chosen strategies is based on a finite number of previous outcomes of the repeated game. However, as was shown in [4], the precise nature of the feedback information is not important; what matters is only the actual amount of information being transmitted to the players. Therefore, it is more natural to interpret this “feedback” as a training stimulus that allows users to correlate their strategies to the actions of other users.

Other similar approaches of allowing the players to learn have recently appeared in the literature, specifically in the context of wireless sensor networks. In [8, 9] the authors analyse a variant of a Kaufman network. However, they indicate that the fluctuations of the game scale faster than the random game. Also, in [10] coordination between sensors appears if they keep track of the previous outcomes of the repeated game, which is a variant of the Gur game.

In this paper, we adapt the simplex game to the general context of wireless networks by allowing the players to have $B \geq 2$ choices (nodes) to connect to. Our main analytic result is theorem 3.2: *A simplex game with B choices is macroscopically equivalent to a minority game with a retuned training parameter.* This virtual reduction of choices has a number of important consequences. For example, one can use it to predict the effect that the addition of an access point will have to the performance of the network, a question which turns out to be quite complicated and does not admit a trivial answer. But, more importantly, this result, together with the rest of our analysis, can be viewed as a first step towards establishing a dictionary with which to translate the considerable literature of learning games and, more specifically, the minority game to the language of networks.

1.1 Outline

Our presentation is organised as follows: in section 2 we offer the system model that motivates our discussion and introduce the simplex game with B choices. The definition of the game takes place in sections 2.1 and 2.2 while, in section 2.3, we calculate some benchmarks for its performance.

In section 3, we specify the iteration mechanism with which the players actually learn and become more experienced in the simplex game (section 3.1) and we also analyse the evolutionary steady states of the game (section 3.2) to obtain exact analytic measures for its performance and efficiency. In section 4 we discuss the effect that the addition of a node has to the performance of the network.

The proof of theorem 3.2 is presented in appendices A and B; also, for completeness, we present the basic steps of the replica symmetric calculation in appendix C.

1.2 Notational conventions

For two vectors v, w of the same dimension, we will denote their Euclidean product by: $v \cdot w \equiv \sum_i v_i w_i$. Also, we will denote the n-dimensional simplex by Δ_n and its standard embedding will be in $\mathbb{R}^{n+1}$ via:

$$\Delta_n = \{x \in \mathbb{R}^{n+1} : 0 \leq x_i \leq 1 \, \forall i \text{ and } \sum_i x_i = 1\}.$$

2. SYSTEM MODEL

Our principal motivation is a very simple model for a system of B WLAN access points (nodes) and N users able to connect to any of those B nodes. If N_r (with $0 \leq N_r \leq N$) users decide to connect to node $r = 1, \ldots B$, we then define the capacity of that node to be equal to

$$C_r = \frac{c_0}{N_r} \tag{2.1}$$

where c_0 is a constant. Note that this quantity diverges for $N_r = 0$; however, this is an unrealistic instance of the model, and it will not appear for a large enough number of users. In practice, this throughput can be thought of being achieved through round-robin scheduling or by allocating a certain number of carriers to each user. However, the specifics of this model are not essential; the crux of the matter is simply the fact that the capacity of each node decreases with N_r, i.e. users that are in a relative minority (small N_r) are rewarded with better quality of service.

This model can also be encountered in a different context, namely in the case of an uplink multicarrier CDMA. If we assume perfect power-control for each user, then the SINR for the users connected with the carrier $r \in \{1, \ldots B\}$ will be

$$\mathrm{SINR}_r = \frac{s}{\gamma + N_r} \tag{2.2}$$

where s is the signal to interference ratio and γ is the noise to interference ratio. In this paper, we will be interested in this model for large numbers of users N and we will therefore make the important assumption that the variations in N_r is small compared to its mean N/B. Therefore, we may linearise the above equations to the following:

$$C_r \approx \frac{Bc_0}{N} - \frac{B^2 c_0}{N^2}\delta N_r \tag{2.3}$$

where $\delta N_r = N_r - N/B$ is the variation of N_r from its mean. Even though this is an approximation, we will see that it has virtually no impact on the behaviour of the game (see equation 3.2) thanks to the fact that the state with a node having unreasonably low numbers of users connected to it is inherently unstable.

Now, on an informal level, our game-theoretic analogue of this model is the following: N selfish players wish to place a bet among B choices in an effort to maximise their capacity (equation 2.1 or equation 2.3). To facilitate them in this decision, public information is being transmitted to all (e.g. the last performance ranking of the access points in our system model, the current needs of the network, etc.). This information will be described by a random variable m of uniform distribution[1], which the players process with the aid of S *strategies* (decisional schemes). These strategies are selected with probability p_{is} = Prob(player i selects strategy s) and, according to their performance, the players will readjust these probabilities as they learn which strategies are most effective.

2.1 Introducing the simplex game

This setting is an immediate generalization of the original minority game which consisted of a large number N of heterogeneous agents that were asked to make a binary choice (e.g. "buy" vs. "sell"), based on public information and the behavioural schemes that the agents had developed over time [5].

Obviously then, the first step in order to adapt this game to our system model would be to expand the "choice set" $\{-1, 1\}$ of our agents, in order to accommodate for the B access points they may choose from. To model a game with B choices, it is convenient to introduce the *characteristic B-states*, which are B vectors of unit norm $\{q_r\}_{r=1}^B \subseteq \mathbb{R}^{B-1}$ that constitute the vertices of a $(B-1)$-simplex[2] centred at the origin (more details on the properties of these vectors can be found in appendix A). These characteristic states are actually the namesake of the *simplex game* wherein players will be placing their bets among the vertices of a $(B-1)$-simplex:

Definition 2.1. *A* **simplex game** *with B choices is a collection $\mathscr{G}$ of the following data:*

- *The set of* players*: $\mathscr{N} = \{1, 2 \ldots N\}$.*
- *The set of B* choices *or* bets*: $\mathscr{B} = \{q_r\}_{r=1}^B$ (where the q_r are characteristic B-states).*
- *The* training *set: $\mathscr{M} = \{1, \ldots M\}$; we will also refer to the ratio $\alpha = \frac{M}{N}$ as the* **training parameter** *of the game $\mathscr{G}$.*
- *The set of* strategy labels*: $\mathscr{S} = \{1, \ldots S\}$.*

Finally, we consider the sample space $\Omega = \mathscr{S}^N \times \mathscr{M}$ endowed with the probability measure P*:*

$$\mathrm{P}(s_1 \ldots s_n, m) = \frac{1}{M} p_{1s_1} \ldots p_{Ns_N} \tag{2.4}$$

where, for all i, $0 \le p_{is} \le 1$ and $\sum_{s\in\mathscr{S}} p_{is} = 1$ (i.e. $\mathrm{P}_i(s) = p_{is}$ is itself a probability measure on $\mathscr{S}$). We will then refer to an event $\omega \in \Omega$ as an **instance of** $\mathscr{G}$*.*

To be more precise, a *strategy* should convert the training signal to a betting suggestion (i.e. it should be defined as a map $\mathscr{M} \to \mathscr{B}$) and each player $i \in \mathscr{N}$ would have S of them. Formally, this can be efficiently encoded in a *strategy matrix* $c : \mathscr{N} \times \mathscr{S} \times \mathscr{M} \to \mathscr{B}$ so that, given a signal $m \in \mathscr{M}$, $c_{is}^m \equiv c(i, s, m) \in \mathscr{B}$ is just the choice of player i if he employs his s^{th} strategy. Clearly, there are $C = B^{NSM}$ different strategy matrices and we will denote the set of all strategy matrices by $\mathscr{C}$.

2.2 Playing the simplex game

We are now in a position to study the players' bids and their respective payoffs:

Definition 2.2. *Let $(\mathscr{G}, c)$ be a B-choice simplex game whose players have been equipped with the strategy matrix $c \in \mathscr{C}$, and let $\omega = (s_1 \ldots s_N, m) \in \Omega$ be an instance of $\mathscr{G}$:*

1. *The* bid *of player i is the random variable $b_i : \Omega \to \mathbb{R}^{B-1}$ where $b_i(\omega) = c(i, s_i, m) \equiv c_{is_i}^m$. Similarly, the* aggregate bid *is defined as: $b = \sum_{i\in\mathscr{N}} b_i$.*
2. *The* payoff *that player i receives is the random variable $u_i : \Omega \to \mathbb{R}$ defined by: $u_i = -b_i \cdot b$, i.e. $u_i(\omega) = -b_i(\omega) \cdot \sum_{j\in\mathscr{N}} b_j(\omega)$*

Clearly then, we can see that the aggregate bid is the quantity of interest in the simplex game, since it succinctly encodes the actual state of the game, i.e. how many players made a particular choice. In fact, given the aggregate bid b at an instance ω of the game, some linear algebra combined with the properties of the characteristic states (lemma A.1) will reveal that the number N_r of players whose bet was q_r is just:

$$N_r = \frac{N}{B} + \frac{B-1}{B} q_r \cdot b \tag{2.5}$$

Therefore, in order to study the players' distribution, it suffices to study the statistical behaviour of the aggregate bid b (note also that the second term in the above equation is just δN_r, as defined in equation 2.3).

Moreover, the aggregate bid also determines the players' payoff, defined so as to reward those that took the "path less trodden". One way to see this would be to calculate the total payoff: $u = -\sum_{i\in\mathscr{N}} u_i = -\sum_{i\in\mathscr{N}} b_i \cdot \sum_{j\in\mathscr{N}} b_j = -b^2 \le 0$; i.e. minorities are rewarded while majorities are penalised. Alternatively, a little algebra shows that the payoff to the N_r players that chose q_r is $\frac{N - BN_r}{B-1}$: this is positive if less

[1] Without loss of generality, we may assume that there is no *a priori* betting bias.

[2] Recall that a n-simplex is the generalization of an equilateral triangle (2-simplex) to higher dimensions: regular tetrahedron (3-simplex), pentatope (4-simplex), etc. Already, for $B = 2$, we recover the choice set $\{-1, 1\}$ of the original minority game.

than the average number of players $\frac{N}{B}$ chose q_r (minority reward) and negative otherwise (majority penalty).

Now, as far as our choice of payoff function is concerned, our motivation is twofold: to begin with, this "Euclidean" payoff is the direct generalization of the original minority payoff used in [12], and this will be very important to us later on when we perform the theoretical analysis of the simplex game. More importantly however, thanks to equation 2.5, this payoff function is (up to a normalisation constant) precisely the linear approximation to the users' capacity in our system model (equation 2.3) and, for this reason, this payoff will be dubbed *linear*. We have already argued that this approximation is well justified for small variations δN_r and this can be clearly seen in figure 1 where the game is iterated according to the learning scheme of section 3. Not surprisingly, the game's variance exhibits the same steady state and convergence properties, i.e. players learn exactly the same things, whether they use the capacity payoff or its linear approximation.

2.3 Vital Statistics of the simplex game

We will devote the rest of this section to some statistical calculations that will serve as our benchmarks when we iterate the game. To begin with, even though the SG has already been described as a stochastic process, there is still some inherent randomness in the choice of the strategy matrix $c \in \mathscr{C}$. Now, since there are no a priori *good* strategies[3], we will assume that c is uniformly distributed in $\mathscr{C}$; this is equivalent to assuming that, given a training signal $m \in \mathscr{M}$, the probability of the s^{th} strategy of player i to dictate choice $q_r \in \mathscr{B}$ is: $\text{Prob}(c_{is}^m = q_r) = \frac{1}{B}$.

With this ansatz, we get the following array of expectations:

$$\mathbf{E}(b|c,m) = \sum_{\mathbf{s}} p_{1s_1} \cdots p_{Ns_N} \sum_i c_{is_i}^m = \sum_i \sum_s p_{is} c_{is}^m$$

and, thus:

$$\mathbf{E}(b) = \frac{1}{MC} \sum_m \sum_i \sum_s p_{is} \cdot \sum_{c \in \mathscr{C}} c_{is}^m = 0$$

(since $\sum_r q_r = 0$).

Furthermore, with regards to the variance:

$$\begin{aligned}
\mathbf{E}(b^2) &= \mathbf{E}\left(\sum_i b_i^2 + \sum_{i \neq j} b_i \cdot b_j\right) \\
&= N + \frac{1}{MC} \sum_m \sum_{i \neq j} \sum_{s,s'} p_{is} p_{js'} \sum_{c \in \mathscr{C}} c_{is}^m \cdot c_{js'}^m \\
&= N \qquad (2.6)
\end{aligned}$$

because the random variable $c_{is}^m \cdot c_{js'}^m$ only assumes the values 1 (with probability $\frac{1}{B}$) and $-\frac{1}{B-1}$ (with probability $\frac{B-1}{B}$).

Thus, the (intensive) variance of the aggregate bid in this stochastic process is $\frac{\sigma^2(b)}{N} = 1$ and we will be referring to this benchmark as the *fully random* case.

This benchmark is useful by itself but, instead of averaging out over all possible strategy matrices $c \in \mathscr{C}$, it is better to employ the ergodicity assumption and consider expectations given a strategy matrix $c \in \mathscr{C}$ which, typically, has $\sum_m c_{is}^m = 0$ (i.e. is such that $\mathbf{E}(b|c) = 0$) [4]. Given such a typical strategy, and to leading order N, the variance becomes:

$$\begin{aligned}
\sigma^2(b|c) &= \mathbf{E}(b^2|c) \\
&= N + \frac{1}{M} \sum_m \sum_{i \neq j} \sum_{s,s' \in \mathscr{S}} p_{is} p_{is'} \, c_{is}^m \cdot c_{js'}^m \\
&\underset{N}{\sim} N + \frac{1}{M} \sum_{m \in \mathscr{M}} \mathbf{E}(b|c,m)^2 - NG(\mathbf{p}) \qquad (2.7)
\end{aligned}$$

where $G(\mathbf{p}) = \frac{1}{N} \sum_i \sum_s p_{is}^2$ measures the self-overlap of strategies and "$\underset{x}{\sim}$" is the asymptotic notation: $f(x) \underset{x}{\sim} g(x)$ if $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1$.

In contrast to the fully random case, the variance in this stochastic process depends nontrivially on the probability measures p_{is} that evolve over time; thus, in the next section we will be interested in the minima of this quantity, since they correspond to the optima of the game's performance.

3. EVOLUTION IN THE SIMPLEX GAME

The main strength of inductive reasoning [1] is that players become more sophisticated over time because their decisional schemes evolve. In our setting, this means that, as the game is played over and over, the probability measures p_{is} are updated in a way that reflects the satisfaction of player i with his s^{th} strategy. Through this evolutionary procedure, players become increasingly sophisticated, and it has been shown in [12] that the game's dynamics converge to a *steady state* which is actually a Nash equilibrium.

3.1 Iterating the simplex game

The key feature of this iteration procedure is that players reward strategies that pay well and penalise them otherwise. However, in order to measure the performance of a strategy, the game has to be actually played and, for that reason, the reward of a strategy will depend on the actions that *every* player took. In fact, this is how players come to interact with one another: the players' overall actions indirectly affect the way any individual player thinks and, through that process, players eventually learn how others think and become more sophisticated.

To make this idea precise, let $(\mathscr{G}, c)$ be a B-choice simplex game with strategy matrix c, and consider an instance $\omega = (s_1 \dots s_N, m) \in \Omega$ of $\mathscr{G}$. In this setting, a function $W_{is} : \Omega \to \mathbb{R}$ will be called a *reward for the s^{th} strategy of player i*, if $W_{is}(\omega) \geq 0$ precisely when the payoff u_i has: $u_i(s, s_{-i}, m) \geq 0$ [5]. In other words, if employing strategy s in a particular instance ω of the game would have yielded a positive payoff $u_i(s, s_{-i}, m) \geq 0$ to player i, this strategy receives a positive reward $W_{is}(\omega) \geq 0$. Then, if $\omega(t) \in \Omega$, $t \in \mathbb{N}$ is a sequence of instances of the game, we may recursively define the *score* $U_{is}(t)$ of the strategy (i,s) at time t by:

$$U_{is}(t+1) = U_{is}(t) + W_{is}(\omega(t)) \qquad (3.1)$$

[3]Indeed, if there were a strategy that performed consistently better than others, players would invariably follow it and lose because of the minority payoff.

[4]This is a typical strategy because it has no betting bias: as m runs through $\mathscr{M}$, c_{is}^m runs through all values of $\mathscr{B}$. Since $\sum_r q_r = 0$, it is easy to see by the central limit theorem that the matrices of $\mathscr{C}$ are sharply concentrated around strategies of this sort.

[5]This notation is the standard game-theoretic shorthand:

$$(s, s_{-i}, m) \equiv (s_1 \dots s_{i-1}, s, s_{i+1} \dots s_N, m).$$

This scoring function is what will keep track of the performance of a strategy as the game is iteratively played (as described by the sequence $\omega(t)$) and, clearly, the most natural choice W_{is} for a strategy's reward is the payoff itself:

$$\begin{aligned} W_{is}(\omega) &= \frac{1}{M} u_i(s, s_{-i}, m) \\ &= -\frac{1}{M} c_{is}^m \cdot \left[b(\omega) + \left(c_{is}^m - c_{is_i}^m \right) \right] \end{aligned} \quad (3.2)$$

Here, the rescaling factor $\frac{1}{M}$ has been introduced because players evolve at a slower pace than the game's iterations (on an intuitive level, players would have to compare their strategies against all M possible training signals).

Now, we are finally ready to incorporate a strategy's score into the probability with which it is selected. To that end, we will adhere to the model of *exponential learning* according to which, every player has a learning rate Γ_i [6] and adjusts his probability measure p_{is} based on the score of his strategies:

$$p_{is}(t) = \frac{e^{\Gamma_i U_{is}(t)}}{\sum_{s' \in \mathscr{S}} e^{\Gamma_i U_{is'}(t)}} \quad (3.3)$$

This "exponential learning" scheme can be seen to be equivalent to the standard replicator dynamics that have been studied extensively in game theory (see e.g. [12] for a discussion that can be immediately generalised to our case). Therefore, the fixed points of the dynamical system 3.3 will actually be evolutionary stable Nash equilibria of the game. Hence, once the iteration sequence converges to a specific stage, the game will be characterised by minimal volatility and the users will have no incentive to leave, on account of the stage being a Nash equilibrium.

Of course, to make precise sense of the above, we need to define rigourously this iteration procedure; to that end, we have:

Definition 3.1. *Let $(\mathscr{G}, c)$ be a B-choice simplex game with strategy matrix $c \in \mathscr{C}$, and let ω be an instance of $\mathscr{G}$. Then, an* iteration sequence for $\mathscr{G}$ and ω *is a sequence $(\mathscr{G}(t), \omega(t))$ where $\mathscr{G}(t)$ is a B-choice simplex game with the same underlying set structure as $\mathscr{G}(t-1)$, but with a probability measure* P_t *determined by equation 3.3 and $\omega(t)$ is an instance of $\mathscr{G}(t)$, drawn according to* P_t.

At this point, it is important to state explicitly that the choice of strategy matrix $c \in \mathscr{C}$ does not change throughout an iteration sequence of a particular simplex game: even though the players continually rate their decision schemes and either abandon or embrace them, they do not develop new ones as time passes. In fact, this is the key difference between the variances in equations 2.6 and 2.7 which describe the fully random case and the simplex game, respectively (see also figure 1).

3.2 Evolutionary Steady States

Essentially, equation 3.3 describes a discrete dynamical system; as such, it has been studied extensively in [12] and [6][7]. The main result there is that the probability measures $\mathbf{p} = \{p_{is}\}$ converge to a *steady state* which is actually a *Nash equilibrium* of the game (i.e. a player has no incentive to change his way of thinking if other players stick to theirs). In figure 1 we have performed 64 iterations of a simplex game with $B = 5$, $S = 2$ and training parameter $\alpha = \frac{1}{4}$, following the training scheme outlined above with both the linear payoff of definition 2.2 and the payoff $u_i \propto \frac{1}{N_r} - \frac{1}{N/B}$ entailed by equation 2.1. This performance is then compared to that of the fully random case; in this way, we readily observe the convergence to a socially optimal steady state. Indeed, within approximately M steps of the iteration sequence, the game has practically converged to a steady state whose global efficiency is dramatically better than that of the fully random case. In other words, players (albeit selfish) do learn to operate efficiently, despite their totally incompatible needs and the complete lack of central control.

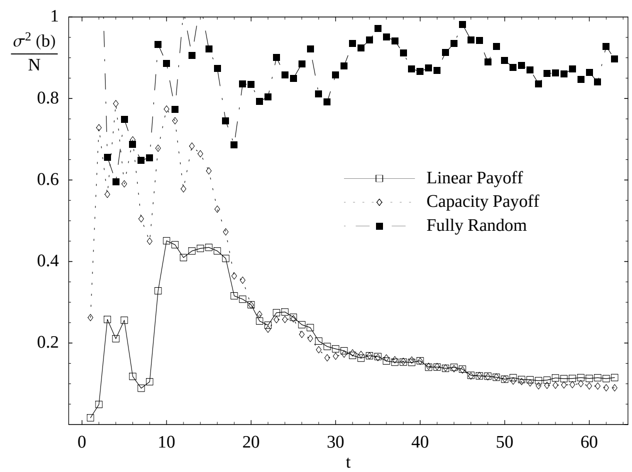

Figure 1: The variance of the aggregate bid over 64 iterations of the simplex game with linear payoff(equation 2.3; empty boxes) and capacity payoff (equation 2.1; diamonds) versus the fully random case (full boxes). The parameter values chosen here are $B = 5$, $S = 2$ and $\alpha = 1/4$, with a learning rate of $\Gamma_i = 100$.

A detailed discussion of these issues can be found e.g. in [12]; for our purposes, what is most interesting is that these steady states can be recovered as the minima of a certain quadratic form. Indeed, it is shown in [6] that the evolution dynamics for p_{is} of equation 3.3 admit a *Lyapunov function*, i.e. a function $L(\mathbf{p}, c)$ that is minimised along the system's trajectories; hence, its minima will correspond to the steady states of equation 3.3. Amazingly enough, this Lyapunov function turns out to be (up to an irrelevant constant) just the variance (equation 2.7) of the aggregate bid for a given

[6]The learning rate only affects the rate of convergence to the steady state, not the steady state itself [12]; for this reason we will ignore its role in our discussion concerning the actual steady states of the game. However, we note that in the "hard" case $\Gamma \to \infty$, these dynamics simply reduce to the well-understood "best response" scheme.

[7]Actually, their analysis was done for the $B = 2$ case but the scalar product generalization that we use for the payoff function makes their results immediately transferable to our setting.

strategy matrix c:

$$L(\mathbf{p}, c) = \sigma^2(b|c) - N \underset{N}{\sim} \frac{1}{M} \sum_{m \in \mathscr{M}} \mathbf{E}(b|c,m)^2 - NG(\mathbf{p}) \qquad (3.4)$$

where, as before: $G(\mathbf{p}) = \frac{1}{N} \sum_i \sum_s p_{is}^2$.

Thus, we will be interested in the quantity $L_0(\mathscr{G}, c) = \min_{\mathbf{p}}\{L(\mathbf{p}, c)\}$, i.e. the minimum of L with respect to p_{is}. In appendices B and C, we will see that the Lyapunov function L can be interpreted as the energy of a certain physical spin system; motivated by this analogy, we will refer to $L_0(\mathscr{G}, c)$ as the *minimum energy* of the game $(\mathscr{G}, c)$. Then, in order to make further progress, we introduce the so-called *partition function*:

$$\mathscr{Z}(\beta, c) = \int_{\mathscr{D}} e^{-\beta L(\mathbf{p},c)} \, \mathrm{d}p$$

from which we can harvest $L_0(\mathscr{G}, c)$ using corollary B.2 of appendix B as follows:

$$L_0(\mathscr{G}, c) = -\lim_{\beta \to \infty} \frac{1}{\beta} \log \mathscr{Z}(\beta, c). \qquad (3.5)$$

We now need to make the following important assumption

Assumption 1 (Self-Averaging Property). *In the limit $N \to \infty$, the quantity $L_0(\mathscr{G}, c)$ is asymptotically equal to its average over all strategy matrices $c \in \mathscr{C}$. Thus,* in high probability*:*

$$L_0(\mathscr{G}, c) = \left\langle L_0(\mathscr{G}, c') \right\rangle_{c' \in \mathscr{C}} \equiv L_0(\mathscr{G}).$$

To evaluate the average of the logarithm in equation 3.5 we will rely on the following assumption, as discussed for example in [14]:

Assumption 2 (Replica Continuity). *The average $\langle \mathscr{Z}^n(\beta, c)\rangle$ evaluated for $n \in \mathbb{N}$ can be analytically extended for real n in the vicinity of 0^+. Specifically,*

$$\langle \log \mathscr{Z}(\beta, c)\rangle = \left.\frac{\partial}{\partial n} \langle \mathscr{Z}^n(\beta, c)\rangle\right|_{n=0}$$

As a result we have

$$L_0(\mathscr{G}) = -\lim_{\beta \to \infty} \frac{1}{\beta} \frac{\partial}{\partial n} \left.\langle \mathscr{Z}^n(\beta, c)\rangle\right|_{n=0}$$

Indeed, not only will $L_0(\mathscr{G})$ yield the steady states for the iteration sequence of a simplex game, but, perhaps more importantly, it will also give a lower bound for the game's *volatility* $\sigma^2(b)$ (in fact, this lower bound will turn out to be an excellent estimate for the volatility itself).

We can finally make a precise statement of our main result on the reduction of multi-choice games to binary ones. Indeed, if $\mathscr{G}$ is a simplex game with B choices, define its *binary reduction* $\mathscr{G}_{\mathit{eff}}$ to be the simplex game of 2 choices which, except for an enlarged training set $\mathscr{M}_{\mathit{eff}}$ that has $M_{\mathit{eff}} = M(B-1)$ (and, therefore, an effective training parameter of $\alpha_{\mathit{eff}} = \frac{M_{\mathit{eff}}}{N} = \alpha(B-1)$), otherwise consists of the same data as $\mathscr{G}$. We then obtain:

Theorem 3.2 (Reduction of Choices). *The minimum energy of a B-choice simplex game $\mathscr{G}$ is asymptotically equal to the minimum energy of its binary reduction $\mathscr{G}_{\mathit{eff}}$, i.e. $L_0(\mathscr{G}) \sim L_0(\mathscr{G}_{\mathit{eff}})$ as $N \to \infty$. Thus, the macroscopic quantities (volatility etc.) of a simplex game with B choices can be obtained from an associated minority game with a retuned training parameter: $\alpha_{\mathit{eff}} = \alpha(B-1)$.*

We will leave the details of the proof to appendix B and, additionally, we will calculate the analytic expression for $L_0(\mathscr{G})$ in appendix C, following [12] and resorting to the assumption of so-called *replica symmetry*. Within that framework, we obtain the following approximation:

$$\frac{\sigma^2(b)}{N} = 1 + \frac{1}{N} L_0(\mathscr{G}) = \Theta(\alpha - \alpha_c) \left(1 - \sqrt{\frac{\alpha_c}{\alpha}}\right)^2 \qquad (3.6)$$

where Θ is the Heaviside step function, and $\alpha_c \equiv \alpha_c(S, B)$ is the critical value of the training parameter beyond which the variance vanishes within the replica symmetric approximation.

Now, even though equation 3.6 only provides us with an approximation for the variance $\sigma^2(b)$, it turns out that this is actually a formidable measure of the game's actual performance. This agreement is readily seen in figure 2, which depicts the variance of a simplex game with $B = 5$ choices and $S = 2$ strategies per player in its steady state.

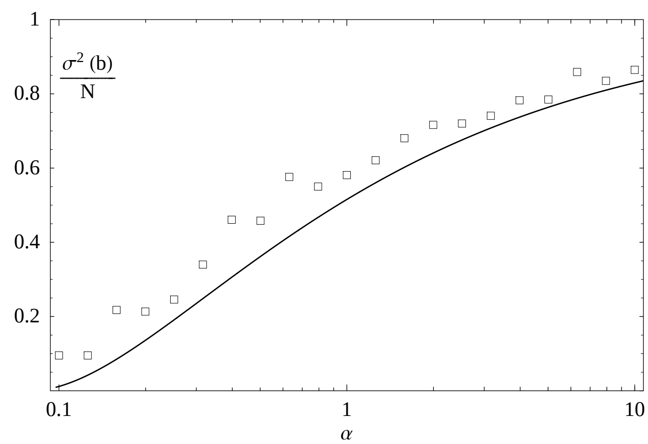

Figure 2: Volatility as a function of the training parameter α from numerical simulations (boxes) and its theoretical estimate (line). Here again we have used the parameter values $B = 5$ and $S = 2$.

4. THE EFFECT OF CHOICES

Already, figures 1 and 2 are a strong indication that, even as selfish users try to maximise their gain in the simplex game, they actually start to operate more efficiently and the game converges to a socially optimal state. However, a fundamental question that still remains has to do with the way that the number of choices affects the overall performance of the game: does abundance of choices confuse players who no longer know how to operate efficiently, or does it reduce individual fluctuations and allow the game to converge to a steady state more easily?

Naively, one could argue both ways but, thanks to the efficiency measure of equation 3.6 we may provide an exact answer. Indeed, recall that from equation 2.5 and the fact that $\mathbf{E}(b) = 0$, we immediately obtain that: $\mathbf{E}(N_r) = \frac{N}{B}$.

Furthermore:

$$\mathbf{E}(N_r^2) = \frac{N^2}{B^2} + \frac{(B-1)^2}{B^2}\,\mathbf{E}\left((q_r\cdot b)^2\right)$$

and, since:

$$\mathbf{E}\left((q_r\cdot b)^2\right) = \frac{1}{B}\,\mathbf{E}\left(\sum_{r=1}^{B}(q_r\cdot b)^2\right)$$

by symmetry considerations, we may combine the properties of the characteristic B-states (lemma A.1) and equation 3.6 to finally obtain the variance of the number of users per access point:

$$\frac{\sigma^2(N_r)}{N} = \frac{B-1}{B^2}\frac{\sigma^2(b)}{N} = \frac{B-1}{B^2}\Theta(\alpha-\alpha_c)\left(1-\sqrt{\frac{\alpha_c}{\alpha}}\right)^2 \tag{4.1}$$

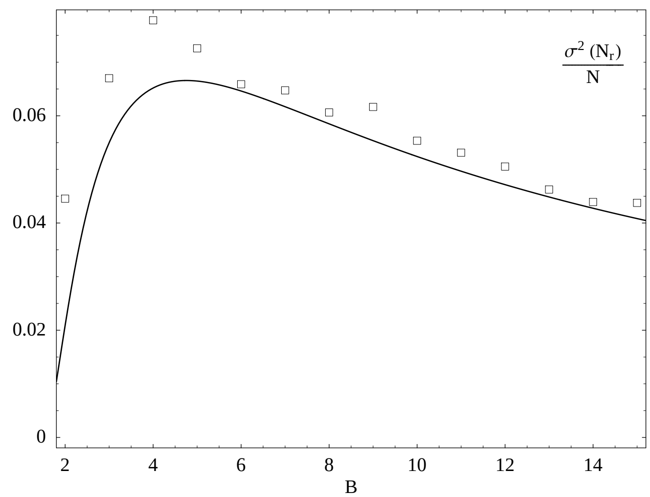

Figure 3: The variance of the number of users connected to a particular access point as a function of the total number of access points.

In figure 3 we have exhibited precisely this behaviour for a simplex game consisting of 100 players and a training set with size $M = 64$ (as always, we have restricted ourselves to the simplest case of $S = 2$ strategies per player). Surprisingly enough, we see that there is a critical number of access points ($B_c = 4$ for the parameter values of figure 3) where the network's performance is at an absolute worse (high fluctuations on the number of users per access point).

This can be explained by considering the competing factors that determine the game's performance: abundance of choices versus players' confusion. Indeed, if there are few enough access points, addition of further nodes will actually decrease the game's performance because the players are unable to process the training signal quickly enough in order to make educated guesses, i.e. they get confused by too much information, and this results in poor performance. However, beyond a critical number of access points, the players' confusion is quite mild compared to the smoothing effect that a large number of access points provides to the network's performance, and we see that the variance decreases (asymptotically, it vanishes).

Of course, this critical number of access points depends sharply on the training parameter α and the number S of strategies per player. In fact, from equation 4.1, we can easily see that, if the number of nodes is rather small (say 4 or 5), network performance will actually suffer if we increase the quality of the training signal (i.e. the training parameter α) because we will be in the region were players are being confused by too much information. On the other hand, if we also increase the number S of strategies per player, the players will now have enough decision schemes with which to process the signal presented to them and the variance will, again, decrease. So, in short, unless the players are quite sophisticated and have a lot of strategies that need to be ranked, they learn to operate much more effectively with a training signal of small size, simply because they are better able to process it to their advantage in the game's extremely complex environment.

5. CONCLUSION

In summary, we have introduced and analysed a non-cooperative game played by heterogeneous agents with bounded rationality. Unlike other applications of game theory to telecommunications, these players do not adapt their choices to environmental changes but, rather, change their way of thinking. In this sense, our model is closely related to the minority game encountered in econophysics (e.g. [1, 5]). We actually generalise this concept in order to allow players to choose from a multitude of different resources, thus providing the foundations for application to a WiFi context or carriers in an OFDMA case. Then, by using the replica mehtod, we have calculated the variance of the game's aggregate bid and have found it to be significantly smaller than the fully random situation where each player makes random choices, a fact that is confirmed by our numerical simulations. These promising results suggest that this method can be used to study network problems involving competition, such as frequency hopping in OFDMA in order to minimise collisions.

Acknowledgements

This research was supported by the European Commission under grants MIRG-CT-2005-030833 and also FET-FP6-IST-034413.

APPENDIX

A. CHARACTERISTIC STATES

In order to define the characteristic B-states, let us consider a $(B-1)$-simplex Δ_{B-1} (affinely) embedded in $\mathbb{R}^{B-1}$, so that its barycentre lies at the origin while its vertices lie on the unit sphere S^{B-2}; we will then refer to its vertices $\{q_r\}_{r=1}^{B}$ as *characteristic B-states.* It is straightforward to set up a specific embedding $q : \Delta_{B-1} \to \mathbb{R}^{B-1}$ that satisfies the above requirements, but the expressions are a bit cumbersome and we will avoid them. Fortunately, we only need some easily obtainable calculational properties of the characteristic states, encapsulated in the following:

LEMMA A.1. *If $\mathscr{B} = \{q_r\}_{r=1}^{B} \subseteq \mathbb{R}^{B-1}$ denotes a set of characteristic B-states then, for all $r, l = 1, 2 \ldots B$:*

1. $q_r \cdot q_l = \frac{B}{B-1}\delta_{rl} - \frac{1}{B-1}$
2. $\sum_{r=1}^{B} q_r = 0$

3. *for all* $\xi \in \mathbb{R}^{B-1}$*:* $\sum_{r=1}^{B} (q_r \cdot \xi)^2 = \frac{B}{B-1}\xi^2$.

PROOF. Motivated by the standard embedding $\Delta_{B-1} \hookrightarrow \mathbb{R}^B$, let $\{e_r\}_{r=1}^B$ be the usual basis of $\mathbb{R}^B$ and set $e_0 = \frac{1}{B}\sum_{r=1}^B e_r$. Then, if $\tilde{q}_r = \sqrt{\frac{B}{B-1}}(e_r - e_0)$, it is not hard to check that the set $\{\tilde{q}_r\}_{r=1}^B$ is a set of characteristic B-states for the hyperplane $E_0 = \{\xi\colon \xi \cdot e_0 = 0\}$ of $\mathbb{R}^B$ that is perpendicular to e_0.

In this embedding, all properties follow trivially. To begin with:

$$\begin{aligned}\tilde{q}_r \cdot \tilde{q}_l &= \frac{B}{B-1}(e_r - e_0)\cdot(e_l - e_0)\\ &= \frac{B}{B-1}\delta_{rl} - \frac{1}{B-1}\end{aligned}$$

Furthermore:

$$\left(\sum_{r=1}^B q_r\right)^2 = \sum_{r=1}^B q_r^2 + \sum_{r\neq l} q_r \cdot q_l = 0$$

and, finally, property (3) simply has to be established on E_0, i.e for ξ such that $\xi \cdot e_0 = 0$. Indeed:

$$\sum_{r=1}^B (\xi \cdot \tilde{q}_r)^2 = \frac{B}{B-1}\sum_{r=1}^B (e_0\cdot\xi - e_r \cdot \xi)^2 = \frac{B}{B-1}\xi^2.$$

Since all properties are invariant with respect to affine embeddings (i.e. they are properties of the simplex and not of its embedding), our proof is complete. ☐

Property (1) is actually equivalent to our definition of characteristic states and implies the rest. However, from a calculational standpoint, property (3) is much more important since it is used in a key step of the proof of our main result on the virtual reduction of choices (theorem 3.2).

B. REDUCTION OF CHOICES

In order to prove the reduction theorem 3.2, we will turn to *replica theory*, a well-established tool of Statistical Mechanics that greatly facilitates calculations of the so-called ground states (what we refer to as "minimum energy"). We will use assumptions 1 and 2 as well as the following lemma:

LEMMA B.1 (STEEPEST DESCENT). *If D is a bounded domain and f is a measurable function on D, then:*

$$\log \int_D e^{xf(t)}\,\mathrm{d}t \sim x \max_{t\in D} f(t)$$

as $x \to \infty$.

This lemma is just a formal statement of the *steepest descent* method for asymptotic integration. We will not give a proof here (see e.g. [7]) but, instead, we will state an obvious consequence:

COROLLARY B.2. *By descending steeply to large x, we get:*

$$\max_{t\in D} f(t) = \lim_{x\to\infty}\frac{1}{x}\log\int_D e^{xf(t)}\,\mathrm{d}t.$$

and this yields equation 3.5.

Now, drawing our motivation from Statistical Mechanics, the Lyapunov function L can be considered as the *energy* of the *continuous-valued spin system* $\mathbf{p} = \{p_{is}\}$. Therefore, we introduce the *partition function*:

$$\mathscr{Z}(\beta, c) = \int_{\mathscr{D}} e^{-\beta L(\mathbf{p},c)}\,\mathrm{d}p$$

where $\mathscr{D}$ is the N-fold product of the $(S-1)$-simplices: $\Delta_{S-1}^i = \{p_{is} : \sum_s p_{is} = 1\}$ and $\mathrm{d}p$ is the usual Lebesgue measure on $\mathscr{D}$. Then, if we average over the space $\Omega = \mathscr{C}$ of strategy matrices with uniform probability measure, an application of lemma B.1 and assumptions 1 and 2 yields:

$$L_0(\mathscr{G}) = -\lim_{\beta\to\infty}\lim_{n\to 0}\frac{1}{\beta}\frac{\partial}{\partial n}\left\langle \mathscr{Z}^n(\beta, c)\right\rangle \tag{B.1}$$

We will first evaluate $\langle \mathscr{Z}^n(\beta)\rangle$ for $n \in \mathbb{N}$ (see e.g. [14]) using the identity:

$$\mathscr{Z}^n(\beta, c) = \prod_{\mu=1}^n \mathscr{Z}_\mu(\beta, c)$$

where $\mathscr{Z}_\mu(\beta, c)$ is the partition function of the μ^{th} *replica* $\{p_{is\mu}\}$ of the system. Subsequently, we will use assumption 2 to obtain an analytic continuation of this expression for real values of n in the vicinity of 0. As a result:

$$\begin{aligned}\langle \mathscr{Z}^n(\beta, c)\rangle &= \left\langle \textstyle\prod_{\mu=1}^n \mathscr{Z}_\mu(\beta, c)\right\rangle\\ &= \int_{\mathscr{D}^n} e^{N\beta \sum_\mu G_{\mu\mu}(\mathbf{p})}\left\langle e^{-\frac{\beta}{M}\sum_\mu\sum_m \mathbf{E}_\mu(b|c,m)^2}\right\rangle \mathrm{d}p\end{aligned}$$

where $G_{\mu\nu}(\mathbf{p}) = \frac{1}{N}\sum_i\sum_s p_{is\mu}p_{is\nu}$ and $\mathbf{E}_\mu$ is the expectation in the μ^{th} replica of the system:

$$\mathbf{E}_\mu(b|c,m) = \sum_i\sum_s p_{is\mu}c_{is}^m.$$

To proceed from here, we will employ a calculational property of Gaussian integrals that is referred to in the literature as the *Hubbard–Stratonovich* transformation. The fundamental identity is $\mathbf{E}_z\, e^{\mathrm{i}qz} = e^{-\frac{q^2}{2}}$ where

$$\mathbf{E}_z\, f = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(z)e^{-\frac{z^2}{2}}\,\mathrm{d}z$$

is the expectation with respect to a Gaussian variable of zero mean and unit variance: $z \sim N(0,1)$. We then get[8]:

$$\begin{aligned}&\left\langle e^{-\frac{\beta}{M}\sum_{\mu,m}\mathbf{E}_\mu(b|c,m)^2}\right\rangle\\ &= \left\langle \mathbf{E}_{\{z_\mu^m\}}\, e^{\mathrm{i}\sqrt{\frac{2\beta}{M}}\sum_{\mu,m} z_\mu^m\cdot\mathbf{E}_\mu(b|c,m)}\right\rangle\\ &\qquad = \mathbf{E}_{\{z_\mu^m\}}\left\langle \exp\left(\mathrm{i}\textstyle\sum_{i,s,m}\underbrace{\textstyle\sum_\mu\sqrt{\frac{2\beta}{M}}p_{is\mu}z_\mu^m}_{\xi_{is}^m}\cdot c_{is}^m\right)\right\rangle\end{aligned}$$

This is actually the crucial step of the calculation that necessitated the introduction of the characteristic B-states. To streamline the presentation, we prove:

[8]We will actually use *vectorial* Gaussian random variables $z_\mu^m \in \mathbb{R}^{B-1}$, each of whose coordinates has unit variance and zero mean, i.e.: $z_{\mu,k}^m \sim N(0,1)$. The only thing that changes is that we need to use the scalar product instead of the regular product.

LEMMA B.3. *Let $\xi_k \in \mathbb{R}^{B-1}$, $k = 1, \dots K$ be a finite collection of vectors and let c_k be independent uniformly distributed random variables on the set $\mathscr{B} = \{q_r\}_1^B$. Then:*

$$\left\langle e^{\mathrm{i}\sum_k \xi_k \cdot c_k}\right\rangle = e^{-\frac{1}{2(B-1)}\sum_k \xi_k^2} + \mathscr{O}(\xi^3)$$

PROOF. By expanding the exponential, we obtain:

$$\begin{aligned}
&\left\langle e^{\mathrm{i}\sum_k \xi_k \cdot c_k}\right\rangle \\
&= \left\langle 1 + \mathrm{i}\sum_k \xi_k \cdot c_k - \tfrac{1}{2}\left(\sum_k \xi_k \cdot c_k\right)^2 + \mathscr{O}(\xi^3)\right\rangle \\
&= 1 + \mathrm{i}\sum_k \xi_k \cdot \langle c_k\rangle - \tfrac{1}{2}\left\langle \sum_k (\xi_k \cdot c_k)^2\right\rangle \\
&- \tfrac{1}{2}\sum_{k\neq k'} \langle \xi_k \cdot c_k\rangle \langle \xi_k' \cdot c_k'\rangle + \mathscr{O}(\xi^3)
\end{aligned}$$

due to the stochastic independence of the c_k's. However, the uniform distribution for c_k implies that: $\langle c_k\rangle = \frac{1}{B}\sum_{r=1}^B q_r = 0$; moreover:

$$\left\langle (\xi_k \cdot c_k)^2\right\rangle = \tfrac{1}{B}\sum_{r=1}^B (\xi_k \cdot q_r)^2 = \tfrac{1}{B-1}\xi_k^2$$

by lemma A.1. We then conclude that:

$$\begin{aligned}
\left\langle e^{\mathrm{i}\sum_k \xi_k \cdot c_k}\right\rangle &= 1 - \tfrac{1}{2(B-1)}\sum_k \xi_k^2 + \mathscr{O}(\xi^3) \\
&= e^{-\frac{1}{2(B-1)}\sum_k \xi_k^2} + \mathscr{O}(\xi^3)
\end{aligned}$$

and this completes the proof. $\square$

By the definition of ξ_{is}^m we have $\xi_{is}^m = \mathscr{O}\left(M^{-\frac{3}{2}}\right) = o\left(\frac{1}{N}\right)$ and, hence, we can use the previous lemma to get:

$$\begin{aligned}
&\left\langle \exp\left(\mathrm{i}\sum_{i,s,m}\xi_{is}^m \cdot c_{is}^m\right)\right\rangle = \\
&\quad = \exp\left\{-\tfrac{1}{2(B-1)}\sum_{i,s,m}(\xi_{is}^m)^2\right\} + o\left(\tfrac{1}{N}\right) \\
&\quad \sim \exp\left\{-\tfrac{\beta}{\alpha(B-1)}\sum_m \sum_{\mu,\nu} G_{\mu\nu}(\mathbf{p}) z_\mu^m \cdot z_\nu^m\right\}
\end{aligned}$$

where $\alpha = \frac{M}{N}$ is the *training parameter* of the game. Then, if we set: $J_{\mu\nu}(\mathbf{p}) = \delta_{\mu\nu} + \frac{2\beta}{\alpha(B-1)}G_{\mu\nu}(\mathbf{p})$, the expectation over the z-variables yields the Gaussian integral:

$$\begin{aligned}
&\mathbf{E}_{\{z_{\mu,k}^m\}}\left\langle \exp\left(\mathrm{i}\sum_{i,s,m}\xi_{is}^m \cdot c_{is}^m\right)\right\rangle \sim \\
&\sim \int_{\mathbb{R}^{nM(B-1)}} e^{-\frac{1}{2}\sum_m \sum_{k=1}^{B-1}\sum_{\mu,\nu} J_{\mu\nu}(\mathbf{p}) z_{\mu,k}^m z_{\nu,k}^m}\,\mathrm{d}\tilde{z} \\
&= \left(\int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle z,z\rangle_{\mathbf{J}(\mathbf{p})}}\,\mathrm{d}\tilde{z}\right)^{M(B-1)} \\
&= \exp\left\{\tfrac{M(B-1)}{2}\log\det(\mathbf{J}(\mathbf{p}))\right\}
\end{aligned}$$

where we have introduced the bilinear pairing $\langle z, w\rangle_{\mathbf{J}} \equiv \sum_{\mu,\nu} J_{\mu\nu} z_\mu w_\nu$ and used the fundamental "change of variables" identity of multiple Gaussian integrals:

$$\int e^{-\frac{1}{2}\langle z,z\rangle_{\mathbf{J}}}\,\mathrm{d}\tilde{z} = \sqrt{|\det \mathbf{J}|}$$

(where $\mathrm{d}\tilde{z}$ is the ordinary Lebesgue measure $\mathrm{d}z$ normalised to: $\int_{\mathbb{R}} e^{-\frac{z^2}{2}}\,\mathrm{d}\tilde{z} = 1$).

Combining all of the above, we obtain:

$$\begin{aligned}
\langle \mathscr{Z}^n(\beta)\rangle \sim \int_{\mathscr{D}^n} \exp\Big\{ N\beta\Big[&\operatorname{tr}[\mathbf{G}(\mathbf{p})] - \\
&- \tfrac{\alpha(B-1)}{2\beta}\log\det\left(\mathbf{I} + \tfrac{2\beta}{\alpha(B-1)}\mathbf{G}(\mathbf{p})\right)\Big]\Big\}\,\mathrm{d}p \quad \text{(B.2)}
\end{aligned}$$

This expression is valid for any simplex game $\mathscr{G}$ with B choices and training parameter α, the probability measure P being reflected in the matrix $G_{\mu\nu}(\mathbf{p})$. Therefore, applying this formula to the binary reduction $\mathscr{G}_{\mathit{eff}}$ of $\mathscr{G}$ which has 2 choices, training parameter α_{eff} and is otherwise identical to $\mathscr{G}$, we obtain the averaged partition function:

$$\begin{aligned}
\langle \mathscr{Z}^n(\beta)\rangle \sim \int_{\mathscr{D}^n} \exp\Big\{ N\beta\Big[&\operatorname{tr}[\mathbf{G}(\mathbf{p})] - \\
&- \tfrac{\alpha_{\mathit{eff}}}{2\beta}\log\det\left(\mathbf{I} + \tfrac{2\beta}{\alpha_{\mathit{eff}}}\mathbf{G}(\mathbf{p})\right)\Big]\Big\}\,\mathrm{d}p
\end{aligned}$$

Since $\alpha_{\mathit{eff}} = \alpha(B-1)$, the two expressions are asymptotically equal and, by equation B.1, we immediately recover $L_0(\mathscr{G}) \sim L_0(\mathscr{G}_{\mathit{eff}})$, as claimed in theorem 3.2.

C. THE REPLICA SADDLE-POINT

Our result on the reduction of choices in the simplex game, already allows us to obtain an analytic expression for the optimal state $L_0(\mathscr{G})$, simply by adapting the results of [12]:

$$\tfrac{\sigma^2(b)}{N} = 1 + \tfrac{1}{N}L_0(\mathscr{G}) = \Theta(\alpha - \alpha_c)\left(1 - \sqrt{\tfrac{\alpha_c}{\alpha}}\right)^2 \qquad \text{(C.1)}$$

where $\alpha_c = \frac{\zeta^2(S)}{B-1}$ and $\zeta(S)$ is the expected value of the minimum of S Gaussian random variables $z_s \sim N(0,1)$:

$$\zeta(S) = \tfrac{S}{2^{S-1}}\sqrt{\tfrac{2}{\pi}}\int z e^{-z^2}\operatorname{erfc}^{S-1}(z)\,\mathrm{d}z.$$

For the sake of completeness, we will outline the replica calculation, mainly because we wish to also discuss the role of *replica symmetry*.

We will thus consider the exponential representation of the δ-function: $\delta(x) = \frac{1}{2\pi}\int e^{\mathrm{i}kx}\,\mathrm{d}k$ and introduce a δ-function for every entry of $G_{\mu\nu}(\mathbf{p})$:

$$\delta(\mathbf{Q} - \mathbf{G}(\mathbf{p})) = \left(\tfrac{N}{2\pi}\right)^{n^2}\int e^{\mathrm{i}N\sum_{\mu,\nu} k_{\mu\nu}(Q_{\mu\nu} - G_{\mu\nu}(\mathbf{p}))}\,\mathrm{d}k$$

Then, by descending to the large N limit in order to integrate asymptotically over Q and k, we obtain:

$$\begin{aligned}
&\tfrac{1}{N}\log\langle \mathscr{Z}^n(\beta)\rangle \sim -N\beta \times \\
&\times \Bigg[\tfrac{\alpha(B-1)}{2\beta}\log\det\left(\mathbf{I} + \tfrac{2\beta}{\alpha(B-1)}\mathbf{Q}\right) - \operatorname{tr}(\mathbf{Q}) \\
&- \mathrm{i}\sum_{\mu,\nu} k_{\mu\nu}Q_{\mu\nu} - \frac{1}{\beta}\log\int_{\Delta^n} e^{-\mathrm{i}\sum\limits_{\mu,\nu} k_{\mu\nu}\sum_s p_{s\mu}p_{s\nu}}\,\mathrm{d}p\Bigg]
\end{aligned}$$

where the matrices $\mathbf{Q}$, $\mathbf{k}$ have been chosen so as to extremise the function Λ within the brackets.

The *replica symmetry* assumption is that these *saddle-point* matrices have two kinds of elements, one in the diagonal and one off-diagonal; more precisely, we will seek extrema of the form $Q_{\mu\nu} = q + (Q - q)\delta_{\mu\nu}$ and $k_{\mu\nu} = \mathrm{i}\alpha\beta^2\frac{B-1}{2}(r + (R - r)\delta_{\mu\nu})$ (the scaling factor has been introduced for future convenience). Of course, this is a rather strong assumption since there is no a priori reason for the replicas to eventually converge to the same state. In [6], this assumption has been shown not to be valid and the authors perform the first step of *replica symmetry breaking* (1RSB) within the setting of the Parisi solution. However, it is also shown that the replica symmetric ground state does not differ significantly from the 1RSB state. However, replica symmetry breaking is beyond the scope of this short account of replica-theoretic techniques, and we will not carry out the 1RSB analysis here.

Regardless, under the replica symmetric assumption:

$$\Lambda = \frac{\alpha(B-1)}{2\beta}\log\det\left(q+\left(1+\frac{2\beta}{\alpha}\frac{Q-q}{B-1}\right)\delta_{\mu\nu}\right)-nQ$$
$$+\,n\alpha\beta\frac{B-1}{2}(QR-qr)+n^2\alpha\beta\frac{B-1}{2}$$
$$-\,\frac{1}{\beta}\log\int_{\Delta^n} e^{\alpha\beta^2\frac{B-1}{2}\left((R-r)\sum_\mu \mathbf{p}_\mu^2+r\left(\sum_\mu \mathbf{p}_\mu\right)^2\right)}\mathrm{d}p$$

where $\mathbf{p}_\mu$ is shorthand for the vector $\sum_s p_{s\mu}e_s$ (in the standard basis of $\mathbb{R}^S$). Now, in order to study the last term of this expression, we will use once again the Hubbard–Stratonovich transformation:

$$e^{\alpha r\beta^2\frac{B-1}{2}\left(\sum_\mu \mathbf{p}_\mu\right)^2} = \mathbf{E}_z\, e^{-\sqrt{\alpha r(B-1)}\,z\cdot\sum_\mu \mathbf{p}_\mu}$$

Then, if we let $V_z(p)=\sqrt{\alpha r(B-1)}\,p\cdot z-\alpha\beta\frac{B-1}{2}\cdot(R-r)p^2$, we get:

$$\log\int_{\Delta^n} e^{\alpha\beta^2\frac{B-1}{2}\left((R-r)\sum_\mu \mathbf{p}_\mu^2+r\left(\sum_\mu \mathbf{p}_\mu\right)^2\right)}\mathrm{d}p$$
$$=\log\mathbf{E}_z\int_{\Delta^n} e^{-\beta\sum_\mu V_z(\mathbf{p}_\mu)}\,\mathrm{d}p$$
$$=\log\mathbf{E}_z\, e^{n\log\int_\Delta \exp\{-\beta V_z(\mathbf{p})\}\,\mathrm{d}p}$$
$$=n\,\mathbf{E}_z\log\int_\Delta e^{-\beta V_z(\mathbf{p})}\,\mathrm{d}p+o(n)$$

Now, because of assumption 2, we only need to calculate $\Lambda_0=\lim_{n\to 0}\frac{1}{n}\Lambda$. So, if we retrace our steps and perform some elementary calculations along the way, we find:

$$\Lambda_0=\frac{q}{1+\chi}+\frac{\alpha}{2\beta}(B-1)\log(1+\chi)-Q$$
$$+\,\alpha\beta\frac{B-1}{2}(QR-qr)$$
$$-\,\frac{1}{\beta}\mathbf{E}_z\left[\log\int_\Delta \exp\left(-\beta\tilde{H}_z(p)\right)\right]\qquad\text{(C.2)}$$

where $\chi=\frac{2\beta}{\alpha}\frac{Q-q}{B-1}$ and the parameters Q,q,R,r are such as to satisfy the *replica symmetric saddle-point* conditions: $\frac{\partial\Lambda_0}{\partial Q}=0,\ldots$

These equations have to be solved in the limit $\beta\to\infty$ to which we will descend by taking advantage of the asymptotic approximation of lemma B.1. To wit, we first let $p_*(z)\in\Delta_{S-1}$ be the point of Δ_{S-1} where V_z attains its minimum[9] and set $\zeta=\mathbf{E}_z[p_*(z)\cdot z]$ and $\phi=\mathbf{E}_z[p_*^2(z)]$. Then, one can show that both $(Q-q)$ and $(R-r)$ are $\mathscr{O}(\frac{1}{\beta})$, and we obtain the following asymptotic solutions:

$$Q\sim\phi \qquad\qquad R=r+\frac{1}{\alpha\beta(B-1)}\frac{2\chi}{1+\chi}$$
$$q\sim\phi+\frac{1}{\beta}\frac{\zeta}{\sqrt{\alpha r(B-1)}} \qquad\qquad r=\frac{4}{\alpha^2(B-1)^2}\frac{q}{(1+\chi)^2}$$

Finally, after some geometric considerations, one sees that $z\cdot p_*(z)=\min_{s\in\mathscr{S}}\{z_1\ldots z_S\}$, an expression that yields $\phi=1$ and also:

$$\zeta\equiv\zeta(S)=\frac{S}{2^{S-1}}\sqrt{\frac{2}{\pi}}\int z e^{-z^2}\mathrm{erfc}^{S-1}(z)\,\mathrm{d}z$$

Therefore, for finite χ (which happens for $\alpha\geq\alpha_c=\frac{\zeta^2(S)}{B-1}$), expression C.2 becomes:

$$\Lambda_0=\Theta(\alpha-\alpha_c)\left(1-\sqrt{\frac{\alpha_c}{\alpha}}\right)^2-1\qquad\text{(C.3)}$$

and, with $\Lambda_0=\frac{1}{N}L_0(\mathscr{G})$, we obtain equation 3.6:

$$\frac{\sigma^2(b)}{N}=1+\frac{1}{N}L_0(\mathscr{G})=\Theta(\alpha-\alpha_c)\left(1-\sqrt{\frac{\alpha_c}{\alpha}}\right)^2.$$

[9]The function $p_*(z)$ is only well-defined almost everywhere but still remains a measurable function.

D. REFERENCES

[1] William B. Arthur, *Inductive reasoning and bounded rationality (the El Farol problem)*, Am. Econ. Assoc. Papers Proc. **84** (1994), 406–411.

[2] N. Bonneau, E. Altman, M. Debbah, and G. Caire, *An evolutionary game perspective to ALOHA with power control*, Proceedings of the 19th International Teletraffic Congress, Beijing, China (2005).

[3] N. Bonneau, M. Debbah, E. Altman, and A. Hjorungnes, *Wardrop equilibrium for CDMA systems*, Rawnet, 2007.

[4] Andrea Cavagna, *Irrelevance of memory in the minority game*, Phys. Rev. E. **59** (1999), 3783–3786.

[5] Damien Challet and Yi-Cheng Zhang, *On the minority game: Analytical and numerical studies*, Physica A **246** (1997), 407–418.

[6] Andrea de Martino and Matteo Marsili, *Replica symmetry breaking in the minority game*, Journal of Physics A: Mathematical and General **34** (2001), no. 12, 2525–2537.

[7] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Springer-Verlag, 1998.

[8] Aram Galstyan, Bhaskar Krishnamachari, and Kristina Lerman, *Resource allocation and emergent coordination in wireless sensor networks*, AAAI Workshop on Sensor Networks (2004).

[9] Aram Galstyan and Kristina Lerman, *Adaptive Boolean networks and minority games with time-dependent capacities*, Phys. Rev. E **66** (2002), 0151103R.

[10] R. Iyer and Leonard Kleinrock, *Qos control for sensor networks*, Proceedings of IEEE ICC, 2003.

[11] Allen B. MacKenzie and Stephen B. Wicker, *Selfish users in ALOHA: A game theoretic approach*, IEEE Communications Magazine **126** (2001), 126.

[12] Matteo Marsili, Damien Challet, and Riccardo Zecchina, *Exact solution of a modified El Farol's bar problem: Efficiency and the role of market impact*, Physica A **280** (2000), 522–553.

[13] Farhad Meshkati, Mung Chiang, Vincent Poor, and Stuart C. Schwartz, *A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems*, IEEE Journal on selected areas in communications **24** (2006), 1115–1129.

[14] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro, *Spin glass theory and beyond*, World Scientific, 1987.

[15] P. Michiardi and R. Molva, *A game theoretic approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks*, Proceedings of WiOpt'03, Sophia-Antipolis, France, March 2003.

[16] Aris L. Moustakas, Steven H. Simon, and Anirvan M. Sengupta, *MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis*, IEEE Journal of Information Theory **49** (2003), no. 10, 2545–2561.

[17] Toshiyuki Tanaka, *A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors*, IEEE Journal of Information Theory **48** (2002), no. 11, 2888–2910.