Population games with networking applications
Hamidou Tembine

HAL Id: tel-00451970
https://theses.hal.science/tel-00451970

Submitted on 1 Feb 2010

THÈSE

présentée à l'Université d'Avignon et des Pays de Vaucluse
pour obtenir le diplôme de DOCTORAT

SPÉCIALITÉ : Informatique

École Doctorale 380 « Sciences et Agronomie »
Laboratoire d'Informatique (EA 931)

Population games with networking applications

par
Hamidou TEMBINE

Soutenue publiquement le 18 Septembre 2009 devant un jury composé de :

M.	Philippe MICHELON	Professeur, Université d'Avignon	Président
M.	Tamer BAŞAR	Professeur, University of Illinois at Urbana-Champaign	Rapporteur
M.	Nahum SHIMKIN	Professeur, Technion	Rapporteur
M.	Pierre BERNHARD	Professeur Emérite, INRIA	Examinateur
M.	Konstantin AVRACHENKOV	Chargé de recherche, INRIA	Examinateur
M.	Eitan ALTMAN	Directeur de Recherche, INRIA, Sophia Antipolis	Directeur
M.	Rachid EL-AZOUZI	Maître de Conférences, Université d'Avignon	Co-Directeur

Laboratoire d'Informatique d'Avignon

Contents

iii

Acknowledgments

I would like to thank my advisors, Eitan Altman and Rachid El Azouzi for their continuous guidance and support throughout my PhD studies. I am grateful to them for having given me the opportunity to begin this work, the freedom to carry it out, and the support and encouragement to bring it to a conclusion. I am grateful to LIA members, Marc El Beze, the scientific leader of LIA, for having hosted me during this thesis. A large part of this thesis owes its existence to the work done in collaboration with Jean Yves Le Boudec, Yezekael Hayel, and William H. Sandholm. My sincere gratitude for all the delightful discussions. I wish to thank Tamer Başar for hosting me as PhD Internship at University of Illinois at Urbana Champaign in Fall 2008. It gave me the opportunity to learn new tools and methods, and to extend my work to "evolutionary games with continuous action space" and application to access control, rate control and network security. Many thanks to Jean Yves Le Boudec who greatly helped me in understanding of mean field interactions during my internship at Ecole Polytechnique Federale de Lausanne (EPFL) in Summer 2008 and in Winter 2009 . His patience and generous feedback on many problems we worked on helped me greatly in writing the chapters 6 and 7 and the new class of games : *differential population games* that is derived from. Many thanks to William H. Sandholm for helpful comments on "differential population games" during my stay at University of Wisconsin in August 2009.

It is a great privilege for me to have Konstantin Avrachenkov, Tamer Başar, Pierre Bernhard, Philippe Michelon and Nahum Shimkin as members of my PhD thesis committee. I wish to thank them for having accepted this task. I would like to thank Tamer Başar and Nahum Shimkin for agreeing to serve as reviewers, for having carefully read the manuscript. I am thankful to Fréderic Bonnans, Pierre Cardaliaguet, Patrick Combettes, Michael Eisermann, Stéphane Gaubert, Jean Charles Gilbert, Lucien Guillou, Barthel Gottfried, Anatoli Iouditski, Siegmund Kosarew, Jean Pierre Ponssard, Dinah Rosenberg, Eilon Solan, Sylvain Sorin, Tristan Tomala, Yannick Viossat, Patrick Witomski etc for constructive comments during my graduate and undergraduate studies that lead me to think about several problems in economics, biology, optimization, control, stochastic processes and game theory. I am in particular indebted to my colleagues and coauthors, with whom I spent many years, to Ephie Deriche, Jocelyne Gourret, Patricia Hjelt, Rebecca Lonberger, Simone Mouzac for their administrative assistance during the last two years. I am deeply grateful to French Ministry of Research, INRIA, EPFL, European Project BIONETS, Project WiNEM, Project POPEYE for having provided the financial support which made this work possible. Last but the least, gratitude to all, who helped directly or indirectly.

Chapter 1

Introduction

Recently, there has been a surge in research activities that employ game theory to model and analyze the performance of various networks, such as communication networks, computer networks, social networks, biologically inspired networks etc. There already exist several successful examples where game theory provides deeper understanding of complex network dynamics and leads to better design of efficient, scalable, and robust networks. Still, there remain many interesting open research problems yet to be identified and explored, and many issues to be addressed. Important analysis and applications have been done in the context of *static games*. Motivated by the dynamic behavior of most of the long-term systems and the understanding of prediction, learning and evolution, *dynamic game theory* has found several applications. Those include cooperative and non-cooperative models of repeated games, sequential games, stochastic games, differential games, evolutionary games etc.

Evolutionary games in large population provides a simple framework for describing strategic interactions among large numbers of players. Traditionally, predictions of behavior in game theory are based on some solution concept, typically Cournot equilibrium (64), Bertrand equilibrium (47) or some extension/refinement thereof. These notions require the assumption of equilibrium knowledge, which posits that each user correctly anticipates how other players will act or react. The equilibrium knowledge assumption is too strong and is difficult to justify in particular in contexts with large number of players. As an alternative to the equilibrium approach, an explicitly dynamic updating choice is proposed, a model in which players myopically update their behavior in response to their current strategic environment. This dynamic procedure does not assume the automatic coordination of player's actions and beliefs, and it can derive many specifications of players' choice procedures. These procedures are specified formally by defining a revision of pure strategies called *a revision protocol*. A revision protocol takes current payoffs (also called *fitness* in behavioral ecology) and aggregate behavior as inputs; its outputs are conditional switch rates which describe how frequently players in some class playing strategy *a* who are considering switching strategies move to another strategy *b*, given that the current payoff vector and subpopulation state. Revision protocols are flexible enough to incorporate a wide variety of paradigms, including ones based on imitation, adaptation, learning, optimization, etc. The revision protocols describe the procedures players follow in adapting their behavior in the dynamic evolving environment such as evolving networks.

Birth of Evolutionary Game Theory in Engineering

The evolutionary games formalism is a central mathematical tool developed by biologists for predicting population dynamics in the context of *interactions between populations*. This formalism identifies and studies two concepts: the *Evolutionary Stability*, and the *Evolutionary Game Dynamics*. The unbeatable strategy has been defined by Hamilton (95; 96) which is the analogous of strong equilibrium (resilient against multilateral deviations) in large systems. A weaker of notion of locally unbeatable strategy, the *Evolutionarily Stable State or Strategy* (ESS), has been defined by the biologists Maynard Smith & Price (186). The ESS is characterized by a property of robustness against invaders (mutations). More specifically, (i) if an ESS is reached, then the proportions of each population do not change in time. (ii) at ESS, the populations are immune from being invaded by other small populations. This notion is stronger than Nash equilibrium in which it is only requested that a single user would not benefit by a change (mutation) of its behavior. The ESS concept helps to understand mixed strategies in games with symmetric payoffs. A mixed strategy can be interpreted as a composition of the population. An ESS can be also interpreted as a Nash equilibrium of the one-shot game but a (symmetric) Nash equilibrium cannot be an ESS. As is shown in (223), ESS has strong refinement properties of equilibria such as proper equilibrium, perfect equilibrium etc. Before the ESS concept, Hamilton has introduced the concept of *Unbeatable Strategy* (95; 96), which is stronger than ESS. Although ESS and unbeatable strategy have been defined in the context of biological systems, it is highly relevant to engineering as well (230). In the biological context, the replicator dynamics is a model for the change of the size of the population(s) as biologist observe, where as in engineering, we can go beyond characterizing and modeling existing evolution. The evolution of protocols can be engineered by providing guidelines or regulations for the way to upgrade existing ones and in determining parameters related to deployment of new protocols and services.

There have been a lot of work on non-cooperative modeling of power control using game theory (76; 166). There are two advantages in doing so within the framework of evolutionary games:

- it provides the stronger concept of equilibria, the ESS, which allows us to identify robustness against deviations of a fraction of players, and

- it allows us to apply the generic convergence theory of evolutionary game dynamics, and stability results that we introduce in next chapters.

Homogenous population

We consider the standard setting of evolutionary games : there is a large populations of players; each member of the population has the same pure action set $\mathcal{A}$ (a finite set).

Mixed strategies and population profile

Denote by $\Delta(\mathcal{A})$ the $(|\mathcal{A}| - 1)-$dimensional simplex of $\mathbb{R}^{|\mathcal{A}|}$. Let $x(t)$ be the $|\mathcal{A}|-$ dimensional vector whose element $x_a(t)$ is the population share of strategy a at time t. Thus we, have $\sum_{a \in \mathcal{A}} x_a(t) = 1$ and $x_a(t) \geq 0$. We frequently use an equivalent interpretation where $x(t)$ is a *mixed strategy* used by all players at time t; by a mixed strategy we mean that a player chooses at time t an action a with probability $x_a(t)$. With either interpretations, at each local interaction occurring at time t a given player can expect that the other player would play strategy a with

probability $x_a(t)$. The vector $x(t)$ will also be called the *state* of the population at time t. Define the expected payoff of a player with the action a when the population profile $x(t)$ by $f_a(x(t))$.

Equilibrium and Evolutionary Stability

The evolutionarily stable state or strategy (ESS) concept have the property that if it is reached, then the proportions of each population do not change in time. Suppose that, initially, the population state has been $x \in \Delta(\mathcal{A})$ during a long time. The average payoff in the population is $\sum_{a \in \mathcal{A}} x_a f_a(x)$. Now suppose that a small group of mutants enters this population playing according to a different profile *mut* (and persists using it during a time longer than the delays). If we call $\epsilon \in (0, 1)$ the size of the subpopulation of mutants after normalization, then the population profile after mutation will be $\epsilon \, mut + (1 - \epsilon)x$. After mutation, the average payoff of non-mutants who are randomly matched to mutants is given by $\sum_{a \in \mathcal{A}} x_a f_a(mut)$, and the average payoff of non-mutants will be given by

$$\epsilon \sum_{a \in \mathcal{A}} x_a f_a(mut) + (1 - \epsilon) \sum_{a \in \mathcal{A}} x_a f_a(x).$$

Analogously, we can construct the average payoff of mutant

$$\epsilon \sum_{a \in \mathcal{A}} mut_a \, f_a(mut) + (1 - \epsilon) \sum_{a \in \mathcal{A}} mut_a f_a(x).$$

A population $x \in \Delta(\mathcal{A})$ is an ESS if for all $mut \neq x$, there exists some $\epsilon_{mut} \in (0, 1)$, which may depend on *mut*, such that for all $\epsilon \in (0, \epsilon_{mut})$

$$\epsilon \sum_{a \in \mathcal{A}} mut_a f_a(mut) + (1 - \epsilon) \sum_{a \in \mathcal{A}} mut_a f_a(x) \tag{1.1}$$

$$< \epsilon \sum_{a \in \mathcal{A}} x_a f_a(mut) + (1 - \epsilon) \sum_{a \in \mathcal{A}} x_a f_a(x)$$

That is, x is ESS if, after mutation, non-mutants are more successful than mutants, in which case mutants cannot invade and will eventually get extinct. The number ϵ_{mut} is called invasion barrier See Weibull (1995). It is the maximum rate of mutants against which s is resistant. If x is an ESS then x is a Nash equilibrium. Equivalently x is an ESS if and only if it meets best response conditions:

$$\sum_{a \in \mathcal{A}} mut_a f_a(x) \leq \sum_{a \in \mathcal{A}} x_a f_a(x), \; \forall \, mut, \tag{1.2}$$

$$\forall \, mut \neq x, \; \sum_{a \in \mathcal{A}} mut_a f_a(x) = \sum_{a \in \mathcal{A}} x_a f_a(x) \Rightarrow \tag{1.3}$$

$$\sum_{a \in \mathcal{A}} mut_a f_a(mut) < \sum_{a \in \mathcal{A}} x_a f_a(mut) \tag{1.4}$$

For population games with non-linear payoffs (i.e when $x \longmapsto f_a(x)$ is non-linear function), define the solution and evolutionary stability concepts as follows: A population profile x is an **equilibrium state** if

$$\sum_{a \in \mathcal{A}} x_a f_a(x) \geq \sum_{a \in \mathcal{A}} mut_a f_a(x), \; \forall \, mut$$

This variational inequality is equivalent to :

$$(*) \ \forall a \in \mathcal{A}, \ x_a > 0 \implies f_a(x) = \max_{b \in \mathcal{A}} f_b(x)$$

This last condition is sometimes referred to "Wardrop first principle" (233) of optimality and can be easily obtained using the indifference condition at mixed equilibria. The condition $(*)$ is sometimes written in the form: $\text{supp}(x) \subseteq \arg\max_a f_a(x)$.

We say that a population profile $x = (x_a)_{a \in A}$ is evolutionarily stable if for each deviation strategy called "mutant strategy" $mut = (mut_a)_{a \in A} \neq x$, there exists some $\epsilon_{mut} > 0$ such that $\forall \epsilon \in (0, \epsilon_{mut})$,

$$\sum_{a \in A} (x_a - mut_a) f_a(\epsilon \ mut + (1 - \epsilon)x) > 0. \tag{1.5}$$

The strategy x is a neutrally evolutionary stable state (NESS) if the above inequality is non-strict. If the inequality is non-strict and $\epsilon = 0$ then the resulting inequality says that x is an equilibrium state. Thus, the ESS notion is stronger than Nash equilibrium.

A population profile $x \in \Delta(\mathcal{A})$ is an unbeatable state if for any $mut \neq x$,

$$\sum_{a \in \mathcal{A}} (x_a - mut_a) f_a(\epsilon mut + (1 - \epsilon)x) > 0, \ \forall \epsilon \in (0, 1) \tag{1.6}$$

That is, x is an unbeatable state if, after mutation of any size (it can be all the population), non-mutants are more successful than mutants. In other words, any mutants of any size cannot invade the population. Note that an unbeatable state is an evolutionarily stable state, which implies neutrally stable state, which is an equilibrium of the one-shot game i.e if the population profile is at this state and no player has incentive to unilaterally change his/her action. The following inclusion holds:

$$\Delta_{Unbeatable \ state} \subset \Delta_{ESS} \subset \Delta_{Neutrally \ stable} \subset \Delta_{Equilibrium \ state} \tag{1.7}$$

Note that evolutionarily stable state and unbeatable state may not exists (see for example the class of Rock-Paper-Scissor games (99; 90; 65; 234)). The following theorem gives another necessary and sufficient condition to be an ESS for evolutionary games with bilinear payoff functions. A proof can be found in (94).

Theorem 1.0.0.1. *Let q be a symmetric (Nash) equilibrium for the matrix game with payoffs (F, F^t) where F^t is the transposition matrix of $F = (F(i,j)_{i,j})$ and $BRP(q)$ be the pure best response to q i.e*

$$BRP(q) = \left\{ j \mid \sum_k F(j,k)q_k = \max_l \sum_k F(l,k)q_k \right\}.$$

Define $\bar{q}$ as

$$\bar{q}_j = \begin{cases} q_j & \text{if } j \in BRP(q) \\ 0 & \text{otherwise} \end{cases}$$

Let $\tilde{F}$ the submatrix obtained from F by taking only the index $i, j \in BRP(q)$. Then q is an evolutionarily stable strategy if and only if

$$\sum_{k \in BRP(q)} (p_k - \bar{q}_k) \sum_{j \in BRP(q)} \tilde{F}_{k,j}(p_j - \bar{q}_j) < 0$$

for all $p \neq q$.

Note that the last condition is equivalent to

$$\forall y \in Y, \quad \sum_{k,j \in BRP(q)} y_k y_j \tilde{F}_{k,j} < 0$$

where

$$Y := \left\{ z \in \mathbb{R}^{|BRP(q)|} \backslash \{0\}, \quad \sum_{j \in BRP(q)} z_j = 0, \text{ and } \bar{q}_j = 0 \implies z_j \geq 0 \right\}$$

and $|BRP(q)|$ is the cardinal of the finite set $BRP(q)$.

ESS is an important refinement of Nash equilibria as shown in the following example.

Example

Consider the following matrix game with two players. Each has one block of strategies 1 nd one pure strategy 2. The block contains at least two strategies. The corresponding payoffs are given in the tabular where the player I chooses a row $r1$ (the block) or $r2$ and player II chooses a column $c1$ (block) or $c2$.

		Player II	
		c1	c2
Player I	r1	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$
	r2	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\max(\alpha,\beta),\max(\alpha,\beta)$

where $\alpha, \beta \in \mathbb{R}$. The game has an infinity of Nash equilibria: (i) one of the player chooses the first strategy (in the block) and the second an arbitrary strategy, (ii) Both players choose the second strategy. But the game has only one ESS (the second strategy). An allocation of payoffs is said *Pareto optimal* if the outcome cannot be improved upon without hurting at least one player. In this case the ESS is also Pareto optimal because the payoff obtained at ESS is the maximum payoff that each player can have.

When the block has only two pure actions, the game becomes the following: Each player has three strategies $1, 2$ or 3. The corresponding payoffs are given in the tabular where the player I chooses a row $r1, r2$ or $r3$ and player II chooses a column $c1, c2$ or $c3$.

	c1	c2	c3
$r1$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$
$r2$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$
$r3$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\max(\alpha,\beta),\max(\alpha,\beta)$

ESS is not unbeatable

We ask the following question: What happens when the size of mutants is greater than the invasion barrier of an ESS? To answer to question, we use the unbeatable strategy, a concept defined by Hamilton (1967), as a population profile which is resilient to any deviant strategy of any size. An unbeatable strategy is always an ESS, though an ESS is not necessarily unbeatable, as it may be beaten by large migrations into the population. The following example illustrates a situation where ESS is not unbeatable

	$c1$	$c2$	$c3$
$r1$	$\max(\alpha,\beta),\max(\alpha,\beta)$	$\max(\alpha,\beta),\max(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$
$r2$	$\max(\alpha,\beta),\max(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\max(\alpha,\beta),\min(\alpha,\beta)$
$r3$	$\min(\alpha,\beta),\min(\alpha,\beta)$	$\min(\alpha,\beta),\max(\alpha,\beta)$	$\min(\alpha,\beta),\min(\alpha,\beta)$

Are strong Nash equilibria evolutionarily stable strategies?

Aumann was one of the first one to introduce a number of concepts in game theory. One of these concepts was a strong equilibrium (34), which is a pure Nash equilibrium, in which not only single players cannot benefit from changing their strategy (to a different pure strategy), but no non-empty subset of players can form a coalition, where a coalition (group) means that all of them can change their strategies together, and all gain from the change.

In a strong equilibrium (see also *acceptable point*), no coalition (of any size) can deviate and improve the utility of every member of the coalition (while possibly lowering the utility of players outside the coalition). Clearly, every strong equilibrium is a Nash equilibrium, but the converse does not hold. In cases where a strong equilibrium exists, it seems to be a very robust notion. Considering strong equilibrium allows us to separate the effect of selfishness (which remains in strong equilibria) from that of lack of coordination (which disappears, since a strong equilibrium is resilient to deviations in coalitions).

In a $n-$player one-shot normal form game we say that a pure strategy profile is a $k-$strong equilibrium if it is robust by deviation of the group of size at most $k(k \leq n)$. A pure Nash equilibrium is resilient to pure deviations of groups of size k, if there is no group of size at most k, such that the strategy profile is not resilient to a pure deviation by the subset of the group. A $k-$strong equilibrium is a pure Nash equilibrium that is resilient to pure deviations of coalitions of size at most k. We will discuss how to define strong equilibrium in randomized strategies. In a Nash equilibrium no player can improve its own payoff by unilaterally changing its action. Thus, a Nash equilibrium corresponds to $1-$strong equilibrium. It is easy to see that strong equilibrium is equivalent to a $n-$strong equilibrium.

We ask the following question: Are strong Nash equilibria evolutionarily stable strategies? We show that the answer is negative. To see this, consider an evolutionary game with constant payoff. It is clear that the game has many strong equilibria. None of theses equilibria are evolutionarily stable strategies.

Evolutionary game dynamics

As an alternative to the equilibrium approach, evolutionary game dynamics propose an explicitly dynamic updating choice, a model in which players myopically update their behavior in response to their current strategic environment. This dynamic procedure does not assume the automatic coordination of users' actions and beliefs, and it can derive many players' choice procedures. These procedures are specified formally by defining a revision of pure strategies. The revision of strategies describes the procedures users follow in adapting their behavior to in the dynamic evolving environment such as in evolving networks.

There is a large number of population dynamics have been used in the context of non-cooperative games. Examples are the excess payoff dynamics, the fictitious play dynamics such as Brown-von Neumann-Nash dynamics (Brown & von Neumann, 1950), gradient-based

game dynamics (Rosen, 1965), replicator dynamics (Taylor & Jonker, 1978), pairwise comparison dynamics such as Smith dynamics (Smith, 1984), fictitious play (Gilboa & Matsui, 1991), projection dynamics, best response dynamics (Fudenberg & Levine, 1991), Fudenberg & Tirole, 1991), Boltzman dynamics and logit dynamics (Fudenberg & Levine, 1998), G-function based dynamics (Vincent & Brown, 2005) etc. Much literature can be found is the extensive survey on evolutionary game dynamics in (99) and in the book of Sandholm on Population Games and Evolutionary Dynamics (168). The *replicator dynamics* is one of the most studied evolutionary game dynamics. It has been used for describing the evolution of road traffic congestion in which the fitness is determined by the strategies chosen by all drivers (167). It has also been studied in the context of the association problem in wireless networks in (181).

Most of applications using evolutionary game dynamics was inter-disciplinary between economics, biology and behavioral ecology. Recently, game theory evolving has found many applications in the field of wireless communications and networks. The standard formulations of evolutionary games consider only pairwise interactions such as random matching. However, interactions which are not pairwise frequently arise in communication networks, such as the cases of Code Division Multiple Access (CDMA) system (209), Orthogonal Frequency-Division Multiple Access (OFDMA) in Worldwide Interoperability for Microwave Access (WiMAX) environment (11) etc. It has been applied to problems such as congestion control (147), distributed cooperative sensing over cognitive radio networks (232; 160), code division multiple access (CDMA), Orthogonal Frequency-Division Multiple Access (OFDMA) based Worldwide Interoperability for Microwave Access (WiMAX) environment (11), reciprocal and non-reciprocal interference control, mobile medium access control and channel selection (213) and capacity region of Additive White Gaussian Noise (AWGN) (88; 242), Multihoming and association problems (184; 209). These examples will be discussed in chapters 3 and 4.

The purpose of this manuscript is threefold.

Part I: Delayed Evolutionary Game Dynamics with Migration

The first part of this manuscript contains evolutionary stability and dynamic foundation of evolutionary games. The chapter 2 presents evolutionary games with delayed payoffs. We introduce and study *delayed evolutionary game dynamics* in which each pure strategy has its own time delay. Delayed evolutionary game dynamics are in general a system of first order non-regular nonlinear differential equations or differential inclusions with time delays. We use the theory of delayed differential equations (DDE), which is a special class of functional differential equations to study bifurcation, periodicity, stability, convergence and chaocity of the system under time delays. Delays play an important role in many situations in networking such as in flow control and congestion problems. We examine the effects of time delays in various contexts : congestion control protocols, dissemination information, Hawk and Dove game and multiple access game with regret cost using various delayed evolutionary game dynamics such delayed replicator dynamics, delayed imitation by dissatisfaction dynamics, delayed best response dynamics, delayed Brown-von Neumann-Nash dynamics, delayed projection dynamics, delayed generalized Smith dynamics, delayed evolutionary game dynamics with migration etc.

The chapter 3 extends the basic pairwise interaction model of evolutionary games to cover more general interactions including non-reciprocal interactions, and random number of players. In order to make use of the wealth of tools and theory developed in biology literature, many works in the area of computer networks ignore cases where local interactions between populations involve more than two players in a local interaction. This restriction limits the modeling power of evolutionary games which are not useful in a network operating at heavy load, such as internet traffic and dense networks. This motivated us to consider more than two players interacting locally. Moreover, the interactions can be *non-symmetric*. The resulting framework leads to evolving games with non-linear expected payoffs (in term of population profile). Equi-

librium analysis as well as dynamic behavior of networking games are presented. Conditions for existence and uniqueness of equilibrium are given in some class of games including submodular games, supermodular games, potential games, games with monotone payoffs etc. We also discuss about selection of equilibria using population dynamics.

We study the following applications:

- *Power control in Wideband CDMA wireless networks:* Wideband Code Division Multiple Access (W-CDMA) is wideband spread-spectrum channel access method that utilizes the direct-sequence spread spectrum method of asynchronous code division multiple access to achieve higher speeds and support more users compared to most time division multiple access (TDMA)(144). Power control in wireless networks has become an important research area. Since the technology in the current state cannot provide batteries which have small weight and large energy capacity, the design of tools and algorithms for efficient power control is crucial. For a comprehensive survey of recent results on power control in wireless networks an interested reader can consult e.g., (144) and the reference therein. Power control protocols based on game theory have been designed for already ten years starting with the pioneering work (76; 115). Non-cooperative games provide a convenient framework for decentralization and distributed decision making in those applications, where as cooperative approaches of game theory have allowed to handle issues concerning fairness in power allocation. Most applications of game theory to power control consider mobile terminals as players of the same type and study strategic one-shot games with a fixed number of players. Here, we consider a population game with many simultaneous local interactions, where the game is played infinitely under some self-organizing process called "hybrid dynamic" and where each interaction concerns a random number of players (213; 11).

- *OFDMA based IEEE 802.16 Networks:* OFDMA (Orthogonal Frequency Division Multiple Access) is recognized as one of the most promising multiple access technique in wireless communication system. This technique is used to improve spectral efficiency and becomes an attractive multiple access technique for 4th generation mobile communication system as WiMAX.

  In OFDMA systems, each user occupies a subset of subcarriers, and each carrier is assigned exclusively to only one user at any time. This technique has the advantage of eliminating intra-cell interference (interference between subcarriers is negligible). Hence the transmission is affected by intercell interference since users in adjacent sectors may have also been assigned to the same carrier. If those users in the adjacent sectors transmitted with high power the intercell interference may severely limit the SINR achieved by the user. Some form of coordination between the different cells occupying the spectral resource are studied in (134; 123). The optimal resource allocation requires complete information about the network in order to decide which users in which cells should transmit simultaneously with a given power. All of these results however, rely on some form of centralized control to obtain gains at various layers of the communication stack. In a realistic network as WiMAX, centralized multicell coordination is hard to realize in practice, especially in fast-fading environments. Decentralized schemes have proposed to the intercell interference in OFDMA-based IEEE 802.16 Networks.

- *Aloha-like protocols:* Random Medium Access Control (MAC) algorithms have played an increasingly important role in the development of wired and wireless networks and the performance and stability of these algorithms, such as slotted-Aloha, Carrier Sense Multiple Access (CSMA) is still an open problem (36). Distributed Medium Access Control, starting from the first version of Abramson's Aloha to the most recent algorithms used in IEEE802.11, have enabled a rapid growth of both wired and wireless networks. They aim

at efficiently and fairly sharing a resource among users even though each user must decide independently (eventually after receiving some messages or listening) when and how to attempt to use the resource. MAC algorithms have generated a lot of research interest, especially recently in attempts to use multi-hop wireless networks to provide high-speed access to the internet with low-cost and low-energy consumption. we focus our attention to wireless networks, where the resources are receivers, base station or access points and where users interact because of interference, i.e., interfering users cannot transmit simultaneously. There is a collision if another user (mobile) transmits with a greater power level at the same range of the receiver.

We apply evolutionary game with *random number* of interacting players to slotted Aloha wireless networks, power control in Wideband CDMA networks and OFDMA-based IEEE 802.16 networks. We then focus on evolutionary stability properties of correlated equilibrium in evolutionary games with arbitrary number of players in each local interaction. The performance and cost of random medium access control with power control is analyzed. This study is extended to the case of Signal-to-Interference-plus-Noise-Ratio (SINR) and Quality of Service (QoS) based Admission Control in which more than one user can have successful transmissions at the same time slot. A class of delayed evolutionary game dynamics for correlated equilibria is given based on assignment functions from the set of signals to set of actions. The chapter 4 presents a class of bio-inspired game dynamics with migration in a *hybrid evolutionary game model* in Code Division Multiple Access (CDMA) wireless data networks, in OFDMA-based WiMAX environment and in association problem between several types of technologies such as UMTS 3G, High Speed (Downlink/Uplink) Packet Access (HSDPA/HSUPA), IEEE 802.16 (WiMAX), IEEE 802.11 (WiFi), Long Term Evolution (LTE) etc. Considering both location of users (regions) and secondary actions association in each location, we transform the migration problem in term of revision of strategies. Each user has its own state (location, specific technology). An individual state has two components: the current location of the user and the current secondary action. In contrast to the standard formulations in which users can not change their class or type, we consider here that the class which corresponds to its location can be changed in time. This captures the mobility of users in the population. We show that the order of revision of strategies is important in the leading dynamics (speed of convergence, Nash stationarity, positive correlation etc) .

Part II: Stochastic Population Games

The second part of this manuscript focuses on stochastic population games of evolving games. The chapter 5 introduces *stochastic population games*. This class of games is described by : (i) classes of large population of players, (ii) each player from each class has its own state and a set of actions available in each state, (iii) each player interacts with other randomly selected players (in the same class or from other classes). The states and actions of each player in an interaction together determine the instantaneous payoff for all involved players. They also determine the transition probabilities to move to the next state. Each player wishes to maximize the expected payoff during its sojourn time in the system.

- *Energy management in hybrid Aloha-like networks* We study energy management in a distributed Aloha network with large number of mobile terminals. Each mobile can choose both the channels and powers (this is in contrast to standard Aloha model in which users are associated to the closest receivers). Each terminal is faced to a random number of interacting players which transmit at the channel. A terminal attempts transmissions during a finite horizon of times depending on the state of its battery energy. At each slot, each terminal have to take a decision on the transmission power based on the battery state. At each state of the battery, there are a finite power levels. At the lowest state of battery no power is available and the mobile have to replaced the battery by a new or to recharge

its battery. A transmission is successful if no other user transmit during the slot or the mobile transmits with a power which is bigger than the power of all others transmitting mobiles at the same receiver. We model each battery state as a Markov decision process. We analyze this problem stochastic population games with individual states.

- **Access control in solar-powered broadband networks**

  Environmental energy is becoming a feasible alternative for many low-power systems, such as wireless sensor and mesh networks. However, this provides an unpredictable and limited amount of energy over time. The power storage elements, such as rechargeable batteries or super-capacitors, become very useful to increase the system lifetime and the system availability. In particular, solar power is made possible with the use of Photovoltaic cells. Comprised of several layers of material, these cells are able to produce electrical power from exposure to sunlight. Since in many geographic areas, nice weather is not guaranteed and is unpredictable, the nodes should be able to recover from blackout periods caused by the unavailability of energy.

  We extend our study to the case where the mobiles use power storage element, such as rechargeable Solar-powered batteries, in order to have energy available for later use. By modeling the energy-level of Solar-powered batteries as a stochastic process, we study noncooperative interactions within large population of mobiles that interfere with each other through many local interactions. Each local interaction involves a random number of mobiles. The actions taken by a mobile determine not only the immediate payoff but also the state transition probabilities of its battery. We define and characterize the evolutionary stable strategies of the stochastic population game.

  The stochastic population game with dynamic rechargeable battery depending on the weather (solar energy) is described as follows: There are many local interactions among individuals belonging to large populations of mobiles. The result of the interaction between mobiles depends on their current individual state. From time to time the individual state of a mobile varies. The action choice of mobiles involved in a local interactions as well as their individual states determine the not only the result of the interactions but also the transition probabilities to the other possible individual states. Each individual is thus faced with an Markov Decision Process (MDP) in which it maximizes the expected average cost criterion. Each individual knows only the state of its own MDP, and does not know the state of the other mobiles it interacts with.

  The destination of some transmission occasionally may receive simultaneously a transmission from another terminal which results in a collision. It is assumed however that even when packets collide, one of the packets can be received correctly if transmitted at a higher power. As state of the MDP of a user we take its energy level. The immediate fitness (rewards) is the number of successful transmissions. By allowing the mobiles to be equipped with rechargeable solar powered batteries, the mobiles may have infinite life time and the criteria that is maximizing is the limit average Cesaro-type payoff.

Part III: Mean Field Limits

The third part of this manuscript focuses on a class of mean field games and asymptotic properties of stochastic games. An interesting and important problem in $n-$player games is that of determining what happens when the number of players increases. Mean field asymptotics have been proposed to analyze this problem and a connection between the equilibria of limiting game and finite game is obtained. In addition, the dynamics of the system evolution is characterized. Those dynamics extend the standard evolutionary game dynamics. The chapter 6 studies the transition between evolutionary games with finite number of players and evolutionary games with infinite number of players. When the size of the population goes

to infinity, a new class of deterministic population dynamics is given and sufficiency conditions for convergence and evolutionary stability are derived. A fluid model of spatial random access games is presented. The chapter 7 provides a rigorous derivation of the asymptotic behavior of the stochastic evolving games as the size of the population grows to infinity. Under restricted class of strategies, the random process consisting of one specific player and the remaining population converges weakly to a jump process driven by the solution of a system of differential equations. The large population asymptotic of the microscopic model is equivalent to a macroscopic stochastic population game in which a local interaction is described by a single player against an evolving population profile. We illustrate our model to derive the equations for a dynamic evolutionary Hawk and Dove game with energy level. We derive a new class of dynamic games called *differential population games*. Different from standard differential game models, this class of games is described by large population of players in which each player from each class is facing a random vector that evolves according to the population dynamics (deterministic or stochastic differential equations/inclusions). We derive equilibrium and optimality solutions using Hamilton-Jacobi-Bellman-Issacs equations. This research is of importance for dealing with complexity in stochastic dynamic optimization of large-scale systems and has methodological implications in many complex systems arising in engineering and socio-economic areas, ecology and evolutionary biology.

Part I

Delayed Evolutionary Game Dynamics with Migration

Chapter 2

Delayed evolutionary game dynamics

In this chapter, we study some evolutionary games where competition between individuals from a large population occurs through many pairwise interactions between randomly selected individuals. There are two features that make evolutionary games attractive to networking and communication systems. Its *evolutionarily stability* concept is better adapted to large populations of players as it describes robustness against deviations of a whole fraction of the population. The second appealing element of evolutionary games, the evolutionary game dynamics, describes the evolution of strategies in time. We study in particular an evolutionary game describing competition between mobile terminals over the access to a channel and we revisit the well known Hawk and Dove game (or Chicken game). In addition to the study of properties of the ESS in these games, we study the effect of time delays on the convergence of various evolutionary dynamics to the ESS. Delays play an important role in many situations in networking such as in flow control and congestion problem. We assume that time delays are not necessarily symmetric. The case of a common time delay for all strategies can be found in Tao & Wang (1997). We provide instability conditions and illustrate the non-stable behavior with numerical examples.

Several previous papers have already studied evolutionary games with pairwise local interactions in the context of wireless networks. Evolutionary games have been studied in the context of unslotted Aloha in (52). They have identified conditions for the existence of non trivial ESS and have computed them explicitly. In (215), we have considered the multiple access game and studied delay effect under various models of evolutionary game dynamics with asymmetric delay based on the theoretic results on stability obtained in (214). In (216), we have extended the model of (215) by including a regret cost, incurred when no user transmits, and studied the impact of that cost on the proportion of mobiles that transmit at equilibrium. In the last three papers, the delay is shown to have negative impact on the stability of the system. For other applications of evolutionary games concepts in networking, see (147; 240; 12; 211; 13) who study congestion control models. The time delay is shown to have negative impact on the stability of the system and "*evolutionarily stable state*" can be *unstable*.

We introduce below asymmetric time delays into classical models of populations dynamics. An action taken today will have its effect some time later. The delays can be symmetric or not. We then obtain delayed payoff (or payoff, fitness, cost) functions. The evolution of the system leads to delayed evolutionary game dynamics. Delayed evolutionary game dynamics

are in general a system of first order non-regular nonlinear differential equations or differential inclusions with time delays. We use the theory of delayed differential equation (DDE) to study stability, convergence and non-convergence of the system under time delays. Evolutionary game dynamics with asymmetric time delays have been introduced in (214; 216; 215; 13). In the case of finite number classes where each class has a finite number of pure actions, a simple class of delayed evolutionary game dynamics is given by

$$\dot{x}_a^e(t) = \sum_{b \in \mathcal{A}^e} x_b^e(t)\beta_a^{e,b}(x(t), \{x(t-\tau_b^e)\}_{e,b}) - x_a^e(t) \sum_{b \in \mathcal{A}^e} x_b^e(t)\beta_b^{e,a}(x(t), \{x(t-\tau_b^e)\}_{e,b}) \qquad (2.1)$$

where τ_a^e denotes the time delay associated to the pure strategy a of class e. See chapter 6 for a derivation of this expression from mean field interactions or from Kolmogorov forward equations.

We now consider simple pairwise interactions. For such games there is a direct relation to a two players matrix game representing the expected fitness obtained in an interaction between two individuals (selected at random); the expectation is with respect to the fraction of the population that uses each strategy. We introduce asymmetric time delays in evolutionary game dynamics and study asymptotic stability of evolutionary stable strategies under these dynamics. We give sufficiency conditions of stability of the ESS which can not invaded by small proportion of deviations. Using delay differential equations properties, we show that the interior ESS of generic two-players games is asymptotically stable for small delays (Theorems 2.1.5, 2.1.6) and apply this result in the Hawk and Dove Game. We show that for a small time delay of the hawk strategy the system is asymptotically stable independently the Dove's time delay. In the common access problem in wireless adhoc networks, we show that the game has a unique ESS in which nodes which transmit and nodes which does not transmit can co-exist. We show how the evolution and the ESS are influenced by the characteristics of the wireless channel and pricing:

- We give a necessary and sufficient stability condition (Propositions 2.2.2.1 and 4.3.3) which depend on the regret cost in the multiple access game,

- We show that ESS is stable for small transmission delay under BNN and replicator dynamics for the regret cost lower to unit. The ESS becomes unstable for large delays.

- We show that the fraction of transmitters of the population of mobiles oscillates around the ESS for all small regret cost under imitate the better dynamics (Proposition 2.2.3.1 and remark 2.2.3.2).

The rest of this chapter is organized as follows. In Section 5.4, we describe our model of population game with delayed payoffs and we give an example of channel access problem in which delays must be take in consideration the interactions and we analyze delays impact on stability of the ESS in many dynamics. Multiple access game is studied in Section 2.2. We give some numerical examples of oscillatory regime in the access game. Section 2.4 extends delayed evolutionary game dynamics to asymmetric population games which one can have simultaneously intra-population interactions and inter-population interactions.

2.1 Setting and notations

We consider the standard setting of evolutionary games : there is a large populations of players; each member of the population has the same pure strategy set $\mathcal{A}$ (a finite set). There are many pairwise interactions. Denote by $\Delta(\mathcal{A})$ the $(|\mathcal{A}|-1)-$dimensional simplex of $\mathbb{R}^{|\mathcal{A}|}$.

2.1.1 Delays and Fitnesses

The fitness for a player at a given time is determined by the action a taken by the player at that time, as well as by the actions of the population it interacts with, that was taken τ_a units ago. More precisely, If player 1 chooses the strategy b and player 2 the strategy a then player 1 receives the payoff $u(b, a)$ (called fitness) only τ_b times later and player 2 receives the payoff $u(a, b)$ after a delay of τ_a.

Below we denote by $f_a(x(t - \tau_a))$ the fitness for an individual using the strategy a at time t when it encounters an individual that used strategy $x(t - \tau_a)$ at time $t - \tau_a$. We have $f_a(x(t - \tau_a)) = \sum_{b \in \mathcal{A}} u(a, b) x_b(t - \tau_a)$, and we set $\bar{f}(x(t)) = \sum_{a \in \mathcal{A}} x_a(t) f_a(x(t - \tau_a))$ to be the average expected fitness of the whole population at time t.

The following example in the context of wireless communications illustrates our model.

Example

Consider a transmitter A and a very close receiver B. Transmitter A can decide whether to transmit with frequency f_1 or f_2. If it transmits with frequency f_i then its signal receives interference at receiver B from transmitters that are located at a distance of D_i from the receiver, which translates to a time delay of τ_i (where $i = 1, 2$). Thus indeed, the action of a player determines the delay it will take for actions of other players to be effective.

2.1.2 Delayed evolutionary game dynamics

Evolutionary game dynamics are models of strategy change commonly used in evolutionary game theory. A strategy which does better than the average or its opponent, increases in frequency at the expense of strategies that do worse than the average or the opposed action. Many evolutionary game dynamics models are used in the literature (see ([168](); [99]()) and the references therein) : replicator dynamics, best response dynamics, Brown-von Neumann-Nash dynamics, Smith dynamics, gradient dynamics, projection dynamics etc. We introduce below delays into classical models of populations dynamics. Let β_a^b conditional switch rate from the strategy b to the strategy a. The flow of the population is specified in terms of the functions β_a^b which determine the rates at which an player who is considering a change in strategies opts to switch to his various alternatives. The *incoming flow* of the action a is $\sum_{b \in A} x_b(t) \beta_a^b(x(t))$, and the *outgoing flow* is $x_a(t) \sum_b \beta_b^a(x(t))$ where $x_a(t)$ represents the fraction of players of the population which use the action a at time t. Let

$$W_f^a(x(t), \{x(t - \tau_k)\}_k) = \sum_{b \in \mathcal{A}} x_b(t) \beta_a^b(x(t), \{x(t - \tau_k)\}_k) - x_a(t) \sum_{b \in \mathcal{A}} \beta_b^a(x(t), \{x(t - \tau_k)\}_k).$$

The evolutionary game dynamics ([234](); [168](); [208]()) is given by

$$\dot{x}_a = W_f^a(x(t), \{x(t - \tau_k)\}_k), \ a \in \mathcal{A}. \tag{2.2}$$

For example the *delayed replicator dynamics*, delayed Brown-von Neumann-Nash dynamics and generalized projection dynamics are obtained respectively for

$$\beta_a^b(x(t), \{x(t - \tau_k)\}_k) = x_a(t) \max(0, f_a(x(t - \tau_a)) - f_b(x(t - \tau_b))),$$

$$\beta_a^b(x(t), \{x(t - \tau_k)\}_k) = \max(0, f_a(x(t - \tau_a)) - \sum_{b \in \mathcal{A}} x_b(t) f_b(x(t - \tau_b))),$$

$$\beta_a^b(x(t), \{x(t - \tau_k)\}_k) = h_a(x(t), \{x(t - \tau_k)\}_k).$$

Delayed replicator dynamics

In the delayed replicator dynamics, the share of a strategy in the population grows at a rate equal to the difference between the delayed payoff of that strategy and the average delayed payoff of the population. The fitness acquired at time t will impact the rate of growth τ_a time later. Thus the delayed replicator dynamics of $x_a(t)$ is given by

$$a \in \mathcal{A}, \ \dot{x}_a(t) = x_a(t) \left[f_a(x(t - \tau_a)) - \sum_{b \in \mathcal{A}} x_b(t) f_b(x(t - \tau_b)) \right].$$

The standard replicator dynamics is obtained when all the delays are zero.

Delayed BNN dynamics

Delays can be introduced in BNN dynamics, Brown & von Neumann (1950) as follows

$$a \in \mathcal{A}, \ \dot{x}_a(t) = g_a(x(t)) - x_a(t) \sum_{b \in \mathcal{A}} g_b(x(t)) \tag{2.3}$$

where

$$g_a(x(t)) = \max \left\{ 0, f_a(x(t - \tau_a)) - \sum_{b \in \mathcal{A}} x_b(t) f_b(x(t - \tau_b)) \right\}.$$

The function $g_a(x(t))$ is the positive part of the excess payoff for strategy a when the population state is $x(t)$ at time t.

There are a close relation between BNN dynamics with Nash's original proofs of his equilibrium theorem (see Hofbauer & Sigmund (2003)). BNN dynamics is interpreted in Hofbauer & Sigmund (2003) as follows: Suppose there are large users in the population in which there is steady influx and outflux. New users joining the system use only strategies that are better than average, and better strategies are more likely to be adopted. On the other hand, randomly chosen users leave the game. More precisely, the strategy $x \in \Delta(\mathcal{A})$ is adopted with probability proportional to the excess payoff $g_a(x)$.

Delayed imitate the better dynamics

We introduce delayed myopic imitation dynamics. We suppose that users in the population review their strategy and imitate the better's strategy on the time (imitation by dissatisfaction). In a symmetric two player game, the delayed dynamic is given by

$$b \in \mathcal{A}, \ \dot{x}_b(t) = x_b(t) \left(\sum_{a \in \mathcal{A}} x_a(t) \left[\rho_{ab}(x(t - \tau_a)) - \rho_{ba}(x(t - \tau_b)) \right] \right), \ b \in \mathcal{A} \tag{2.4}$$

where $\rho_{ab}(x) = g(f_b(x(t - \tau_b)), f_a(x(t - \tau_a)))$, $a, b \in \mathcal{A}$ and $g(a, b) = 0$ if $a < b$ and $g(a, b) = 1$ if $a > b$.

Delayed best response dynamics

A strategy $Y \in \Delta(\mathcal{A})$ is a best response to the trajectory $x(t)$ at time t if $\sum_{a \in \mathcal{A}} Y_a f_a(x(t - \tau_a)) \geq \sum_{a \in \mathcal{A}} v_a f_a(x(t - \tau_a))$, $\forall v \in \Delta(\mathcal{A})$. We note by $BR(x(t))$ the best response strategies set to the strategy $x(t)$ i.e

$$BR(x(t)) = \left\{ Y \in \Delta(\mathcal{A}), \sum_{a \in \mathcal{A}} Y_a f_a(x(t - \tau_a)) \geq \sum_{a \in \mathcal{A}} v_a f_a(x(t - \tau_a)), \forall v \in \Delta(\mathcal{A}) \right\}$$

A fraction of population revise their strategy and choose the best replies $BR(x(t))$ to the current population state $x(t)$ at time t. The best response dynamics is given by $\dot{x} \in BR(x) - x$.

Delayed logit dynamics

The delayed logit dynamics (called also delayed exponential weight dynamics or delayed Boltzman dynamics) is given by

$$a \in \mathcal{A}, \ \dot{x}_a(t) = \frac{e^{\frac{1}{\epsilon} f_a(x(t - \tau_a))}}{\sum_{b \in \mathcal{A}} e^{\frac{1}{\epsilon} f_b(x(t - \tau_b))}} - x_a(t) \tag{2.5}$$

It can be shown that the delayed logit dynamics given by

$$a \in \mathcal{A}, \ \dot{x}_a(t) = \frac{e^{f_a(x(t - \tau_a)) T_0}}{\sum_{b \in \mathcal{A}} e^{f_b(x(t - \tau_b)) T_0}} - x_a(t) \tag{2.6}$$

converge to the best response dynamics if $T_0 \rightarrow +\infty$. See Hofbauer and Sigmund (2003).

Delayed ray-projection dynamics

The Ray-projection dynamics with time delays is given by

$$a \in \mathcal{A}, \ \dot{x}_a(t) = h_a(x(t), \{x(t - \tau_b)\}_b) - x_a(t) \sum_{b \in \mathcal{A}} h_b(x(t), \{x(t - \tau_b)\}_b) \tag{2.7}$$

where g is a vectorial function of the the population profile and the vector of expected payoffs.

Delayed orthogonal projection dynamics

The orthogonal-projection dynamics with time delays is given by

$$a \in \mathcal{A}, \ \dot{x}_a(t) = f_a(x(t - \tau_a)) - \frac{1}{n} \sum_{b \in \mathcal{A}} f_b(x(t - \tau_b)) \tag{2.8}$$

where g is a vectorial function of the the population profile and the vector of expected payoffs.

Delayed target projection dynamics

$$\dot{x}(t) = \text{proj}_{TX}\left(f(\{x(t - \tau_k)\}_k)\right) - x(t) \tag{2.9}$$

Definitions

Consider the following delay differential equations

$$\frac{dx_a}{dt} = \dot{x}_a(t) = \Theta_a(x_a(t), x(t - \tau_1), \ldots, x(t - \tau_{|\mathcal{A}|})), \ a \in \mathcal{A}, \ x(.) \in \Delta(\mathcal{A}) \tag{2.10}$$

Denote by τ the maximum of τ_a, $a \in \mathcal{A}$

1. The state x^* is stationary state (or rest point) of the differential equation (2.10) if it is a critical point i.e the right side of (2.10) is zero at x^*.

2. x^* is stable if it is a stationary point with the property that for every neighborhood $\mathcal{V}$ of x^*, there exists a neighborhood $\mathcal{U} \subset \mathcal{V}$ with the property that if $x(t) \in \mathcal{U}$ for $t \in (-\tau, 0)$ then $x(t) \in \mathcal{V}$ for all $t > 0$.

3. x^* is asymptotically stable if it is stable and there exists a neighborhood $\mathcal{W}$ of x^* such that $x(t) \in \mathcal{W}$ for all $t \in (-\tau, 0)$ implies $\lim_{t \to +\infty} x(t) = x^*$

4. x^* is exponentially stable if it is stable and there exists $t_0, L, \eta > 0$ such that

$$\forall \ t \geq t_0, \ |x(t) - x^*| \leq Le^{-\eta t}.$$

From these definitions, it follows that exponential stability implies asymptotic stability implies stability which implies stationary point of system.

2.1.3 Stability analysis of the delayed replicator dynamics

In this subsection, we introduce replicator dynamics for a population and $|\mathcal{A}|$ asymmetric delays with $|\mathcal{A}| \geq 2$.

$$\dot{x}_a(t) = x_a(t)\left[f_a(x(t - \tau_a)) - \sum_{b \in \mathcal{A}} x_b(t) f_b(x(t - \tau_b))\right], \ a \in \mathcal{A}. \tag{2.11}$$

Suppose that equation (2.11) has an interior stationary point x^*. It is known Gopalsamy (1992, page 188) or from the Hartman and Grobman theorem adapted to delay differential equation that the steady state x^* is asymptotically stable for (2.11) around the stationary point x^* if the trivial solution of the linearized version is asymptotically stable.

The linearized equation of (2.11) at the stationary point x^* is given by

$$\dot{z}_a = x_a^*\left[f_a(z(t - \tau_b)) - \sum_{b \in \mathcal{A}} x_b^* f_b(z(t - \tau_b))\right], \ a \in \mathcal{A} \tag{2.12}$$

where $z(t) = x(t) - x^*$. This equation can be written as

$$\dot{z}(t) = K \sum_{l \in \mathcal{A}} B^l z(t - \tau_l)$$

where $z(t) = {}^t(z_1(t), \ldots, z_{|\mathcal{A}|}(t))$ and $B^l = \left(b_{ik}^l \right) \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{A}|}$,

$$b_{ik}^l = \begin{cases} x_i^*(1 - x_i^*)u_{ik} & \text{if } l = i \\ -x_i^* x_l^* u_{ik} & \text{if } l \neq i. \end{cases}$$

Characteristic equation

$$\det \left(\lambda I - K \sum_{l \in \mathcal{A}} B^l e^{-\tau_l \lambda} \right) = 0. \tag{2.13}$$

A necessary and sufficient condition of stability of (2.12) is that all roots of the equation (3.7) have negative real parts.

Two actions

Suppose that each individual of the population only uses a pure strategy. Denote $\xi(t) = x_1(t)$ the proportion of individuals in the population using the first strategy at time t, then replicator dynamic of $\xi(t)$ is given by

$$\dot{\xi}(t) \quad = -K\delta\xi(t)(1 - \xi(t)) \times \left[\frac{(u_{12} - u_{11})}{\delta} \xi(t - \tau_1) + \frac{(u_{21} - u_{22})}{\delta} \xi(t - \tau_2) - \xi^* \right] \tag{2.14}$$

where $\delta = (u_{21} - u_{11}) + (u_{12} - u_{22})$, $\xi^* = \dfrac{u_{12} - u_{22}}{\delta}$. If the payoff matrix

$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$$

satisfies

$$\left(u_{12} > u_{22}, \ u_{21} > u_{11} \right) \text{ or } \left(u_{12} < u_{22}, \ u_{21} < u_{11} \right) \tag{2.15}$$

then $0 < \xi^* < 1$ and ξ^* is a unique interior stationary point so called fixed point, rest point or equilibrium point.

Theorem 2.1.4. *Under the first condition of (2.15):* $\left(u_{12} > u_{22}, \ u_{21} > u_{11} \right)$, ξ^* *is the unique ESS of the game.*

Proof. If the first condition of (2.15) is satisfied then $(\xi^*, 1 - \xi^*)$ is a symmetric mixed Nash equilibrium in the two player game with the payoff matrix U. Moreover, one has,

$$\forall \, mut \neq \xi^*, \quad (\xi^*, 1 - \xi^*)U\begin{pmatrix} mut \\ 1 - mut \end{pmatrix} - (mut, 1 - mut)U\begin{pmatrix} mut \\ 1 - mut \end{pmatrix} \tag{2.16}$$

$$= \delta(\xi^* - mut)^2 \tag{2.17}$$

Hence $(\xi^*, 1 - \xi^*)$ is an ESS. $\qquad \square$

The linearized version of (2.14) at the stationary point x^* is given by

$$\dot{z}(t) = -\gamma \left[\alpha z(t - \tau_1) + \beta z(t - \tau_2) \right] \tag{2.18}$$

where $\gamma = K\xi^*(1 - \xi^*)$, $\alpha = u_{12} - u_{11}$, $\beta = u_{21} - u_{22}$.

Let

$$\dot{z}(t) = -az(t - \tau_1) - bz(t - \tau_2) \tag{2.19}$$

where $a, b \in \mathbb{R}$ and τ_1, τ_2 are positive. The characteristic equation of (2.19) is given by

$$s = -ae^{-\tau_1 s} - be^{-\tau_2 s} \tag{2.20}$$

Gopalsamy (1992) proved that when a, b, τ_1, τ_2 are positive, a sufficient condition for all roots of (2.20) to have negative real parts is $a\tau_1 + b\tau_2 < 1$ and a necessary condition for the same is $a\tau_1 + b\tau_2 < \dfrac{\pi}{2}$. The following theorem give sufficient conditions of stability of (2.19) at zero.

Theorem 2.1.5 (Berezansky& Braverman(2006), Li et al.(1999))**.** *Suppose at least one of the following conditions holds*

1. $a + b > 0$, $|a|\tau_1 + |b|\tau_2 < \dfrac{a + b}{|a| + |b|}$

2. $a > 0$, $a\tau_1 < \dfrac{a - |b|}{a + |b|}$

3. $b > 0$, $b\tau_2 < \dfrac{b - |a|}{|a| + b}$

Then equation (2.19) is exponentially stable.

Theorem 2.1.6. *Suppose that $0 < b < a$ and $\tau_1 < \dfrac{1}{a + b}$. Then, all roots of (2.20) have negative real parts (the equation (2.19) is asymptotically stable).*

Proof. By hypothesis, one has $A = \dfrac{b}{a} \in (0, 1)$ and $r_1 := a\tau_1 < \dfrac{1}{1 + A}$. Equation (2.20) can rewritten as

$$\lambda = -e^{-r_1\lambda} - Ae^{-r_2\lambda} \tag{2.21}$$

where $\lambda = \dfrac{s}{a}$, $r_2 = a\tau_2$. Since all roots of equation (2.21) have negative real parts when $r_1 = 0$, if the conclusion fails, then there must be exist some $r_1 \in (0, \dfrac{1}{1 + A}]$ such that equation (2.21) has purely imaginary roots $\pm iy$, $y > 0$ satisfying

$$\cos(r_1 y) = -A\cos(r_2 y) \text{ and } y - \sin(r_1 y) = A\sin(r_2 y) \tag{2.22}$$

One has

$$A^2 - 1 = y^2 - 2y\sin(r_1 y)$$

that is

$$f(y) := \frac{y^2 + 1 - A^2}{2y} = \sin(r_1 y) \tag{2.23}$$

Since $-1 \leq \sin(r_1 y) \leq 1 \Rightarrow y \in (1 - A, 1 + A)$.

$$f(y) = \frac{y}{2}\left(1 + \frac{1 - A^2}{y^2}\right) \geq \frac{y}{2}\left(1 + \frac{1 - A^2}{(1 + A)^2}\right) \geq y\frac{1}{1 + A} > r_1 y \geq \sin(r_1 y) \tag{2.24}$$

a contradiction with (2.23). □

Corollary 2.1.6.1. *If one of the following conditions holds, then ξ^* is stable*

1. $\alpha + \beta = \delta > 0, \ K\left(|\alpha|\tau_1 + |\beta|\tau_2\right) < \dfrac{\delta}{\xi^*(1 - \xi^*)\left(|\alpha| + |\beta|\right)}$

2. $\alpha > 0, \ K\tau_1 < \dfrac{\alpha - |\beta|}{\alpha\xi^*(1 - \xi^*)\left(\alpha + |\beta|\right)}.$

3. $\beta > 0, \ K\tau_2 < \dfrac{\beta - |\alpha|}{\beta\xi^*(1 - \xi^*)\left(|\alpha| + \beta\right)}.$

4. $0 < \beta < \alpha, \ K\tau_1 < \dfrac{1}{\delta\xi^*(1 - \xi^*)}$

The case of symmetric delay

Consider the following linear delay differential equation

$$z(t) = -\omega z(t - \tau) \tag{2.25}$$

with $\tau, \omega > 0$ and its characteristic equation

$$\lambda + \omega e^{-\lambda\tau} = 0 \tag{2.26}$$

Necessary and sufficient condition of asymptotic stability as a function of the delay is given by the following lemma. A proof can be found in Gopalsamy (1992, Proposition 1.2.8)

Lemma 2.1.6.2. *A necessary and sufficient condition for all roots of (2.26) to have negative real parts is* $2\omega\tau < \pi$

Suppose that $\tau_1 = \tau_2 = \tau$, then the trivial solution of (2.18) is asymptotically stable if

$$K\tau < \frac{\delta\pi}{2\delta_1\delta_2} \tag{2.27}$$

and unstable if $K\tau > \dfrac{\delta\pi}{2\delta_1\delta_2}$ where $\delta_1 = u_{21} - u_{11}$, $\delta_2 = u_{12} - u_{22}$ and $\delta = \alpha + \beta = \delta_1 + \delta_2$.

The case $K\tau = \dfrac{\delta\pi}{2\delta_1\delta_2}$ is called bifurcation point. By using lemma 2.1.6.2 with $\omega = \gamma\delta$, we conclude that the dynamic (2.14) is asymptotically stable at the stationary point $\xi^* = \dfrac{\delta_2}{\delta}$ if K and τ satisfy (2.27).

2.2 Multiple Access Game with Regret Cost

Multiple Access Game introduces the problem of medium access. We assume that mobiles are randomly placed over a plane. All mobiles use the same fixed transmission range of r. The channel is ideal for transmission and all errors are due to collision. A mobile decides to transmit a packet or not to transmit to a receiver when they are within transmission range of each other. Interference occurs as in the ALOHA protocol: if more than one neighbors of a receiver transmit a packet at the same time then there is a collision. The Multiple Access Game is a nonzero-sum game, the mobiles have to share a common resource, the wireless medium, instead of providing it. We suppose that a mobile has a receiver in its range with probability K. When a mobile m

transmits to its receiver $R(m)$, all mobiles within a circle of radius r centered at the receiver $R(m)$ cause interference to the node m for its transmission to receiver $R(m)$. This means that more than one transmission within a distance r of the receiver in the same slot result in a collision of all packets at the receiver. Each of the mobiles has two possible strategies: either to transmit (T) or to stay quiet (S). If mobile m transmits a packet, it incurs a transmission cost of $\delta \geq 0$. The packet transmission is successful if the other mobiles don't transmit (stays quiet) in that given time slot, otherwise there is a collision, and the cost (of collision risk) is $\Delta \geq 0$. If there is no collision, mobile m gets a reward of V from the successful packet transmission after the τ_T times later. We assume that the reward V is greater than the cost of transmission Δ. When all mobiles stay quiet, they have to pay a regret cost κ. The regret cost κ describes the behavior of mobiles when they are aware of the backoff delays. The interaction between two mobiles is represented in Figure 3.27. The value of successful transmission is normalized to unit.

	T	S
T	$-\Delta K$	$K(V - \Delta)$
S	0	$-\kappa K$

Figure 2.1: *Access game between two mobiles.*

2.2.1   Nash Equilibria, ESS and Pareto optimality

An allocation of payoffs is said Pareto-optimal if the outcome cannot be improved upon without hurting at least one user. This matrix game has two pure Nash equilibria (T, S) and (S, T) and a unique mixed Nash equilibrium given by $(\dfrac{1 - \Delta + \kappa}{1 + \kappa}, \dfrac{\Delta}{1 + \kappa})$. Note that the pure strategies (T, S) and (S, T) are also optimal in the Pareto sense. The mixed equilibrium converges the pure strategy T when the regret cost is large.

Proposition 2.2.1.1. *The mixed strategy $(\dfrac{1 - \Delta + \kappa}{1 + \kappa}, \dfrac{\Delta}{1 + \kappa})$ is the unique ESS of the multiple access game.*

Proof. This comes from the theorem 2.1.4. The strategy $(\dfrac{1 - \Delta + \kappa}{1 + \kappa}, \dfrac{\Delta}{1 + \kappa})$ is the unique interior Nash equilibrium. It is the unique symmetric Nash equilibrium. Thus, it is the only candidate to be ESS because symmetric Nash equilibria set contains ESS set. At a mixed equilibrium, the strategies T and S must have the same fitness. We check the condition of ESS given in (1.3). For all $\xi \neq \dfrac{1 - \Delta + \kappa}{1 + \kappa}$

$$\left(\frac{1 - \Delta + \kappa}{1 + \kappa} - \xi, \xi - \frac{(1 - \Delta + \kappa)}{1 + \kappa} \right) A \left(\begin{array}{c} \xi \\ 1 - \xi \end{array} \right)$$

$$= \frac{K}{1 + \kappa} \left((1 + \kappa)\xi - 1 + \Delta - \kappa \right)^2 > 0$$

We conclude that $(\dfrac{1 - \Delta + \kappa}{1 + \kappa}, \dfrac{\Delta}{1 + \kappa})$ an ESS. $\qquad \square$

Note that at the ESS, the two subpopulations obtain the same negative payoff equal to $-K\dfrac{\Delta}{1 + \kappa}$.

2.2.2 Stability bound for ESS under replicator dynamics

If $(\xi(t), 1 - \xi(t))$ denotes de repartition of the population in the multiple access game, then the fitness of the subpopulation using the strategy T at time t is $K(-\xi(t - \tau_T) + 1 - \Delta)$, and the fitness of the strategy S is $-\kappa K(1 - \xi(t))$. The replicator dynamic equation becomes

$$\dot{\xi}(t) = -K\xi(t)(1 - \xi(t)) \left[\xi(t - \tau_T) + \kappa\xi(t) - 1 + \Delta - \kappa \right] \tag{2.28}$$

and its linearized version around the stationary point ξ^*,

$$\dot{z}(t) = -\frac{K\xi^*(1 - \xi^*)}{1 + \kappa} \left[\kappa z(t) + z(t - \tau_T) \right], \ \tau_T > 0, \tag{2.29}$$

is asymptotically stable if and only if all roots of the characteristic equation

$$\lambda + \frac{K\xi^*(1 - \xi^*)}{1 + \kappa} \left[\kappa + e^{-\tau_T \lambda} \right] = 0 \tag{2.30}$$

Proposition 2.2.2.1. *If $\kappa \geq 1$ then all roots of the characteristic equation (2.30) has negative real parts.*

Proof. • Suppose there exists a root λ of the characteristic equation (2.30) with $\Re(\lambda) > 0$ and $\kappa > 1$ then

$$\Re(\lambda) + K\xi^*(1 - \xi^*)\frac{\kappa}{1 + \kappa} > K\xi^*(1 - \xi^*)\frac{\kappa}{1 + \kappa} > \frac{K\xi^*(1 - \xi^*)}{1 + \kappa} \tag{2.31}$$

$$> -\frac{K\xi^*(1 - \xi^*)}{1 + \kappa}e^{-\tau\Re(\lambda)}\cos(\tau\Im(\lambda)) \tag{2.32}$$

a contradiction with (2.30). Thus all roots of the characteristic equation (2.30) have real parts strictly negative.

• Suppose that $\kappa = 1$.

$$\begin{cases} \Re(\lambda) + K\xi^*(1 - \xi^*)\dfrac{\kappa}{1 + \kappa} = -K\xi^*(1 - \xi^*)\dfrac{\kappa}{1 + \kappa}e^{-\tau\Re(\lambda)}\cos(\tau\Im(\lambda)) \\ \Im(\lambda) = K\xi^*(1 - \xi^*)\dfrac{\kappa}{1 + \kappa}e^{-\tau\Re(\lambda)}\sin(\tau\Im(\lambda)) \end{cases} \tag{2.33}$$

If $\Re(\lambda) > 0$ then $\Re(\lambda) + K\xi^*(1 - \xi^*)\dfrac{\kappa}{1 + \kappa} = -K\xi^*(1 - \xi^*)\dfrac{\kappa}{1 + \kappa}e^{-\tau\Re(\lambda)}\cos(\tau\Im(\lambda)) \leq K\xi^*(1 - \xi^*)\dfrac{\kappa}{1 + \kappa}$ i.e $\Re(\lambda) \leq 0$: impossible. If $\Re(\lambda) = 0$ then $\Im(\lambda) = 0$ and the trivial solution is stable.

$\square$

Proposition 2.2.2.2. *Suppose that $\kappa < 1$. Then all solutions of the characteristic equation (2.30) has negative real parts if and only if*

$$\tau \leq \frac{(1 + \kappa)^2}{K\Delta(1 - \Delta + \kappa)}\sqrt{\frac{1 + \kappa}{1 - \kappa}} \times \cos^{-1}(-\kappa) \tag{2.34}$$

Note that when κ goes to 1, the bound $\dfrac{(1 + \kappa)^2}{K\Delta(1 - \Delta + \kappa)}\sqrt{\dfrac{1 + \kappa}{1 - \kappa}}\cos^{-1}(-\kappa)$ goes to infinity (stability without conditions).

25

Proof.   • Suppose that $\tau < \tau_0 := \dfrac{(1+\kappa)^2}{K\Delta(1-\Delta+\kappa)}cos^{-1}(-\kappa)\sqrt{\dfrac{1+\kappa}{1-\kappa}}$. When $\tau = 0$ the system

(2.30) becomes $\Im(\lambda) = 0$ and $\Re(\lambda) = -K\dfrac{\Delta(1-\Delta+\kappa)}{(1+\kappa)^2} < 0$.

Let $b = K\dfrac{\Delta(1-\Delta+\kappa)}{(1+\kappa)^2}$, $a = \kappa b$. If (2.30) has a solution with real part strictly positive, then there must be some $0 < \tau_1 < \tau_0$ such that (2.30) has purely imaginary roots $\pm\Im(\lambda) \times i$ with $\Im(\lambda) > 0$ satisfying

$$\begin{cases} a = -b\cos(\tau_1 y) \\ y = b\sin(\tau_1 y) \end{cases} \Leftrightarrow \begin{cases} a = -b\cos(\tau_1 y) \\ y = b\sin(\tau_1 y),\ y > 0 \end{cases} \Leftrightarrow \begin{cases} y = \sqrt{b^2 - a^2} \\ \tau_1 = \tau_0 \end{cases}$$

We proved that $\tau_1 = \tau_0$ which is a contradiction. Thus, if $0 < \tau < \tau_0$ then all solutions of the characteristic equation (2.30) have negative real parts.

  • Suppose that $\tau > \tau_0$. Denote by $s = \lambda/b$, $r = \tau b > 0$, $r_0 = \tau_0 b$. The equation (2.30) can be written as the zeros the function

$$\phi(s,\kappa) := s + \kappa + e^{-rs}$$

We want to prove that all zeros of $\phi(.,\kappa)$ have negative real parts. The function ϕ is an analytic function in s and κ. Let

$$\mathcal{O} := \{s,\ \Re(s) \geq 0,\ |s| \leq 1\}$$

One has $r_0 \geq \dfrac{\pi}{2}$. Indeed,

$$\arccos(-p) - \dfrac{\pi}{2}\sqrt{1-p^2} > 0,\ p \in (0,1).$$

Thus when $r > r_0$ and $r \neq 2k\pi + \dfrac{\pi}{2}$, $k \in \mathbb{N}$ the function $\phi(s,0)$ has no zeros in the boundary of $\mathcal{O}$. Using Rouche's theorem $\phi(s,0)$ and $\phi(s,\kappa)$ have same sum of the orders of zeros. When $r > r_0$ the zeros of $\phi(s,0)$ is at least 2. Thus, $\phi(s,\kappa)$ has at least one root with positive real part.

□

2.2.3   Imitate the better dynamics

The imitate the better equation can be expressed as

$$\dot{\xi}(t) = \xi(t)(1-\xi(t))h(\xi(t-\tau),\xi(t))\text{ where}$$

$$h(\xi(t-\tau),\xi(t)) = \begin{cases} 1 & \text{if}\quad \xi(t-\tau) + \kappa\xi(t) < 1-\Delta+\kappa \\ 0 & \text{if}\quad \xi(t-\tau) + \kappa\xi(t) = 1-\Delta+\kappa \\ -1 & \text{if}\quad \xi(t-\tau) + \kappa\xi(t) > 1-\Delta+\kappa \end{cases}$$

Proposition 2.2.3.1. *Suppose that $\kappa < 1$. For all $\tau > 0$, and for all initial conditions $\forall\, t \in (-\tau, 0]$, $\xi(t) = \xi_0 \in (0, \dfrac{1-\Delta+\kappa}{1+\kappa})$ the solution of the imitation equation oscillates around the ESS.*

Proof. Consider the following differential equations

$$
\begin{cases}
\dot{\xi}(t) = \xi(t)(1 - \xi(t)) & \text{if} \quad \xi(t-\tau) + \kappa\xi(t) < 1 - \Delta + \kappa \\
\dot{\xi}(t) = 0 & \text{if} \quad \xi(t-\tau) + \kappa\xi(t) = 1 - \Delta + \kappa \\
\dot{\xi}(t) = -\xi(t)(1 - \xi(t)) & \text{if} \quad \xi(t-\tau) + \kappa\xi(t) > 1 - \Delta + \kappa \\
\forall\, t \in (-\tau, 0],\ \xi(t) = \xi_0 &
\end{cases}
$$

$$
(**) \quad
\begin{cases}
\dot{w}(t) = w(t)(1 - w(t)) \\
w(0) = \xi_0
\end{cases}
,\quad \xi_0 \in \left(0, \frac{1 - \Delta + \kappa}{1 + \kappa}\right).
$$

The solution of $(**)$ is given by $w(t) = 1 - \dfrac{1}{1 + \chi e^t}$ where $\chi = \dfrac{\xi_0}{1 - \xi_0}$. The function w goes to 1 when t goes to infinity. The functions ξ and w are equal in $(0, t_1)$ where $t_1 = t_r + t_*$ and t_* the first time such that $w(t) = \xi^*$ (this time exists because w is a continuous function, starts at ξ_0 which is lower than ξ^* and goes to 1),

$$
t_1 = \inf\{t > 0,\ \xi(t - \tau) + \kappa\xi(t) > 1 - \Delta + \kappa\}.
$$

We can explicitly compute t_* and t_1,

$$
t_* = \log\left(\frac{\xi^*}{\chi(1 - \xi^*)}\right),\quad
t_1 = \log\left(\frac{\nu + \sqrt{\nu^2 + 4\xi^*(1 - \xi^*)e^{-\tau}}}{2\chi(1 - \xi^*)e^{-\tau}}\right)
$$

where $\nu = -(1 - \xi^*)(1 + e^{-\tau}) + \left(\dfrac{1 + \kappa e^{-\tau}}{1 + \kappa}\right)$. One has $t_* + \tau > t_1 > t_*$ and $\xi_1 = \xi(t_1) = w(t_1) > \xi^*$. Let consider the equation

$$
\begin{cases}
\dot{v}(t) = -v(t)(1 - v(t)) \\
v(t_1) = \xi_1
\end{cases}
$$

Then, $\xi(t) = v(t) = \dfrac{1}{1 + \dfrac{e^t}{\chi e^{2t_1}}}$ on (t_1, t'_1) where

$$
t'_1 = \log\left(\chi e^{2t_1} \frac{\alpha + \sqrt{\alpha^2 + 4(1 - \xi^*)\xi^* e^{-\tau}}}{2e^{-\tau}\xi^*}\right),\quad
\alpha = -\xi^*(1 + e^{-\tau}) + \left(\frac{1 + \kappa e^{-\tau}}{1 + \kappa}\right).
$$

$\xi(t'_1) = v(t'_1)$ is lower than ξ^*. $\qquad\square$

Remarks 2.2.3.2. *Using the same arguments as in proposition 2.2.3.1, we extend the result to all initial conditions. Suppose that that $\kappa < 1$. Then, for all $\tau > 0$, and for all initial conditions $\xi(t) = \xi_0 \in (0, 1)$ on $(-\tau, 0)$, the solution of imitation equation oscillates around the ESS.*

2.2.4 Numerical illustrations

Impact of regret cost

The ESS depends on the parameter κ. When κ is large, the ESS point is closed to $(1, 0)$: the pure strategy *transmit*. Our first numerical experiment studies the convergence of these dynamics for the case of the unit growth parameter K and a small delay $\tau = 0.02$ as a function of the regret cost κ : we check the speed of convergence and the stability of the dynamics as a function of the cost κ. The state $2/3$ is a stationary point for these parameters, for which $2/3$ of the population choose to transmit. We took $\Delta = 1/3$. Figure 2.2 (resp. Figure 2.3) represents the trajectories of the population using the strategy T in the imitation dynamic (resp. replicator dynamic) when the initial condition is $\xi(t) = 0.02, \forall t \in (-\tau, 0)$.
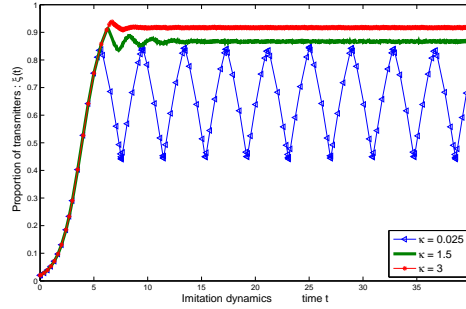
Figure 2.2: *Effect of κ on stability of delayed imitation dynamics, $\tau = 1$.*
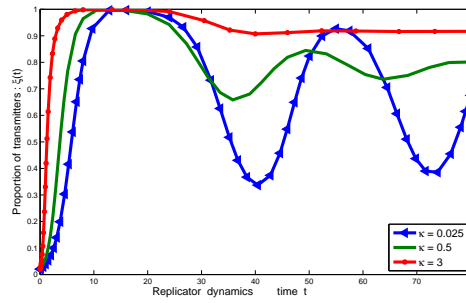
Figure 2.3: *Effect of κ on stability of replicator dynamics, $\tau = 8$.*

Impact of delay of transmission with regret cost

The ESS does not depend on the delay but the delay has a big influence on the stability of the system. Our second numerical experiment studies the convergence of these dynamics for the case of the unit growth parameter K as a function of delay τ. We took $\Delta = 1/3$ and $\kappa = 0.002$. The resulting trajectories of the population using the strategy T is represented as a function of time. We evaluate the stability varying the delay τ between 0.02 and 15 time units in the replicator dynamic in Figure 2.4 and Figure 2.5. For $\tau = 0.02$, we have stability but the convergence speed is slow. The other extreme is illustrated for $\tau = 15$ which the trajectory oscillates rapidly and the amplitude is seen to be greater than 2/3. The system is unstable. Figure 2.6 represents the fraction of the population using the strategy T in the imitation dynamic in which the system is unstable for all $\tau_T > 0$. The amplitude of oscillation growth with the delay.

Impact of the delay of transmission without regret cost

We evaluate the stability varying the delay τ_T between 0.02 and 10 time units in the replicator dynamic without regret cost in Fig. 2.7. For $\tau_T = 0.02$, we have stability but the convergence speed is slow. The other extreme is illustrated for $\tau_T = 10$ which the trajectory oscillates rapidly and the amplitude is seen to be greater than 2/3. The system is unstable. Figure 2.8 represents the trajectories of the population using the strategy T in the imitation dynamic in which the system is unstable for all $\tau_T > 0$. The amplitude of oscillation growth with the delay. The periodicity increases when the delay decreases. The best response at the time t when the population
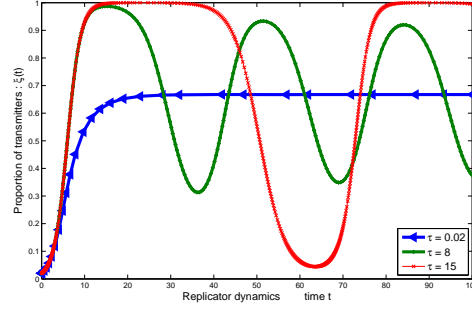
Figure 2.4: *Effect of τ_T in delayed replicator dynamic for $\kappa = 0.002$*
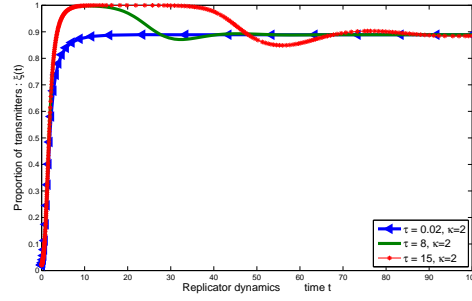
Figure 2.5: *Effect of τ_T in delayed replicator dynamic for $\kappa = 2$.*
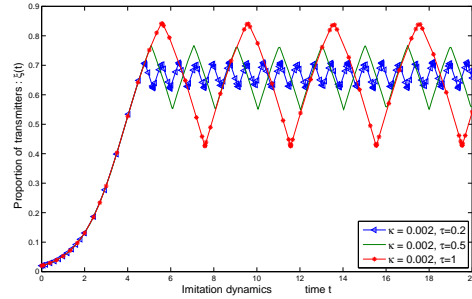
Figure 2.6: *Impact of τ_T in the imitation dynamics*

$\tilde{\xi}(t)$ is

$$BR(\xi(t)) = \mathbb{I}_{[0,1-\Delta]}(\xi(t - \tau_T)) + [0,1]\mathbb{I}_{\{1-\Delta\}}(\xi(t - \tau_T))$$

where $\mathbb{I}_A(.)$ is the indicator function of the set A. In Figure 2.9, the best response dynamics is represented for respectively $\tau_T = 0.04, 1, 2$. The mixed strategy $(1/2, 1/2)$ is chosen for the best response to $(1 - \Delta, \Delta)$. We evaluate the stability varying the delay τ_T between 0.04 and 8.4823 time units in the Brown-von Neumann-Nash in Fig. 2.10. The BNN dynamics becomes

$$\dot{\xi}(t) = \begin{cases} K(1 - \xi(t)^2)(-\xi(t - \tau_T) + 1 - \Delta) & \text{if } \xi(t - \tau_T) < 1 - \Delta \\ 0 & \text{if } \xi(t - \tau_T) = 1 - \Delta \\ K\xi(t)^2(-\xi(t - \tau_T) + 1 - \Delta) & \text{if } \xi(t - \tau_T) > 1 - \Delta \end{cases}$$

29

For $\tau_T = 0.04$, we have stability but the convergence speed is slow. The system is unstable when τ_T becomes greater than 8.4823.
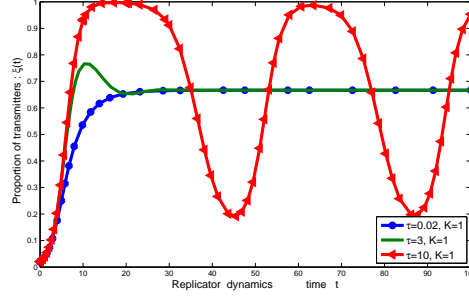
Figure 2.7: *Effect of τ_T on velocity and stability of replicator dynamics.*
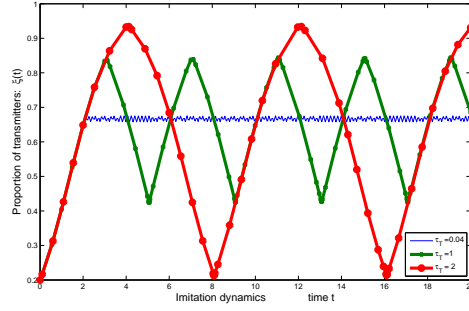
Figure 2.8: *Effect of τ_T on velocity and stability of imitation dynamics.*
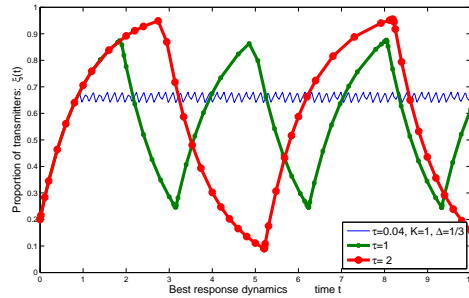
Figure 2.9: *Effect of τ_T on velocity and stability of best response dynamics.*

In these figures we observe cases of stable and of non-stable behavior. All turn out to confirm the stability conditions that we obtained above.

2.2.5 Notes

This chapter is based on our publications (13; 12; 215; 214). We considered evolutionary games with one population of users and studied delays impact on convergence to ESS for different
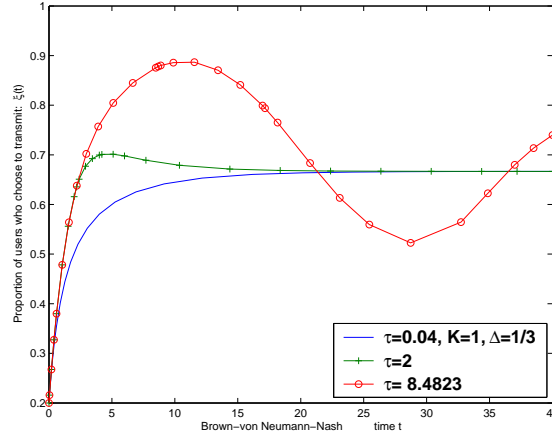
Figure 2.10: *Effect of τ_T on velocity and stability in BNN dynamics.*

types of evolutionary dynamics. Using some results from delay differential equation theory, we observed stability phenomena that are new with respect to non-delayed evolutionary game dynamics. In all the dynamics considered, delays were shown to have negative impact and ESS can be unstable when the delays are large. In the context of the access evolutionary game of mobile terminals, this suggest that updating MAC strategies in the terminals have to be done with care so as to avoid the oscillatory behavior that we observed in the non-stable regime.

2.3 Evolution of Transport Protocols in Wireless and Wired Networks

Todays Internet is well adapted to the evolution of protocols at various network layers. Much of the intelligence of congestion control is delegated to the end users and they have a large amount of freedom in the choice of the protocols they use. In the absence of a centralized policy for a global deployment of a unique protocol to perform a given task, the Internet experiences a competitive evolution between various variants of protocols. The evolution manifests itself through the upgrading of existing protocols, abandonment of some protocols and appearance of new ones. We highlight in this paper the modelling capabilities of the evolutionary game paradigm for explaining past evolution and predicting the future one. In particular, using this paradigm we derive conditions under which (i) a successful protocol would dominate and wipe away other protocols, or (ii) various competing protocols could coexist. In the latter case we also predict the share of users that would use each of the protocols. We further use evolutionary games to propose guidelines for upgrading protocols in order to achieve desirable stability behavior of the system.

When transferring data between nodes, flow control protocols are needed to regulate the transmission rates so as to adapt to the available resources. A connection that looses data units has to retransmit them later. In the absence of adaptation to the congestion, the on going transmissions along with the retransmissions can cause increased congestion in the network resulting in losses and further retransmissions by this and/or by other connections. This type of phenomenon, that leads to several 'congestion collapses' (113), motivated the evolution of the Internet transport protocol, TCP, to a protocol that reduces dramatically its throughput upon congestion detection.

31

The possibilities to deploy freely new versions of protocols on terminals connected to the Internet creates a competition environment between protocols. Much work has been devoted to analyze such competition and to predict its consequences. The two main approaches for predicting whether one version of a protocol would dominate another one are

- **Inter Population Competition (IRPC):** One examines local interactions between connections of different types that interact with each other (by sharing some common bottleneck link). If a connection that corresponds to one version performs better in such an interaction, then the DC approach predicts that it would dominante and that the other version would vanish.

- **Intra Population Competition (IAPC):** In this approach one studies the performance of a version of a protocol assuming a world where all connections use that version. This is repeated with the other version. One then predicts that the version that gives a better world would dominate.

We address the dominance question with the evolutionary game paradigm and provide an alternative answer along with a more detailed analysis of this competition scenario. Our approach predicts whether one can expect one protocol to dominate the other or whether the two protocols can be expected to coexist. It provides with the tools for computing the share of the population that is expected to use each version in case the versions would coexist. Finally, it provides a description of the dynamics of the competition, which may result in a stable behavior that consists of a convergence to some equilibrium, or it may display instabilities and oscillations. By identifying the conditions for a stable behavior, one can provide guidelines for upgrading protocols so as to avoid undesirable oscillating behavior.

We provide a framework to describe and predict evolution of protocols in a context of competition between two types of behaviors: aggressive and peaceful. We compute the ESS for congestion protocols of different degree of aggressiveness. We identify cases in which at ESS, only one population prevails (ESS in pure strategies) and others, in which an equilibrium between several population types is obtained. To study this, we map the problems, whenever possible, into the Hawk and Dove Game. We then study the convergence of the replicator dynamics to it.

We develop a framework for controlling evolutionary dynamics (changing or upgrading protocols) through the choice of a gain parameter governing the replicator dynamics. We address the following two design issues concerning this choice:

(i) the tradeoff between fast convergence and stability. We identify a simple threshold on the gain parameter in the replicator dynamics such that the stability is only determined by whether we exceed or not the threshold.

(ii) the stability as a function of delays. We derive new stability conditions for the replicator dynamics in the Hawk and Dove game with non-symmetric delays and apply it to the evolution of the MAC and transport layer protocols.

We first provide in the next subsection the needed background on evolutionary games. We summarize work on competition between TCP versions in subsection 2.3.1. We then study the ESS for congestion control protocols (subsection 2.3.2). After that, we investigate the impact of the choice of some parameters in the replicator dynamics on the stability of the system in subsection 2.3.4. Finally we give some numerical investigations.

2.3.1 Competition Between Congestion Control Protocols

There are various versions of the TCP protocol among which the mostly used one is New-Reno. The degree of 'aggressiveness' varies from version to version. The behavior of New-Reno is approximately AIMD (Additive Increase Multiplicative Decrease): it adapts to the available capacity by increasing the window size in a linear way by α packets every round trip time and when it detects congestion it decreases the window size to β times its value. The constants α and β are 1 and $1/2$, respectively, in New Reno.

In last years, more aggressive TCP versions have appeared, such as HSTCP (High Speed TCP) (191) and Scalable TCP (122; 72; 5). HSTCP can be modeled by an AIMD behavior where α and β are not constant anymore : α and β have minimum values of 1 and of $1/2$, resp. and both increase in the window size. Scalable TCP is an MIMD (Multiplicative Increase Multiplicative Decrease) protocol, where the window size increases exponentially instead of linearly and is thus more aggressive. Versions of TCP which are less aggressive than the New-Reno also exist, such as Vegas (54).

Several researchers have analyzed the performance of networks in which various transport protocols coexist, see (200; 51; 1; 6; 135). In all these papers, the population size using each type of protocol is fixed.

Some papers have already considered competition between aggressive and well behaved congestion control mechanisms within a game theoretic approach. Their conclusions in a wireline context was that if connections can choose selfishly between a well behaved cooperative behavior and an aggressive one then the Nash equilibrium is obtained by all users being aggressive and thus in a congestion collapse (89; 136).

We introduce in the next two sections two models of competition between TCP versions, both can be modeled within the framework of the Hawk and Dove game. This will allow us to predict whether a given version of TCP is expected to dominate others (ESS in pure strategies, which means that some versions of TCP would disappear) or whether several versions would co-exist. The first model is adapted to competition in wireless networks and the second to wireline networks.

2.3.2 Competition in Wireless Networks

During the last few years, many researchers have been studying TCP performances in terms of energy consumption and average goodput within wireless networks (180; 243). Via simulation, the authors show that the TCP New-Reno can be considered as well performing within wireless environment among all other TCP variants and allows for greater energy savings. Indeed, a less aggressive TCP, as TCP New-Reno, may generate lower packet loss than other aggressive TCP. Thus the advantage of an aggressive TCP in terms of throughput could be compensated with energy efficiency of a more gentle TCP version. (In Section 2.3.3 we shall illusutrate another consideration that affects the competition between TCP versions.) The goal of this section is to illustrate this point, as well as its possible impact on the evolution of the share of TCP versions, through a simple model of an aggressive TCP.

The model. We consider two populations of connections, all of which use AIMD TCP. A connection of population i is characterized with a linear increase rate α_i and a multiplicative decrease factor β_i. Let $x_i(t)$ be the transmission rate of connection i at time t. We consider the following simple model for competition.

- The RTT (round trip times) are the same for all connections.

- There is light traffic in the system in the sense that a connection either has all the resources its needs or it shares the resources with one other connection. (If files are large then this is a light regime in terms of number of connections but not in terms of workload).

- Losses occur whenever the sum of rates reaches the capacity C: $x_1(t) + x_2(t) = C$.

- Losses are synchronized: when the combined rates attain C, both connections suffer from a loss. This synchronization has been observed in simulations for connections with RTTs close to each other (2). The rate of connection i is reduced by the factor $\beta_i < 1$.

- As long as there are no losses, the rate of connection i increases linearly by a factor α_i.

We say that a TCP connection i is more aggressive than a connection j if $\alpha_i \geq \alpha_j$ and $\beta_i \geq \beta_j$. Let $\overline{\beta}_i := 1 - \beta_i$. Let x_n and y_n be the transmission rates of connection i and j, respectively, just before a loss occurs. We have $x_n + y_n = C$. Just after the loss, the rates are $\beta_1 x_n$ and $\beta_2 y_n$. The time it takes to reach again C is

$$T_n = \frac{C - \beta_1 x_n - \beta_2 y_n}{\alpha_1 + \alpha_2}$$

which yields the difference equation:

$$x_{n+1} \quad = \quad \beta_1 x_n + \alpha_1 T_n = q x_n + \frac{\alpha_1 C \overline{\beta}_2}{\alpha_1 + \alpha_2}$$

where $q = \dfrac{\alpha_1 \beta_2 + \alpha_2 \beta_1}{\alpha_1 + \alpha_2}$. The solution is given by

$$x_n = q^n x_0 + \left(\frac{\alpha_1 C \overline{\beta}_2}{\alpha_1 + \alpha_2} \right) \frac{1 - q^n}{1 - q}.$$

HD game: throughput-loss tradeoff

In wireline, the utility related to file transfers is usually taken to be the throughput, or a function of the throughput (e.g. the delay). It does not explicitly depend on the loss rate. This is not the case in wireless context. Indeed, since TCP retransmits lost packets, losses present energy inefficiency. Since energy is a costly resource in wireless, the loss rate is included explicitly in the utility of a user through the term representing energy cost. We thus consider fitness of the form $J_i = Thp_i - \lambda R$ for connection i; it is the difference between the throughput Thp_i and the loss rate R weighted by the so called tradeoff parameter, λ, that allows us to model the tradeoff between the valuation of losses and throughput in the fitness. We now proceed to show that our competition model between aggressive and non-aggressive TCP connections can be formulated as a HD game. We study how the fraction of aggressive TCP in the population at (the mixed) ESS depends on the tradeoff parameter λ.

Since $|q| < 1$, we get the following limit x of x_n when $n \to \infty$:

$$x = \frac{\alpha_1 C \overline{\beta}_2}{\alpha_1 + \alpha_2} \cdot \frac{1}{1 - q} = \frac{\alpha_1 \overline{\beta}_2 C}{\alpha_1 \overline{\beta}_2 + \alpha_2 \overline{\beta}_1}.$$

It is easily seen that the share of the bandwidth (just before losses) of a user is increasing in its aggressiveness. Hence the average throughput of connection 1 is

$$Thp_1 = \frac{1 + \beta_1}{2} \times \frac{\alpha_1 \overline{\beta}_2}{\alpha_1 \overline{\beta}_2 + \alpha_2 \overline{\beta}_1} \times C.$$

The average loss rate of connection 1 is the same as that of connection 2 and is given by

$$R = \frac{1}{T} = \left(\frac{\alpha_1}{\overline{\beta}_1} + \frac{\alpha_2}{\overline{\beta}_2} \right) \frac{1}{C} \text{ where } T = \frac{\overline{\beta}_1 \overline{\beta}_2 C}{\alpha_1 \overline{\beta}_2 + \alpha_2 \overline{\beta}_1}$$

with T being the limit as $n \to \infty$ of T_n.

Let H corresponds to (α_H, β_H) and D to (α_D, β_D) such that $\alpha_H \geq \alpha_D$ and $\beta_H \geq \beta_D$. Then, for $i = 1, 2$, $Thp_i(H, H) = Thp_i(D, D)$. Since the loss rate for any user is increasing in $\alpha_1, \alpha_2, \beta_1, \beta_2$ it then follows that $J(H, H) < J(D, D)$, and $J(D, H) < J(D, D)$. We conclude that the utility that describes a tradeoff between average throughput and the loss rate leads to the HD structure.

The mixed ESS is given by the following probability of using H:

$$x^*(\lambda) = \frac{\eta_1 - \eta_2 \lambda}{\eta_3} \quad \text{where}$$

$$\eta_1 = \left(\overline{\mu} \frac{1 + \beta_1}{2} - \frac{1 + \beta_2}{4} \right) C, \ \eta_2 = \frac{1}{C} \left(\frac{\alpha_1}{\overline{\beta}_1} - \frac{\alpha_2}{\overline{\beta}_2} \right),$$

$$\eta_3 = C(\frac{1}{2} - \mu) \frac{\beta_1 - \beta_2}{2}, \ \mu = \frac{\alpha_2(\overline{\beta}_1)}{\alpha_2(\overline{\beta}_1) + \alpha_1(\overline{\beta}_2)}.$$

where $\overline{\mu} := 1 - \mu$. Note that η_2 and η_3 are positive. Hence, the equilibrium point x^* decrease linearly on λ. We conclude that applications that are more sensitive to losses would be less aggressive at ESS (Braess type paradoxes do not occur here).

2.3.3 Competition in Wireline Networks

Tradeoff between throughput and fairness Consider a number of TCP connections that share a common bottleneck. We study the competition between the New-Reno version of TCP and the Scalable version of TCP that has been proposed (122) in the context wireline networks that are characterized by a very high speed and has long distance (and thus large delays). Scalable TCP is then much more aggressive than the New-Reno TCP as it is an MIMD (Multiplicative Increase Multiplicative Decrease) protocol, where the window size increases exponentially instead of linearly at the absence of losses. It is also more aggressive when losses are detected: it decreases its window size (and thus the transmission) rate to 0.875 times the one it had prior to the loss (instead of to halving it as is done in New-Reno). We begin by examining the Inter and Intra Population Competition.

Intra Population Competition If the connections are all symmetric and use the same AIMD version, then we know that they share equally the bandwidth (7). Moreover if the number of connections or the buffer size is large then the connections use all the available bandwidth so that the throughput per connection is the capacity of the bottleneck link divided by the number of connections that share the bottleneck.

The situation is different when a bottleneck link is shared between symmetric MIMD connections. The connections suffer from a high level of unfairness under various scenarios. We briefly summarize some central findings in (6):

- The case of synchronized losses: if all connections suffer a loss at the same time then the bandwidth share of each connection remains constant in time. This is particularly harmful for new connections that may not be able to grab throughput.

- The case of asynchronous losses: in case that there are some asynchronous random losses (with or without synchronous losses that are due to congestion), the bandwidth is shared more fairly: the average throughput tend to be the same after a very long time.

Remarks 2.3.3.1. *The situation that we described corresponds to symmetric connections. In case the round trip times are not the same, the throughput share between the connections is inversely proportional to the RTT. In MIMD, in contrast, the connection with larger RTT gets almost no bandwidth.*

As for the throughput, MIMD is able indeed to use all the available link capacity (provided that there are buffers; the buffer size needed for full utilization are much smaller than those needed in AIMD. Below we do not consider this aspect.)

Inter Population Competition The way bandwidth is shared between TCP connections of different types (AIMD and MIMD) that share a common bottleneck is complex and depends on various factors. We shall restrict our discussion to links with high bandwidth and large buffering. In those conditions, an MIMD connection gets more throughput than AIMD does. As for fairness, the problems of no fairness (in the sense that the the initial share remains the same) or fairness on a very slow time scale, do not occur between two connections of different types (AIMD and MIMD).

The game formulation Consider a competition between a large population of users where a fraction of them use an AIMD fersion of TCP and another MIMD connection sharing a common bottleneck link We can now use the above description to define. the structure of the evolutionary game. There are many pairwise interactions described through the matrix game of the form given in Figure 2.11 where H (Hawk) stands for the MIMD TCP and D (Dove) stands for AIMD. Let the fitness of a player that uses an action i that interacts with a player that plays j be given as $F(i,j) = \theta(i,j) - f(i,j)$ where $\theta(i,j)$ is the throughput part corresponding to the share of the bottleneck capacity that the player receives and f is a disutility for lack of fairness. $f(i,j)$ is zero except for the case $i = j = H$; $\theta(i,j) = 0.5$ when $i = j$, where as $\theta(H,D) = 1 - \theta(D,H) = a$ where a is some positive constant smaller than 0.5. Denoting $d = 0.5 - f(H,H)$ we obtain the fitnesses as given in Figure 2.11.

Player II

		H	D
Pl. I	H	d	1–a
	D	a	0.5

Figure 2.11: MIMD (H) versus AIMD (H)

Characterizing the ESS We note that the conditions defining a Generalized Hawk and Dove game are satisfied. We thus conclude the following:

- AIMD is never a dominant strategy (i.e. there is no ESS where all the population uses AIMD). To understand this, we note that if all used AIMD then a shift of a small fraction of the population to MIMD would give this fraction a larger fitness, since MIMD would get a larger throughput when interacting with AIMD but it would almost never suffer

from the fairness problem since encounters between MIMD connections would be quite rare.

- MIMD may, on the other hand, be dominant. This occurs if $d > a$. In other words, if users did not care much about the fairness problems of MIMD, then its share in the population could be expected to grow to 1 and AIMD could disappear.

- if $d < a$ then the unique ESS consists of a coexistence between AIMD and MIMD. At ESS, the fraction using MIMD is given by

$$p = \frac{1 - 2a}{1 - 2d}$$

Notes

We have mentioned in the introduction two simplistic approaches for predicting which population would dominate. The Intra Population Competition (IAPC) compares a world with only one type of behavior and compares the corresponding utility; in our case it would amount to observing that $F(D, D) > F(H, H)$ and thus would wrongly suggest that AIMD would dominate the MIMD and their share would be one, where as we know that this possibility never occurs

The Inter Population Competition (IRPC) compares only the interactions that involve H and D and it would predict that MIMD would dominate since it gets better fitness in these interactions. This prediction is again wrong; indeed, we have seen that both population may coexist insight of having $F(H, D) > F(D, H)$ provided that $d < a$.

Delayed replicator dynamics

We introduce here the replicator dynamics which describes the evolution in the population of the various strategies. In the replicator dynamics, the share of a strategy in the population grows at a rate equal to the difference between the payoff of that strategy and the average payoff of the population. More precisely, consider N strategies. Let $\mathbf{x}$ be the N dimensional vector whose ith element x_i is the population share of strategy i. Thus we have $\sum_i x_i = 1$ and $x_i \geq 0$. Below we denote by $J(i, k)$ the expected payoff (or the fitness) for a player using strategy i when it encounters a player with strategy k. With some abuse of notation we define $J(i, \mathbf{x}) = \sum_j J(i, j)x_j$.

Then the replicator dynamics is defined as

$$\dot{x}_i(t) = x_i K\left(J(i, \mathbf{x}) - \sum_j x_j J(j, \mathbf{x})\right) = x_i K\left(\sum_j x_j J(i, j) - \sum_j \sum_k x_j J(j, k)x_k\right)$$

where K is a positive constant and $\dot{x}_i(t) := \mathrm{d}x_i(t)/\mathrm{d}t$. Note that the right hand side vanishes when summing over i. This is compatible with the fact that we study here the share of each strategy rather than the size of the population that uses each one of the strategies.

In Equation (2.35), the fitness of strategy i at time t has an instantaneous impact on the rate of growth of the population size that uses it. An alternative more realistic model for replicator dynamic would have some delay: the fitness acquired at time t will impact the rate of growth τ time later. We then have

$$\dot{x}_i(t) = x_i(t)K\left(\sum_j x_j(t - \tau)J(i, j) - \sum_{j,k} x_j(t)J(j, k)x_k(t - \tau)\right) \tag{2.35}$$

where K is some positive constant. The delay τ represents a time scale much slower than the physical (propagation and queueing) delays, it is related to the time scale of (i) switching from the use of one protocol to another (ii) upgrading protocols.

2.3.4   Architecting evolution

We study the choice of two parameters in the replicator dynamics that impact the stability of the evolution process of protocols: the gain parameter K and the delay τ appearing in Equation (2.35). The standard replicator dynamics (2.35) appearing in the evolutionary game literature is defined with $K = 1$. K's other than one can be interpreted as if the utilities J are multiplied by a constant. Alternatively, it can be seen as scaling time. The parameter K can thus be used to accelerate the rate of convergence in (2.35).

The impact of K and τ on the stability

We consider below the case of two players and two actions. Define

$$\delta_1 = J(B, A) - J(A, A), \quad \delta_2 = J(A, B) - J(B, B),$$

$$\delta = \delta_1 + \delta_2, \quad \theta = \frac{\delta\pi}{2\delta_1\delta_2}$$

Guidelines for an evolution framework.

For $K = 1$, it has been shown in (205) that if the delay τ in (2.35) satisfies $\tau < \theta$ then the mixed ESS is asymptotically stable, and if τ increases beyond θ then the ESS becomes unstable.

We make the observation that Equation (2.35) with K is equivalent to Equation (2.35) with all elements $J(i, m)$ multiply by K. Thus we can use the result of (205) to conclude that the stability condition for general K is given simply by

$$\tau K < \theta. \tag{2.36}$$

This provides us with an important guideline for designing evolutionary protocols. In order for such a protocol to be scalable to any delay, the product of the adaptation speed parameter K and delay τ should be $O(1)$. Thus the larger the delay is, the slower we should react to the fitness of a strategy being used.

We note that this type of scaling is quite familiar in other networking contexts: the internet transport protocol TCP has a throughput that scales according to $1/RTT$ (where RTT is the round trip delay). This scaling is obtained by a self clocking mechanism based on ACKs that trigger new transmissions.

2.3.5   Numerical illustrations

Impact of gain parameter

Our first numerical experiment on evolution of protocols studies the behavior of the replicator dynamics for the case of one delay unit as a function of K: we check the speed of convergence

and the stability of the replicator dynamics as a function of the gain parameter K. We consider the following fixed parameters: we took $\tau = 1$, $\delta = 2/3$, and let K vary between 0.16 and 15. A unique mixed ESS exists for these parameters, for which the fraction of the population using H is $3/4$. The resulting trajectories of the population ratio using the first strategy, H, as a function of time, is given in Fig. 2.14, 2.12, 2.13 . For $K = 0.16$, we have stability but the convergence speed is slow. The other extreme is illustrated for $K = 15$ which is seen to be unstable: it oscillates rapidly and the amplitude is seen to grow slowly.
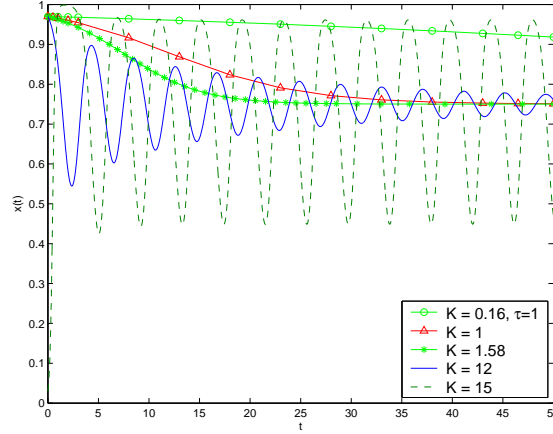
Figure 2.12: *Effect of K on stability*
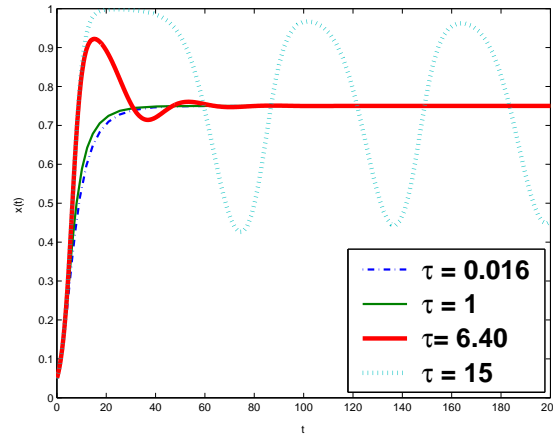
Figure 2.13: *Effect of τ on stability*

Impact of delay

We now keep K constant and evaluate the stability varying the delay between 0.016 and 15 time units. When $\tau = 0.016$ the system is stable but the rate of convergence to the interior equilibrium is not fast. For $\tau = 15$ the system is unstable, the solution oscillates around the equilibrium $x^* = 3/4$.
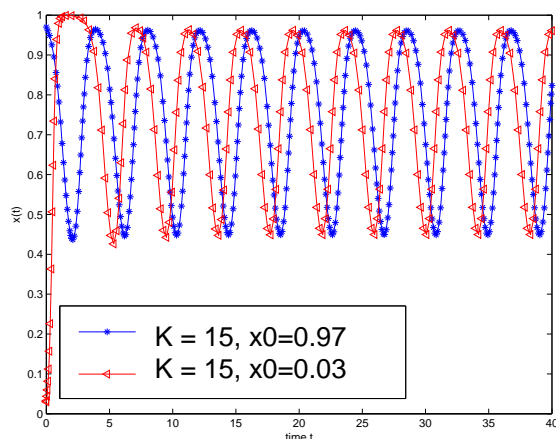
Figure 2.14: *Effect of initial state*

Oscillating Solution and dependence on the initial state.

In Figure 2.14 we display an oscillatory behavior of the population ratios as function of time for two different initial values of $x(0) = 0.03$ and $x(0) = 0.97$ with $K = 15$ and $\tau = 1$. It corresponds to an unstable regime in which the ESS is not attained. The trajectories are seen to converge to periodic ones. The limit trajectories look the same and do not depend on the initial state except for a dependence through the phase. In this unstable regime, more than one protocol co-exist and the ratio of population sizes using the protocols has oscillations with large amplitude.

In both figures 2.14 and 2.13, we observe that we have stability when $\tau K < 4\pi \simeq 12.56$. Indeed, in figure refallfigtau, the parameter $\tau = 1$, hence the condition of stability (2.36) becomes $K < 12.56$. This actually confirm that using $K = 0.16, 1, 1.56, 12$, the system is stable and using $K = 15$, the system is unstable. We observe the same behavior when keeping K constant and varying the delay τ.

2.3.6 Notes

In this section, we have studied evolutionary aspects of congestion control protocols in using the biological paradigm of evolutionary games. We have studied the questions of whether one could expect one type of protocol to wipe away another one or whether we may expect protocols to coexist. In the latter case we provided a quantitative characterization of the share that each protocol could be expected to have in the whole population at equilibrium. We then identified conditions under which there is a convergence to the equilibrium and obtained examined the oscillating behavior that occurs when there is no convergence. The conditions that guarantee convergence can be used as guidelines for deployment of new protocols so that the users upgrades would result in a stable system wide behavior.

2.4 Multi-class delayed evolutionary game dynamics

In this section we briefly describe delayed evolutionary game dynamics with multi-classes of players. Consider several classes of large population of players and finitely many pure strategies in each class. Let $\mathcal{E} = \{1, 2, \ldots, E\}$ be the set of classes. the action set of class e is $\mathcal{A}^e = \{1, 2, \ldots, n_e\}$. Denote $x(t) = (x^e(t))_{e \in \mathcal{E}}$ the population profile at time t. The composition of class e is then describe by $x^e(t) = [x_1^e(t), \ldots, x_{n_e}^e(t)]$ where $x_a^e(t)$ is fraction of players at time t with the strategy a in class e. Denote by $F_a^e(x(t))$ the expected payoff of a user with a in class e when the population profile is $x(t)$ at time t. Denote τ_a^e the time associate to the strategy a of class e. The delayed evolutionary game dynamics is given by

$$\dot{x}_a^e(t) = \sum_{b \in \mathcal{A}^e} x_b^e(t) \beta_a^{e,b}(x(t), \{x(t - \tau_b^e)\}_{e,b}) \tag{2.37}$$

$$-x_a^e(t) \sum_{b \in \mathcal{A}^e} x_b^e(t) \beta_b^{e,a}(x(t), \{x(t - \tau_b^e)\}_{e,b}) \tag{2.38}$$

Delayed Replicator Dynamics with multi-class A fraction of member of a subpopulation of class e grows when its payoff is greater than the expected average payoff of all the class e in the population.

$$\frac{d}{dt} x_a^e(t) = \mu^e x_a^e(t) \left[F_a^e(x(t - \tau_a^e)) - \sum_{b \in \mathcal{A}^e} x_b^e(t) F_b^e(x(t - \tau_b^e)) \right] \tag{2.39}$$

Every equilibrium of evolutionary game with delayed payoffs is a rest point of the dynamic (2.39).

Delayed Brown-von Neuman-Nash Dynamics Denote g_a^e by the positive part of the excess payoff of subpopulation a of class e In continuous time, Brown-von Neuman-Nash Dynamics is given by

$$\frac{d}{dt} x_a^e(t) = \mu^e \left[g_a^e(x(t)) - x_a^e(t) \sum_{b \in \mathcal{A}^e} g_b^e(x(t)) \right] \tag{2.40}$$

where $g_a^e(x(t)) = \max \left[0, F_a^e(x(t - \tau_a^e)) - \sum_{b \in \mathcal{A}^e} x_b^e(t) F_b^e(x(t - \tau_a^e)) \right]$. The equilibria of evolutionary game with delayed payoffs are exactly the rest points of the dynamic (2.40).

Delayed Best Response Dynamics Members of each subpopulation revise their strategy and choose the best replies $BR(x(t))$ at the current population state $x(t)$.

$$\frac{d}{dt} x^e(t) \in \mu^e(BR^e(x(t)) - x^e(t)) \tag{2.41}$$

where $BR^e(x^{-e}(t)) = \arg \max_{y \in \Delta(\mathcal{A}^e)} \left\{ \sum_{b \in \mathcal{A}^e} y_b F_b^e(x(t - \tau_b^e)) \right\}$,

The equilibria of evolutionary game with delayed payoffs are exactly the rest points of the dynamic (2.41).

Delayed Logit dynamics $\dfrac{d}{dt} x_a^e(t) = \mu^e \left[\dfrac{e^{\frac{F_a^e(x(t - \tau_a^e))}{\eta^e}}}{\sum_{b \in \mathcal{A}^e} e^{\frac{F_b^e(x(t - \tau_b^e))}{\eta^e}}} - x_a^e(t) \right]$

41

Imitation by dissatisfaction Users in the population review their strategy and imitate the better strategy at that time. $\frac{d}{dt}x_a^e(t) = \mu^e x_a^e(t) \left[\sum_{b \in \mathcal{A}^e} x_b^e(t)(\rho_{ba}^e(x(t)) - \rho_{ab}^e(x(t))) \right]$ where $\rho_{ba}^e(x(t)) = g(F_a^e(x(t - \tau_a^e)), F_b^e(x(t - \tau_b^e)), g(a,b) = 1$ if $a > b$, otherwise $g(a,b) = 0$.

Delayed orthogonal projection dynamics The orthogonal projection dynamic is a myopic adaptive dynamic in which a subpopulation grows when its expected payoff is greater than the arithmetic average payoff of all the population. $\frac{d}{dt}x_a^e(t) = \mu^e \left[F_a^e(x(t - \tau_a^e)) - \frac{1}{n_e} \sum_{b \in \mathcal{A}^e} F_b^e(x(t - \tau_b^e)) \right]$

Delayed ray-projection dynamics $\frac{d}{dt}x_a^e(t) = \mu^e \left[F_a^e(x(t - \tau_a^e)) - x_a^e(t) \sum_{b \in \mathcal{A}^e} F_b^e(x(t - \tau_a^e)) \right]$

Delayed $\theta-$Smith dynamics

$$
\begin{aligned}
\frac{d}{dt}x_a^e(t) &= \mu^e \sum_{b \in \mathcal{A}^e} x_b^e(t) \max(0, F_a^e(x(t - \tau_a^e)) - F_b^e(x(t - \tau_b^e)))^\theta \\
&\quad - \mu^e x_a^e(t) \sum_{b \in \mathcal{A}^e} \max(0, F_b^e(x(t - \tau_b^e)) - F_a^e(x(t - \tau_a^e)))^\theta
\end{aligned}
\tag{2.42}
$$

The equilibria of the evolutionary game with delayed payoffs are exactly the rest points of the dynamic (2.42).

Delayed target projection dynamics $\dot{x}^e(t) = \text{proj}_{TX^e}(F^e(\{x(t - \tau_k)\}_k)) - x^e(t)$

In figure 2.15 we illustrate the evolution of cooperation with three types in which one can move from one to another stable equilibrium because of time delays. We conclude that the delayed evolutionary game dynamics can have unpredictable trajectories even when the non-delayed dynamics leads to stable points.
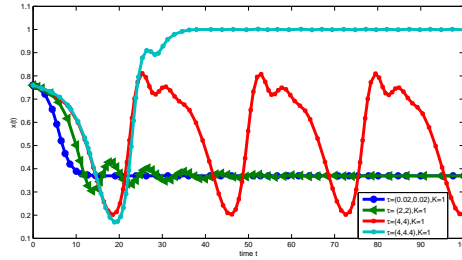
Figure 2.15: *Moving from equilibrium to another.*

Notes

In next chapter we pursue the study of delayed game dynamics and extend this analysis for evolutionary games with variable number of players in each local interaction. The chapter is based to our publications in (213; 11; 211; 218; 210).

Chapter 3

Evolutionary games with random number of interacting players

The classical evolutionary game formalism is a central mathematical tool developed by biologists for predicting population dynamics in the context of interaction between populations. In order to make use of the wealth of tools and theory developed in the biology literature, many works in the area of computer networks (52; 20; 214) ignore cases where local interactions between populations involve more than two individuals. This restriction limits the modeling power of evolutionary games which are not useful in a network operating at heavy load, such as ad-hoc networks with high density. This motivated us in this chapter to consider more than two users interacting locally and the interactions can be *non-reciprocal*.

Our main contributions in this chapter can be summarized in three points:

- The first objective of this work is to extend the evolutionary game framework to allow an arbitrary (possibly random) number of players that are involved in a local interaction (possibly non-reciprocal interaction);

- The second objective of this work is to apply the extended model to access games studied in (216) which we extend to more than two interacting nodes. In the context of MAC games, we study the impact of the node distribution in the game area on the equilibrium stable strategies of the evolutionary game. The interaction between more than two individuals in a population is a new concept in evolutionary game theory and has a lot of application in multiple access game in wireless networks. Considering this kind of games, we use the notion of expected utility as this game is not symmetric, indeed the number of players with which a given one interacts may vary from one to another; and also non-reciprocity property. We consider the following parameters in the access game: transmission cost, collision cost and regret cost. We analyze the impact of these parameters on the probability of successful transmission and give some optimization issues. The notion of *correlated evolutionarily stable state* (see Section 3.4) and coordination mechanism are used to improve the performance(reducing collisions).

- The third objective of this work is to apply evolutionary game models to study the interaction of numerous mobiles in competition in a wireless environment. The power control game in wireless networks is a typical non-cooperative game where each mobile decides about his transmit power in order to optimize its performance. This application of non-cooperative game theory and tools is studied in several articles (145; 171; 146). The main

difference in this article is the use of the evolutionary game theory which deals with population dynamics that is well adapted for studying the power control game in dense wireless networks. Specifically, we focus our first study in a power control game in a dense wireless network where each user transmits using orthogonal codes like in W-CDMA. In the second scenario, we consider also uplink transmissions but inter-cell interferences like in WiMAX cells deployment. The utility function of each mobile is based on carrier (signal)-to-interference ratio and pricing scheme proportional to transmitted power. We provide and predict the evolution of population between two types of behaviors : aggressive (high power) and peaceful (low power). We identify cases in which at ESS, only one population prevails (ESS in pure strategies) and others, in which an equilibrium between several population types is obtained. We also provide the conditions of the uniqueness of ESS. Furthermore, we study different pricing for controlling the evolution of population. We introduced the *benefit of correlation* (BoC) and show that the random access game with two strategies has a BoC of 100%. In addition, we show that *hierarchical solutions* and *canonical correlation devices* are do not improve the performance the number of strategies is at least three.

The rest of this chapter is organized as follows. In Section 3.1 and 3.2, we illustrate some examples of non-reciprocal interactions and limitations of the pairwise interaction approach. In Section 3.3 we describe a general model of population games with random number of interacting players at each local interaction and study both evolutionary stability and evolutionary game dynamics. We then study in Section 3.4 a generalized multiple access game with a random number of players. In Section 3.5, we present the evolutionary game model to study the power allocation game in a dense wireless ad hoc network. In Section 3.6, we propose an evolutionary game analysis for the power allocation game in a context of inter-cell interferences in OFDMA-based WiMax cells. Numerical results of both wireless network architectures are proposed in each section.

3.1 Reciprocal and non-reciprocal interactions

Consider a population of users. Each individual needs occasionally to take some action. When doing so, it interacts with the actions of some M (possibly random number of) other individuals. All players have the same actions available, and same *expected* utility. We note however that the actual realizations need not be symmetric. In particular, (i) the number of players with which a given player interacts may vary from one player to another. (ii) We do not even need the reciprocity property: if player A interacts with player B, we do not require the converse to hold. We provide some examples of multiple access games to illustrate this non-reciprocity.

We now model interference control as local interactions between transmitters; for each transmitter there corresponds a receiver. We shall say that a transmitter A is subject to an interaction (interference) from transmitter B if the transmission from B overlaps that from A, and provided that *the receiver of the transmission from A is within interference range of transmitter B*.

3.1.1 Non-reciprocal pairwise interaction

Consider the example depicted at Figure 3.4. It contains 4 sources (circles) and 3 destinations (squares). A transmission of a source i within a distance r of the receiver R, causes interference to a transmission from a source $j \neq i$ to receiver R. We see that Source A and Source C cause no interference to any other transmission but the transmission from A suffers from interference from source B, and the one from C suffers from the transmission of the top most source (called

D). Source B and D interfere with each other at their common destination. Thus each of the four sources suffers interference from a single other source, but except for nodes B and D, the interference is not reciprocal.
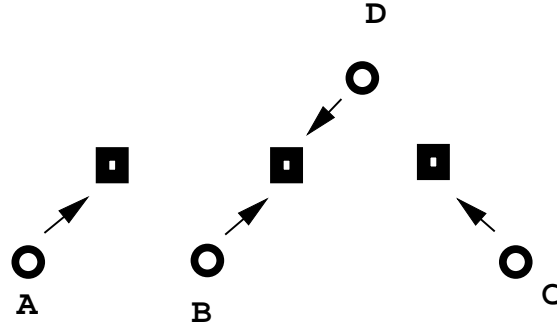
Figure 3.1: Non-reciprocal pairwise interactions

It is easy to see that in this non-reciprocal pairwise interaction, the following configurations leads to equilibria:

- *A* and *D* transmit at the same time slot. There are 2 successful transmissions if the others stay quiet.

- *A* and *C* transmit at the same time slot. There are 2 successful transmissions if the others stay quiet.

- *B* and *C* transmit at the same time slot. There are 2 successful transmissions if the others stay quiet

- *A*, *D* and *C* transmit at the same time slot. There is 1 successful transmission (only from *A*) if *B* stay quiet.

- *A*, *B* and *C* transmit at the same time slot. There is 1 successful transmission (only from *C*) if *D* stay quiet.

- *C*, *D* transmit at the same time slot. There is 1 successful transmission (only from *A*) if *B* stay quiet.

- *B*, *D* transmit at the same time slot. There is no successful transmission in the system at that time.

3.1.2 Non-reciprocal interactions between groups three players

In Figure 3.2 there are four sources and only two destinations. Node A does not cause any interference to the other nodes but suffers interference from nodes B and D. Nodes B, C, D interfere with each other. This is a situation in which each mobile is involved in interference from two other mobiles but again the interference is not reciprocal.

The following configurations leads to equilibria:

- *A* and *C* transmit at the same time slot. There are 2 successful transmissions if the others stay quiet. This is a global optimum.

- Only *B* transmits. There is one successful transmissions if the others stay quiet.
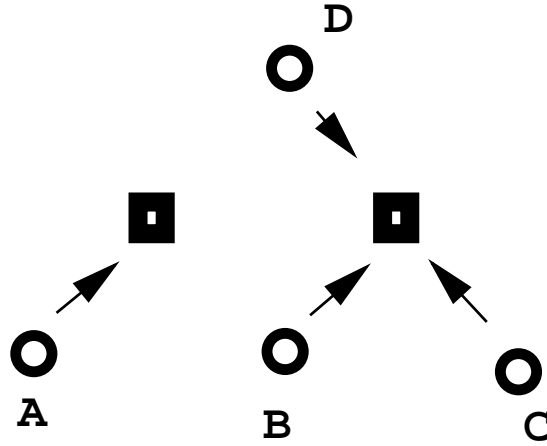
Figure 3.2: *Non-reciprocal interaction between groups three players*

- Only *D* transmits. There is one successful transmissions if the others stay quiet.

- *B* and *D* transmit at the same time slot. There is no successful transmissions.

3.1.3 Interactions between random number of players

In this subsection the number of interfering nodes is not fixed. The mobile A suffers interference from 2 nodes, B and D suffer interference from a single other node and C does not suffer (and does not cause) interference.
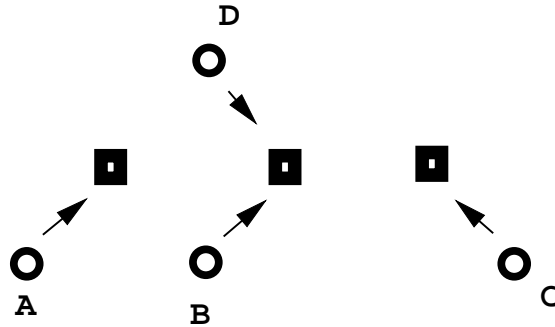
Figure 3.3: *Interactions between a random number of players*

The following configurations leads to equilibria:

- *A* and *C* transmit at the same time slot. There are 2 successful transmissions if the others stay quiet. This is a global optimum.

- *B* and *C* transmit at the same time slot. There are 2 successful transmissions if the mobiles *A* and *D* stay quiet. This is a global optimum.

- *B* and *C* transmit at the same time slot. There are 2 successful transmissions if the others stay quiet. This is a global optimum.

- *D* and *C* transmit at the same time slot. There are 2 successful transmissions if the others mobiles *A* and *B* stay quiet. This is a global optimum.

Note that if *D* or *B* moves from its position. The mobile *A* can play the role of *C* is this analysis and a new game between *B* or *D* and *C* will be played.

All examples exhibit asymmetric realizations and non-reciprocity. We next show how such a situation can still be considered as symmetric (due to the fact that we consider distributions of nodes rather than realizations). Assume that the location of the transmitters follow a Poisson distribution with parameter λ over the two dimensional plane. Consider an arbitrary user A. Let *r* be the interference range. Then the number $K + 1$ of transmitters within the interference range of the receiver of A has a Poisson distribution with parameter $\lambda \pi r^2 / 2$. Since this holds for any node, the game is considered to be symmetric. The reason that the distribution is taken into account rather than the realization is that we shall assume that the actions of players will be taken before knowing the realization.

3.1.4 Spatial non-reciprocal random access games

We assume that all sources (players) have the same actions set (but their decisions can be different). The number of transmitters with which a given transmitter interacts may vary from one transmitter to another. We do not even need the reciprocity property: if transmitter TA interacts with transmitter TB, we do not require the converse to hold. We provide some example of multiple access and interference control to illustrate this non-reciprocity.

We consider local interactions between transmitters; for each transmitter there corresponds a receiver. We shall say that a transmitter TA is subject to an interaction (interference) from transmitter TB if the transmission from TB overlaps that from TA, and provided that *the receiver of the transmission from TA is within interference range of transmitter TB*. Consider the example depicted in Figure 3.4. It contains 3 receivers or destinations represented by squares, with range $\mathcal{N}_j$, $j = 1, 2, 3$. A transmission of a source *i* within a distance *r* of the receiver *Rj*, causes interference to a transmission from a source $j \neq i$ to receiver *Rj* but also *i* causes interference to any transmitter located in the an intersection range of coverage (the destinations of *i* and *j* can be different). We see that sources in area $A1 := \mathcal{N}_1 \backslash (A12 \cup A13 \cup A123)$ cause no interference to any transmitter from $A12 \cup A13 \cup A123$ transmitting to R2 or R3. The transmitters located in area $A1$ cause no interference to any transmitter located in areas A2 or A3. but the transmission from sources in A1 suffers from interference from source located in A12, A13, A123. The sources in A1 interfere with each other at their common destination R1. Thus, the interference is not reciprocal. In term of games, a transmitter in areas $A12 \cup A13 \cup A23 \cup A123$ play simultaneously several one-shot games with a unidimensional action (transmit or not) and receives a vector of payoff. Only the payoff obtained at the receiver in which he transmits can be non-zero.

Analysis of equilibria and optima

We consider arbitrary number of sources in each area. Let n_{ij} be the cardinal of *Aij* and n_{123} be the cardinal of *A123*. Denote by $r_1 := n_1 + n_{12} + n_{13} + n_{123}$ the number of transmitters located in $\mathcal{N}_1 = A12 \cup A13 \cup A1 \cup A123$ (the total number of transmitters covered by receiver R1. Similarly, $r_2 := n_2 + n_{12} + n_{23} + n_{123}$, $r_3 := n_3 + n_{13} + n_{23} + n_{123}$. A strategy profile is a (Nash) equilibrium if no source from any area can unilaterally deviate and improve its probability of success.

Equilibria and Pareto optima
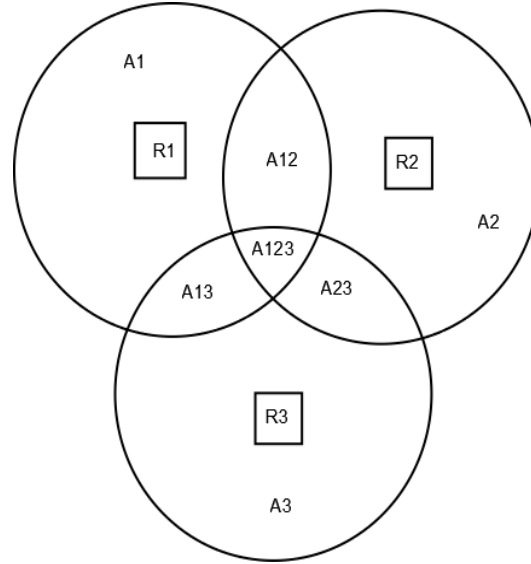
Figure 3.4: *Non-reciprocal interference control*

Consider the following configurations:

C3 Equilibria with three successful transmissions at the same slot:

– From each area $A1$, $A2$, $A3$, only one source transmits: only one source from $A1$, one from $A2$, one from $A3$ transmit and others stay quiet. These configurations are equilibria. There are three successful transmissions. Any permutation of the n_1 sources in $A1$, n_2 sources in $A2$ and n_3 sources in $A3$ when keeping the others sources in areas $A12$, $A13$, $A23$, $A123$ stay quiet lead a pure strategy equilibria that are also Pareto optima. They are global optima (social welfare) in the sense that the total sum of payoffs is maximized (one can have at most three successful transmissions at the slot).

C2 Equilibria with two successful transmissions at the same slot:

– Only one source located in $A12$ transmits, one source from $A3$ transmits and the others stay quiet.

– Only one source located in $A13$ transmits, one source from $A2$ transmits and the others stay quiet. There are two successful transmissions.

– Only one source located in $A23$ transmits, one source from $A1$ transmits and the others stay quiet.

C1 Equilibria with one successful transmission at the same slot:

– Only one source located in $A123$ transmits and all the others sources stay quiet. There

48

is one successful transmission in all the three receivers. Since this source is located at the intersection between $\mathcal{N}_1$, $\mathcal{N}_2$ and $\mathcal{N}_3$, there is interference to any transmission at any receiver. By permutation and selection of one source of the n_{123} sources located $A123$, we obtain the others equilibria at this configuration. These equilibria are not global optima.

C0 At least two sources from each neighborhood $\mathcal{N}_j$, $j = 1, 2, 3$ transmit and the others play any mixed strategies. Hence, there are infinite number of equilibria of this type. These equilibria are not Pareto optima because each source gets zero which is Pareto dominated by the configurations of type $C1 - C3$.

Notice that if there is no pairwise intersections between the neighborhoods $\mathcal{N}_j$, $j = 1, 2, 3$, (the number $n_{12}, n_{123}, n_{23}, n_{13}$ are null) then the interactions becomes reciprocal (if source i interferes with source j then source j interferes with source i) and the games in $\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3$ can be analyzed separately.

Notes

This model can be easily extended to finitely many intersecting neighborhoods. See Chapters 4 and 6.

3.2 Why random number of players?

The classical evolutionary game formalism is a central mathematical tool developed by biologists for predicting population dynamics in the context of interaction between populations. In order to make use of the wealth of tools and theory developed in the biology literature, many works in the area of computer networks (88; 20; 52; 214) ignore cases where local interactions between populations involve more than two individuals. This restriction limits the modeling power of evolutionary games which are not useful in a network operating at heavy load, such as ad-hoc networks with high density. This motivated us to consider a random number of users interacting locally.

We model and study interactions in wireless networks using the theory of evolutionary games which we extend to cover a random number of players. We study access games within a large population of mobiles that interfere with each other through many local interactions. Each local interaction involves a random number of mobiles. The games are not necessarily *reciprocal* as the set of mobiles causing interference to a given mobile may differ from the set of those suffering from its interference. We then apply evolutionary games to non-cooperative uplink power control in wireless networks. Specifically, we focus our study in a power control in W-CDMA and WiMAX wireless systems. We study competitive power control within a large population of mobiles that interfere with each other through many local interactions. We show how the evolution dynamics, the equilibrium and the evolutionary stability behavior are influenced by the characteristics of the wireless channel and pricing characteristics.

3.3 Model and notations

We describe in this part notations of our model.

- There is a large population of users.

- We assume that there are finite number of pure strategies in each population. Each member of the population chooses from the same set of strategies $\mathcal{A} = \{1, 2, \ldots, N\}$.

- Let $Mix = \Delta(\mathcal{A}) := \{(x_a)_{a \in \mathcal{A}} \mid x_a \geq 0, \sum_{a \in \mathcal{A}} x_a = 1\}$ the set of probability distributions over the N strategies. It is also interpreted as the set of distributions of strategies among the population, where x_a represents of proportion of users choosing the strategy a. A distribution x is sometime called the "state" or "profile" of the population. It can also be interpreted as a strategy. M can be interpreted as the set of mixed strategies of a user.

- The number of opponents M of a user is a random variable in the finite set $\{0, 1, \ldots\}$. In the bounded case, we will denote by $k_{\max}$ the maximum number of opponents interacting simultaneously with a user. This value depends on the node density and the transmission range. When making a choice of a strategy, a player knows the distribution of M but not its realization.

- The payoff of all players functions (identical for each member of the population) of the player's own behavior and opponents' behavior. Each user's payoff depends on opponents' behavior through the distribution of opponents' choices and of their number. The expected payoff of a user playing strategy j when the state of the population is x, is given by

$$f_a(x) = \sum_{k \geq 0} \mathbb{P}(\mathcal{K} = k) u_k(a, x, \ldots, x), \ a \in \mathcal{A}$$

where u_k is the payoff function given that the number of opponents is k. Although the payoffs are symmetric, the actual interference or interactions between two players that use the same strategy need not be the same, allowing for non-reciprocal behavior. The reason is that the latter is a property of the random realization whereas the actual payoff already averages over the randomness related to the interactions, the number of interfering players, the topology etc.

- The game is played many times.

Theorem 3.3.1 (Existence of equilibrium). *For any distribution of random number K of players that interact locally such that the expected payoff is continuous in the simplex (or product of simplexes for multipopulation case), the evolutionary game with random number of interacting players has at least one equilibrium state.*

Proof. We show that the game has a symmetric Nash equilibrium. We first remark that the generating function of $\mathcal{K}$ is continuous in $(0, 1)$. Thus, F is lower semi-continuous in $\Delta(\mathcal{A})$ (which is a non-empty, convex and compact subset of Euclidean space $\mathbb{R}^{n+1}$). The existence of symmetric Nash equilibrium in mixed strategies follows from Kakutani fixed point theorem or from the existence of solutions of the following variational inequalities

$$\text{find } x \in \Delta_n \text{ s.t } \sum_{b \in \mathcal{P}} (x_b - mut_b) f_b(x) \geq 0, \ \forall \ mut$$

$\square$

3.4 Slotted Aloha

Multiple Access Game introduces the problem of medium access. We assume that mobiles are randomly placed over a plane. All mobiles use the same fixed transmission range of r. The channel is ideal for transmission and all errors are due to collision. A mobile decides to

transmit a packet or not to transmit to a receiver when they are within transmission range of each other. Interference occurs as in the Aloha protocol: if more than one neighbors of a receiver transmit a packet at the same time then there is a collision. The Multiple Access Game is a nonzero-sum game, the mobiles have to share a common resource, the wireless medium. In this game, the parameter μ represents the probability that a mobile has its receiver $R(i)$ within its range. When a mobile i transmits to $R(i)$, all mobiles within a circle of radius r centered at $R(i)$ cause interference to the node i for its transmission to $R(i)$. This means that more than one transmission within a distance r of the receiver in the same slot cause a collision and the loss of mobile's i packet at $R(i)$.

Each of the mobiles has two possible strategies: either to transmit (T) or to stay quiet (S). If mobile i transmits a packet, it incurs a transmission cost of $\delta \geq 0$. The packet transmission is successful if the other users don't transmit (stays quiet) in that given time slot, otherwise there is a collision and the corresponding cost is $\Delta \geq 0$. If there is no collision, user i gets a reward of V from the successful packet transmission. We suppose that the reward V is greater than the cost of transmission δ. When all users stay quiet, they have to pay a regret cost κ. If $\kappa = 0$ the game is called *degenerate multiple access game*. Figure 3.27 represents an example of interaction of three nodes. The node 1 chooses a row, node 2 chooses a column and node 3 an array. The ESS corresponding to any number of nodes[1] of this game is given in theorem 3.4.1.

	T	S		T	S
T	$(-B_\theta, -B_\theta, -B_\theta)$	$(-B_\theta, 0, -B_\theta)$	T	$(-B_\theta, -B_\theta, 0)$	$(V - \delta, 0, 0)$
S	$(0, -B_\theta, -B_\theta)$	$(0, 0, V - \delta)$	S	$(0, V - \delta, 0)$	$(-\kappa, -\kappa, -\kappa)$

Figure 3.5: *Multiple access game with three nodes. $B_\theta = \Delta + \delta$*

Let $\mathcal{A} := \{T, S\}$ be the set of pure strategies. An equivalent interpretation of strategies is obtained by assuming that individuals choose pure strategies and then the probability distribution represents the fraction of individuals in the population that choose each strategy. We denote by s (resp. $1 - s$) the population share of strategy T (resp. S). The payoff obtained by a node with k opponents when it plays T is $u_k(T, s) = (-\Delta - \delta)(1 - \eta_k) + (V - \delta)\eta_k$ where $\eta_k := (1 - s)^k$, and the node-mutant receives $u_k(S, s) = -\kappa(1 - s)^k$ when it stays quiet. The expected payoff of an anonymous transmitter node-mutant is given

$$
\begin{aligned}
f_T(s) &= \mu \sum_{k \geq 0} \mathbb{P}(\mathcal{K} = k) u_k(T, s) \\
&= \mu \left(-(\Delta + \delta) + (V + \Delta) \sum_{k \geq 0} \mathbb{P}(\mathcal{K} = k)(1 - s)^k \right) \\
&= -\mu(\Delta + \delta) + \mu(V + \Delta) G_K(1 - s).
\end{aligned}
$$

where G_K is the generating function of K. Analogously, we have

$$
f_S(s) := \mu \sum_{k \geq 0} \mathbb{P}(\mathcal{K} = k) u_k(S, s) = -\mu\kappa \sum_{k \geq 0} (1 - s)^k \mathbb{P}(\mathcal{K} = k).
$$

The expected payoff of any individual in the population where s is the proportion of mobiles which transmit, is given by: $s f_T(s) + (1 - s) f_S(s)$.

We next introduce two alternative information scenario that have an impact on the decision making. In the first case, a mobile does not know whether there are zero or more other mobiles

---

[1]The one-shot game with n nodes has $2^n - 1$ Nash equilibria and a unique ESS.

in a given local interaction game about to be played. In the second case the mobile has this information, and consequently he transmits with probability one in case no other potential interferers are present. In addition to studying these two cases we shall also consider a third case called the "massively dense" ad-hoc network in which, whenever a mobile participates wishes to transmit, there is at least one other mobile that is involved in the local interaction game.

We denote $\alpha := \dfrac{\Delta + \delta}{V + \Delta + \kappa}$, which represents the ratio between the collision cost $-\Delta - \delta$ (cost when there is a collision during a transmission) and the difference between global cost perceived by a mobile $-\Delta - \delta - \kappa$ (collision and regret) and the benefit $V - \delta$ (reward minus transmission cost). When the collision cost Δ becomes high, the value α converges to one and when the reward or regret cost becomes high, the value α is close to zero.

Case 1 : Aloha without sensing

A transmitter does not know if there are other transmitters at the range of its receiver. Then, even when it is the only transmitter, it has to decide to transmit or not.

Theorem 3.4.1. *If* $\mathbb{P}(\mathcal{K} = 0) < \dfrac{\Delta + \delta}{V + \Delta + \kappa} =: \alpha$, *then the game has a unique ESS* s_1^* *given by* $s_1^* = \phi^{-1}(\dfrac{\Delta + \delta}{V + \Delta + \kappa})$ *where* $\phi : s \mapsto \sum_{k \geq 0} \mathbb{P}(\mathcal{K} = k)(1 - s)^k$.

Proof. A mixed equilibrium s is characterized by $f_T(s) = f_S(s)$ i.e $\phi(s) = \dfrac{\Delta + \delta}{V + \Delta + \kappa}$. The function ϕ is continuous and strictly decreasing monotone on $(0, 1)$ with $\phi(1) = \mathbb{P}(\mathcal{K} = 0)$ and $\phi(0) = 1$. Then the equation $\phi(s) = \dfrac{\Delta + \delta}{V + \Delta + \kappa}$ has a unique solution in the interval $(\mathbb{P}(\mathcal{K} = 0), 1)$. One has, $f(s, y) - f(mut, y) = \mu(V + \Delta + \kappa)(s - mut)(\phi(y) - \phi(s))$. Since $s - \epsilon mut - (1 - \epsilon)s = \epsilon(s - mut)$, for $y = \epsilon mut + (1 - \epsilon)s$ one has $\sum_{j \in \{T,S\}} (x_j - mut_j)f_j(y) > 0$ (because ϕ is strictly decreasing continuous function) for all $mut \neq s$. This completes the proof. $\square$

When a mobile is never alone in his interference area, i.e. $\mathbb{P}(\mathcal{K} = 0) = 0$; the condition $\alpha > 0$ is satisfied.

Case 2: Aloha with sensing

A mobile knows when it is the only transmitter at the range of its receiver, and when it is it will thus transmit with probability one. We can say then that the action set is T whenever a user has opponents in a local interaction.

Theorem 3.4.2. *An anonymous user without opponents receives the fitness* $f_0 = V - \delta$. *If* $\mathbb{P}(\mathcal{K} = 0) < \dfrac{\Delta + \delta}{V + \Delta}$, *then the game has a unique ESS* s_2^* *given by* $s_2^* = \phi^{-1}\left(\dfrac{\Delta + \delta + \kappa \mathbb{P}(\mathcal{K} = 0)}{V + \Delta + \kappa}\right)$ *where* $\phi : s \mapsto \sum_{k \geq 0} \mathbb{P}(\mathcal{K} = k)(1 - s)^k$.

Proof. The proof is similar as in theorem 3.4.1. $\square$

Case 3: Massively Dense

In this case, we take into account only local interactions between users. Then, in this case, mobiles are never alone to transmit during a slot and we have $\sum_{k \geq 1} \mathbb{P}(\mathcal{K} = k) = 1$.

Theorem 3.4.3. *The game has always an unique ESS which is solution of the following equation* $\sum_{k \geq 1} \mathbb{P}(\mathcal{K} = k)(1-s)^k = \alpha$.

The proof is similar as in theorem 3.4.1 using the monotocity of g.

Proposition 3.4.3.1. *The ESS given in theorems 3.4.1,3.4.2,3.4.3 is asymptotically stable in the replicator dynamics without delays for all non-trivial initial state ($s_0 \notin \{0,1\}$).*

Proof. The replicator dynamics is given by

$$\frac{d}{dt}s(t) = (V + \Delta + \kappa)s(t)(1-s(t))(\phi(s(t)) - \alpha).$$

The function ϕ is decreasing on $(0,1)$ implies that the derivative of the function $s(1-s)(\phi(s) - \alpha)$ at the ESS is negative. These means that the dynamic is *competitive*. Hence, the ESS is asymptotically stable. □

3.4.4 ESS and nodes distribution

In this subsection, we consider different nodes distributions. We study the existence and the uniqueness of ESS in the different nodes distributions. First one, we assume that all mobiles have the same number of neighbors $n-1$, i.e., $\mathbb{P}(\mathcal{K} = j) = \delta_{n-1}(j)$ and seconde one, we assume that nodes are randomly distributed on a plan following a Poisson point process with density λ.

Arbitrary number of opponent nodes

In this part, we suppose that the population of nodes is composed with many local interaction between $n \geq 2$ nodes. Let $\mathcal{A} := \{T, S\}$ the set of strategies and assume that the strategy T has a delay τ_T and the strategy S has the delay τ_S. The payoff of a player using the action $a_i \in \mathcal{A}$ against the other players when they use the multi-strategy $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ is given by $U_i(a)$.

Each user plays the $n-$player following game $\Gamma_n = (\mathcal{N}, \mathcal{A}, (U_i)_{i \in \mathcal{N}})$ where

- $\mathcal{N}$ is the set of users (nodes) and the cardinal of $\mathcal{N}$ is n,

- $\mathcal{A}$ the set of pure actions (the same for every user),

- for every user i in $\mathcal{N}$, the payoff function $U_i : \mathcal{A}^n \to \mathbb{R}$ is given by

$$U_i(a) = \begin{cases} V - \delta & \text{if} & a_i = T \text{ and } a_j = S, \forall j \neq i \\ 0 & \text{if} & a_i = S \text{ and } \{j \in \mathcal{N} \mid a_j = T\} \geq 1 \\ -\Delta - \delta & \text{if} & a_i = T \text{ and } \{j \in \mathcal{N} \mid a_j = T\} \geq 2 \\ -\kappa & \text{if} & a_j = S \ \forall j \in \mathcal{N} \end{cases}$$

Let s be the proportion of nodes in the population using the strategy T. Then $x = (s, 1 - s)$ is the state of the population. Let $\Delta(\mathcal{A}) := \{sT + (1 - s)S \mid 0 \le s \le 1\}$ the set of mixed strategies. The average payoff is

$$f(x, x) = \mu s \left[(-\Delta - \delta)\left(1 - (1 - s)^{n-1}\right) + (V - \delta)(1 - s)^{n-1} \right] - \mu \kappa (1 - s)^n.$$

It is not difficult to see that the one-shot game Γ_n has $2^n - 1$ Nash equilibrium, n of them are optimal in Pareto sense[2]:

- If only one node *transmit* and the others *stay quiet*, then the node which transmit gets the payoff $V - \delta$ and the others receive nothing and has no cost. This configuration is an equilibrium .

- There are exactly n pure equilibria and all these pure equilibria are Pareto optimal.

- k $(1 \le k < n - 1)$ of the n nodes choose to stay quiet and the $n - k$ others are active and play the optimal mixed strategy in the game $\Gamma_{n-k} : \left(1 - \alpha^{\frac{1}{n-k-1}}, \alpha^{\frac{1}{n-k-1}}\right)$ where $\alpha := \dfrac{\Delta + \delta}{V + \Delta + \kappa}$. Thus, there are exactly $\sum\limits_{k=1}^{n-2} \binom{n}{k} = 2^n - (n + 2)$ partially mixed Nash equilibria.

- The game has a unique strictly mixed Nash equilibrium given by $\left(1 - \alpha^{\frac{1}{n-1}}, \alpha^{\frac{1}{n-1}}\right)$

- the allocation of payoff obtained in these (partially or completely) mixed strategy are not Pareto optimal.

Note that the first interference scenario described in the previous section holds here because the number of interferes is fixed and is equal to $n - 1$. Then from Theorem 3.4.1 with the function $\phi(s) = (1 - s)^{n-1}$, the ESS exists and is uniquely defined by $s^* = 1 - \alpha^{\frac{1}{n-1}}$.

This result generalizes the ESS in the two-player case that we have shown in (216) that when there are pairing local interaction in a single population, the completely mixed (Nash) equilibrium of the two nodes stage game Γ_2 is an ESS.

Optimization of the total throughput

We assume here that all mobile has the same number of opponents, that is $n - 1$. We look for the probability of success that can be achieved in a local interaction depending on the Dirac distribution and also cost parameters. At the equilibrium point, the probability of success of a node is given by $s^*(1 - s^*)^{n-1}$ and the total probability to have a successful transmission in a local interaction (total throughput) is given by

$$P_{succ}(\alpha, n) = n \mu s^* (1 - s^*)^{n-1} = n \mu (1 - \alpha^{\frac{1}{n-1}}) \alpha, \tag{3.1}$$

where μ is the probability that a mobile has a receiver in its range. The probability to have a successful transmission $P_{succ}(\alpha, n) = n \mu (1 - \alpha^{\frac{1}{n-1}}) \alpha$ goes to $-\mu \alpha \log(\alpha)$ when the number of nodes goes to infinity. Hence, when n is large, the maximum total throughput of a local interaction is obtained when $\alpha = 1/e$ and is closed to μ/e. The optimal total throughput is

---

[2] An allocation of payoffs is Pareto optimal or Pareto efficiency if there is no other allocation that makes every node at least as well off and at least one node strictly better off.

obtained when $\alpha^* = (1 - \dfrac{1}{n})^{n-1}$ and the corresponding throughput is $P_{succ}(\alpha^*, n) = \mu(1 - \dfrac{1}{n})^{n-1} \longrightarrow \dfrac{\mu}{e}$.

In Figure 3.10 we observe the total throughput depending on the number of interferes n. The parameters considered are $\mu = 0.8$ and $\alpha = 1/3$. We denote that the total throughput is increasing in that case with the number of interferes which it seems non intuitive. The reason is that the number of transmitted mobiles at the ESS, i.e. s^*, is exponentially decreasing with n. Another important result is that it may have a finite number of interferers that maximize the total throughput like in figures 3.7($\alpha = 0.2$). When the ratio α is very small $\alpha = 0.05$ the probability of success is decreasing in n as shown in Fig. 3.6. In Fig. 3.11 we represent the probability of success $P_{succ}(\alpha, n)$ as a function α for several values of n.
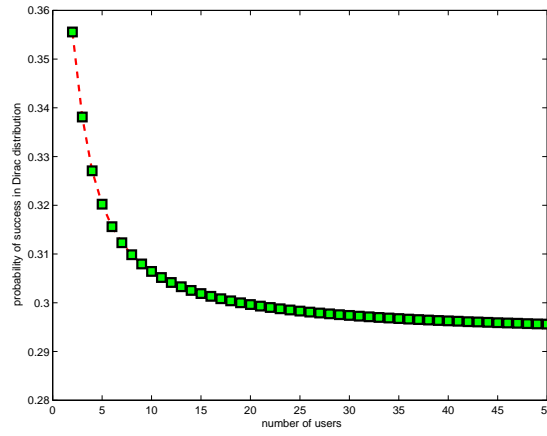
Figure 3.6: *Impact of n in the probability of success in Dirac distribution for $\alpha = 0.05$*
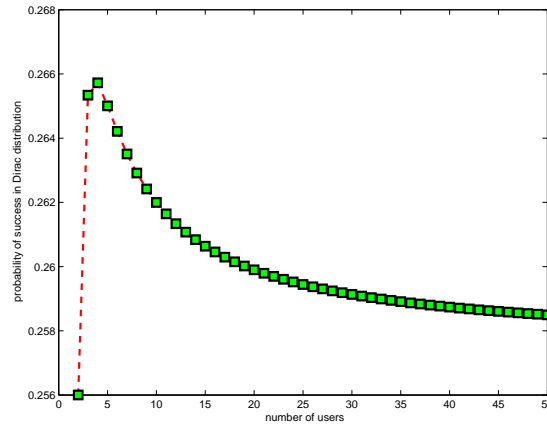
Figure 3.7: *Impact of n in the probability of success in Dirac distribution for $\alpha = 0.2$.*

In figure 3.11, we observe the total throughput for different value of n and applying the result of the previous proposition, we obtain the optimal total throughput depending on α.

Impact of the time delays Now, we study the effect of the time delays on the convergence of
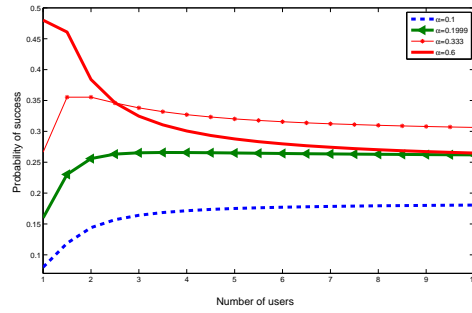
Figure 3.8: Impact of n in the probability of success in Dirac distribution.
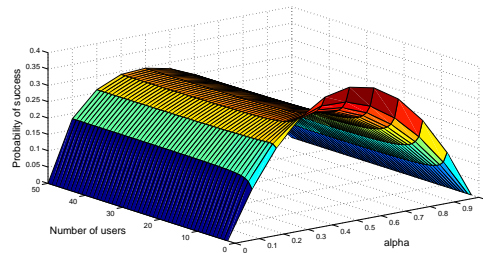
Figure 3.9: Probability of success in Dirac distribution as function of α and n.
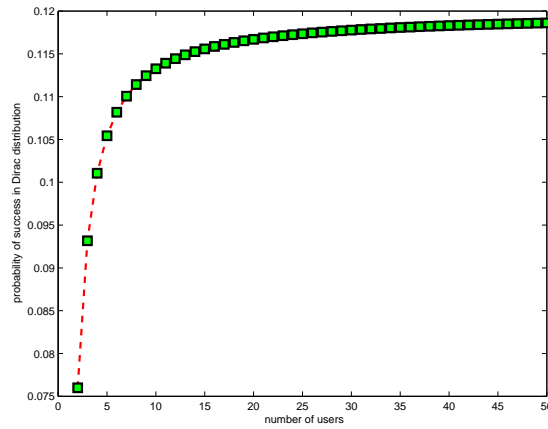
Figure 3.10: Impact of n in the probability of success in Dirac distribution for $\alpha = 1/3$.

replicator dynamics to the evolutionarily stable states in which each pure strategy is associated with its own delay. Let τ_T (resp. τ_S) be the time delay of the strategy (T) (resp. (S)). The delayed replicator equation becomes

$$\dot{s}(t) = \mu\, s(t)(1 - s(t)) \left[f_T(s(t - \tau_T)) - f_S(s(t - \tau_S)) \right] \tag{3.2}$$

where $f_T(s(t)) := -\mu(\Delta + \delta)\left(1 - (1 - s(t))^{n-1}\right) + \mu(V - \delta)\left(1 - s(t)\right)^{n-1}$ and $f_S(s(t)) := -\mu\kappa\left(1 - s(t)\right)^{n-1}$.

In order to study the asymptotically stability of the replicator dynamics (3.2) around the
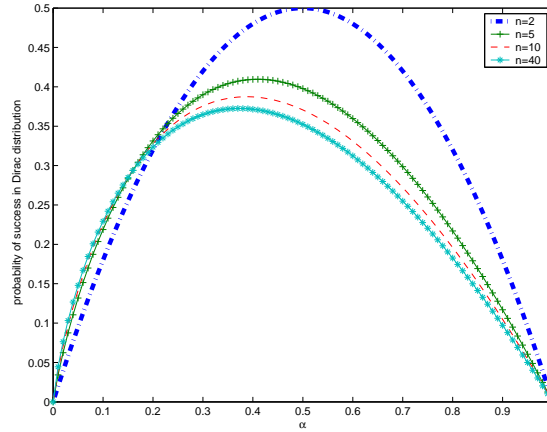
Figure 3.11: Probability of success in Dirac distribution.

unique ESS $s_1^* = 1 - (\dfrac{\Delta + \delta}{V + \Delta + \kappa})^{\frac{1}{n-1}}$, we linearize (3.2) at s_1^*. We obtain the following linear delay differential equation

$$\dot{z}(t) = -\mu(n-1)s(1-s)^{n-1}\left((V+\Delta)z(t-\tau_T) + \kappa z(t-\tau_S)\right) \tag{3.3}$$

where $z(t) = s(t) - s_1^*$. The following theorem gives a sufficient conditions of stability of (3.3) at zero.

Theorem 3.4.5. *Suppose at least one of the following conditions holds*

- $(V+\Delta)\tau_T + \kappa\tau_S < \dfrac{1}{(n-1)s(1-s)^{n-1}\mu}$

- $V + \Delta > \kappa$ *and* $(V+\Delta)\tau_T < \dfrac{V+\Delta-\kappa}{(n-1)s(1-s)^{n-1}\mu(V+\Delta+\kappa)}$

- $V + \Delta < \kappa$ *and* $\kappa\tau_S < \dfrac{-V-\Delta+\kappa}{(n-1)s(1-s)^{n-1}\mu(V+\Delta+\kappa)}$

Then the ESS s is asymptotically stable.

A proof of the theorem 3.4.5 can be obtained using theorem 3 in (214) applying to equation (3.3).

A necessary and sufficient condition of stability of (3.3) at zero when delays are symmetric is given in theorem 3.4.6.

Theorem 3.4.6 (symmetric delay). *Suppose that $\tau_T = \tau_S = \tau$, then, the ESS s_1^* is asymptotically stable if and only if*

$$\tau < \frac{\pi}{2(n-1)\mu s_1^*(1-s_1^*)^{n-1}(V+\Delta+\kappa)}$$

The proof uses the following well known lemma (see (215) and the references therein).

Lemma 3.4.6.1. *The trivial solution of the linear delay differential equation*

$$\dot{z}(t) = -az(t-\tau), \ \tau, a > 0$$

is asymptotically stable if and only if $2a\tau < \pi$.

Numerical solutions The numerical solutions given figure 3.12 are obtained when the random variable K is a Dirac $\delta_{\{n-1\}}$. We took $n = 4 = k_{max}, \Delta = \dfrac{1}{4} = \delta, V = 1$. The initial condition is 0.02 and the delays τ_T and τ_S between 0.02 and 7. For the small delays: $\tau_T = 0.02, \tau_S = 0.02$ and $\tau_T = 3, \tau_S = 2$ respectively, the system is stable. For the delay $\tau_H = 7$ and $\tau_S = 5$, the system is unstable and the proportion of transmitters in the cell oscillates around the ESS.
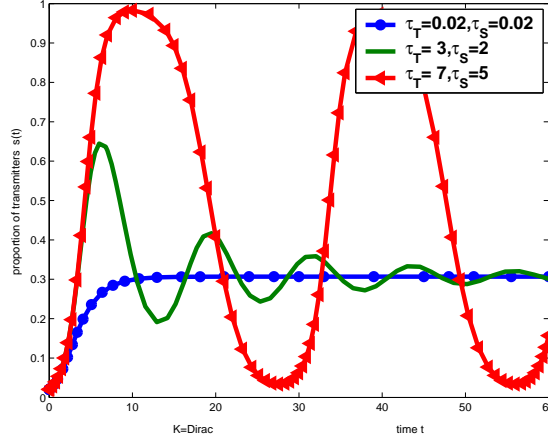
Figure 3.12: Delay effect in Dirac Distribution.

Poisson distribution

We consider that nodes are distributed over a plan following a Poisson distribution with density λ. The probability that a node has k neighbors is given by the following distribution.

Cases 1 and 2: $\mathbb{P}(\mathcal{K} = k) = \dfrac{(\lambda \pi r^2)^k}{k!} e^{-\lambda \pi r^2}$, $k \geq 0$. Case 3: $\mathbb{P}(\mathcal{K} = k) = \dfrac{(\lambda \pi r^2)^{k-1}}{(k-1)!} e^{-\lambda \pi r^2}$, $k \geq 1$.

Considering those node distributions and from previous theorems, the unique ESS s^* for all cases, is solution of the following equation :

$$
\begin{cases}
e^{-\lambda \pi r^2 s_1} = \alpha & \text{for case 1} \\
e^{-\lambda \pi r^2 s_2} = \alpha + \dfrac{\kappa \mathbb{P}(\mathcal{K} = 0)}{V + \Delta + \kappa} & \text{for case 2} \\
(1 - s_3) e^{-\lambda \pi r^2 s_3} = \alpha & \text{for case 3}
\end{cases}
$$

Thus we obtain the following equilibria in the different scenario:

$$
s_1^* = \log\left(\alpha^{-\frac{1}{\lambda \pi r^2}} \right), s_2^* = \log\left((\alpha + \dfrac{\kappa \mathbb{P}(\mathcal{K} = 0)}{V + \Delta + \kappa})^{-\frac{1}{\lambda \pi r^2}} \right)
$$

$$
\text{and } s_3^* = 1 - \dfrac{LambertW(\lambda \pi r^2 \alpha e^{\lambda \pi r^2})}{\lambda \pi r^2},
$$

where $LambertW(s)$ is the LambertW function which is the inverse function of $f(w) = we^w$.

Optimization of the total throughput in Poisson distribution We look for the probability of success that can be achieved in a local interaction depending on distribution parameters and also cost parameters. We consider the Poisson distribution with parameters λ and r.

The probability to have a successful transmission in a local interaction (total throughput) is given by different equation depending on the scenario. In the case 1 we have:

$$P_{succ}(\alpha, \lambda) \quad = \mu s_1^* \sum_{k \geq 0} k \mathbb{P}(\mathcal{K} = k)(1 - s_1^*)^k = \mu s_1^* \sum_{k \geq 0} k \frac{(\lambda \pi r^2)^k}{k!}(1 - s_1^*)^k$$

$$\approx \mu s_1^*(1 - s_1^*)\lambda \pi r^2 \alpha,$$

In the case 2, we have: $P_{succ}(\bar{\alpha}, \lambda) \approx \mu s_2^*(1 - s_2^*)\lambda \pi r^2 \bar{\alpha}$. We derive immediately the following result:

Proposition 3.4.6.2. *The maximum total throughput under poisson distribution is attained when $\alpha = e^{h(\lambda, r)}$ in the case 1 (resp. $\bar{\alpha} = e^{h(\lambda, r)}$ in the case 2) where h is one of the two functions defined by*
$$(\lambda, r) \in \mathbb{R}_+^2 \mapsto \frac{-(1 + 2\lambda \pi r^2) \pm \sqrt{1 + 4(\lambda \pi r^2)^2}}{2}.$$

In the case 3, we have: $P_{succ}(\alpha, \lambda) = \mu s_3^* \sum_{k \geq 1} k P(\mathcal{K} = k)(1 - s_3^*)^k = \mu s_3^* \sum_{k \geq 1} k \frac{(\lambda \pi r^2)^{k-1}}{(k-1)!}(1 -$
$s_3^*)^k \approx \mu \alpha s_3^*(1 + \lambda \pi r^2(1 - s_3^*))$.

Proposition 3.4.6.3. *There exists a unique α_3^* in which the total throughput is maximum when $\alpha = \alpha_3^*$. The α_3^* is given by $\alpha_3^* = (1 - s)e^{-\lambda \pi r^2 s}$ where s is the unique solution in $[0, 1]$ of the following equation :*

$$1 + \gamma - s(2 + 5\gamma + \gamma^2) + s^2(4\gamma + 2\gamma^2) - \gamma^2 s^3 = 0$$

Proof. The derivative of the function $H := \frac{\partial P_{succ}}{\partial s}$ s given by

$$H(s) = (1 + \gamma - s(2 + 5\gamma + \gamma^2) + s^2(4\gamma + 2\gamma^2) - s^3\gamma^2)e^{-\gamma s}.$$

We prove that the above function is strictly decreasing on $[0, 1]$. For that, it is sufficient to study the following function

$$G(s) = 1 + \gamma - s(2 + 5\gamma + \gamma^2) + s^2(4\gamma + 2\gamma^2) - s^3\gamma^2.$$

We have $\frac{\partial G(s)}{\partial s}$ is given by

$$\frac{\partial G(s)}{\partial s} = -(2 + 5\gamma + \gamma^2) + 2s(4\gamma + 2\gamma^2) - 3s^2\gamma^2.$$

It is easy to show that the above function is always negative. Since $H(0) = 1 + \gamma > 0$ and $H(1) = -e^{-\gamma} < 0$ then the function H is positive for $s \in [0, \bar{s})$ and is negative for $s \in (\bar{s}, 1]$ where $\bar{s}$ is the solution of the equation $G(s) = 0$. Since s^* is decreasing function on α, we conclude that function P_{succ} is positive if $s \in [0, \bar{s})$ and is negative $s \in (\bar{s}, 1]$. Since the optimal of function P_{succ} is attained at $\alpha = (1 - \bar{s})e^{-\lambda \pi r^2 \bar{s}}$ □

The probability of success at the ESS in poisson distribution is represented in figures 3.13 and 3.14. We observe in particular case that when the number of interferes increases, i.e. the rate λ in the case of the Poisson distribution, the total throughput increases.

In the figures 3.15 and 3.16, we describe numerical application of our evolutionary game model with Poisson distribution of nodes under the replicator dynamics. We took $n = 4 = k_{\max}, \Delta = \frac{1}{4} = \delta = \kappa, \lambda = 1$ and $V = 1$. The initial condition in all these figures is 0.02. In
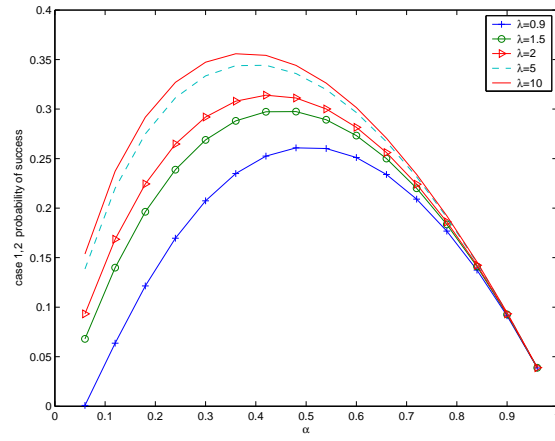
Figure 3.13: Probability of success in Poisson distribution (cases 1,2).
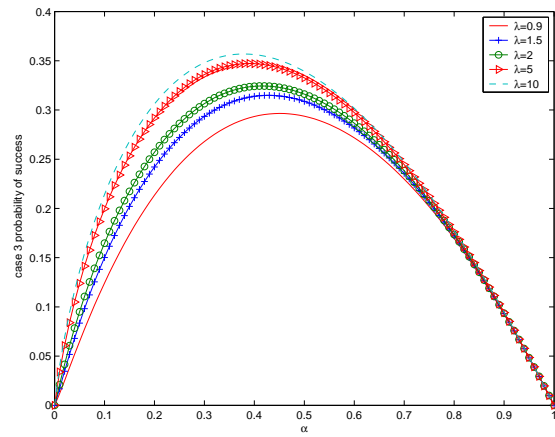
Figure 3.14: Probability of success in Poisson distribution (case 3).
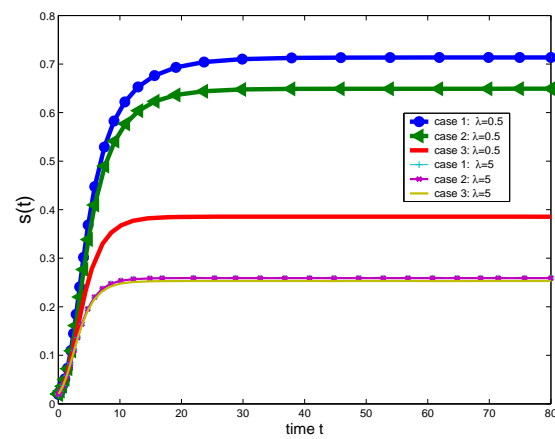
Figure 3.15: Evolution of the fraction of transmitters varying the density parameter λ.

60

the figure 3.15 we compare the evolution of the fraction of transmitters varying the parameter of density λ between 0.1 and 5 for the case 1, 2 and 3 respectively. We observe that we have stability for all cases. In figure 3.16 represents the impact of the parameter μ on the velocity of the system. We took μ between 0.1 and 0.5 without delay. We observe that we have stability but the convergence speed becomes slow when μ decreases.
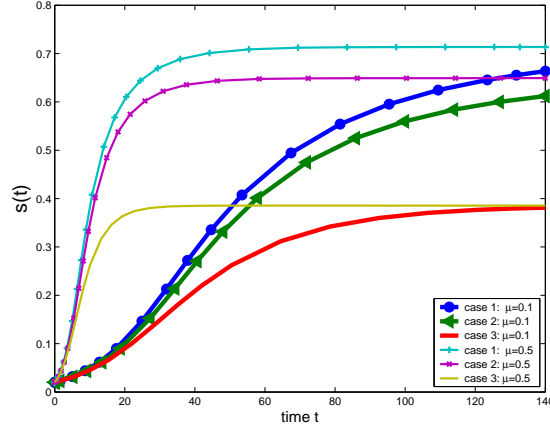
Figure 3.16: *Impact of μ on the velocity of the replicator dynamics without delay*
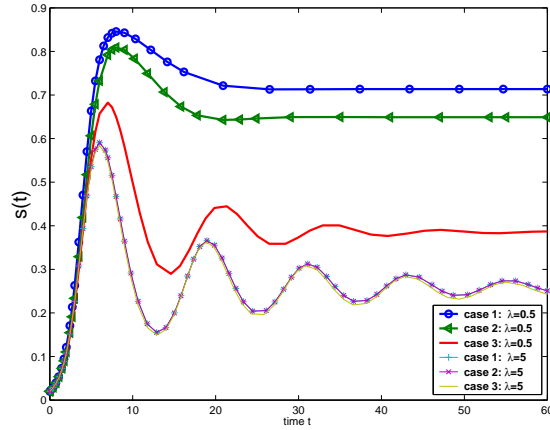
Figure 3.17: *Evolution of transmitters by varying the density parameter λ.*

Now, we study the effect of the time delays on the convergence of replicator dynamics to the evolutionarily stable states in which each pure strategy is associated with its own delay in Poisson distribution. The replicator dynamics becomes

$$\dot{s}(t) = \mu s(t)(1 - s(t)) \left[f_T(s(t - \tau_T)) - f_S(s(t - \tau_S)) \right] \tag{3.4}$$

where $f_T(s(t)) := \mu \left(-(\Delta + \delta) + (V + \Delta)e^{-\lambda \pi r^2 s(t)} \right)$ and $f_S(s(t)) := -\mu \kappa e^{-\lambda \pi r^2 s(t)}$ in the case 1.In order to study the asymptotically stability of the replicator dynamics (3.4) around the unique ESS, we linearize (3.4) at $s^* = s_1^*$. We obtain the following linear delay differential equation

$$\dot{y}(t) = -c_1((V + \Delta)y(t - \tau_T) + \kappa y(t - \tau_S))$$
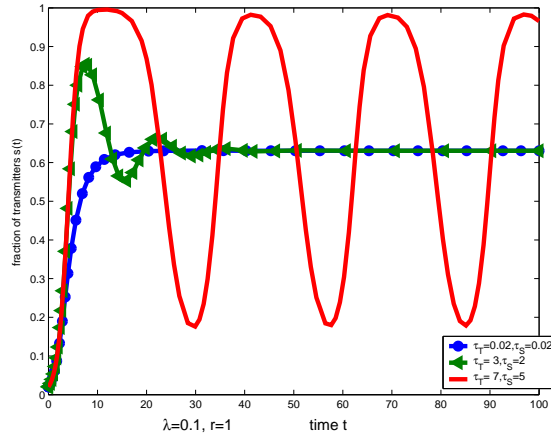
Figure 3.18: Impact of the time delay on replicator dynamics (case 1).

where $c_1 := \mu s^*(1-s^*)\alpha\left(1+s^*(1-s^*)\lambda\pi r^2\right)$, and $y(t) = s(t) - s^*$. The following theorem give sufficient conditions of stability of (3.5) at zero.

Theorem 3.4.7 (see (214)). *Suppose at least one of the following conditions holds*

(i) $(V+\Delta)\tau_T + \kappa\tau_S < \theta_a$,

(ii) $V+\Delta > \kappa$ and $(V+\Delta)\tau_T < \dfrac{(V+\Delta-\kappa)\theta_a}{V+\Delta+\kappa}$,

(iii) $V+\Delta < \kappa$ and $\kappa\tau_S < \dfrac{(-V-\Delta+\kappa)\theta_a}{V+\Delta+\kappa}$ where $\theta_a := \dfrac{1}{s^*(1-s^*)\mu\alpha\left(1+s^*(1-s^*)\lambda\pi r^2\right)}$ Then the ESS s^* is asymptotically stable.

A necessary and sufficient condition of stability of (3.5) at zero when delays are symmetric is given in theorem 3.4.8.

Theorem 3.4.8 (symmetric delay). *Suppose that $\tau_T = \tau_S = \tau$. Then, the ESS s^* is asymptotically stable if and only if*

$$\tau < \frac{\pi}{2s^*(1-s^*)\mu\alpha\left(1+s^*(1-s^*)\lambda\pi r^2\right)(V+\Delta+\kappa)}$$

The fraction of transmitters in the population is represented in figure 3.18 for $\lambda = 0.5$ and $r = 1$. The delays τ_T and τ_S are between 0.02 and 7. The system is stable for $\tau_T = \tau_S = 0.02$ or $\tau_T = 3, \tau_S = 2$. For $\tau_T = 7$ and $\tau_S = 5$ the system is unstable. We display an oscillatory behavior of the population as function of time. The trajectory are seen to converge to periodic ones. All turn out to confirm the stability condition that we obtained in theorem 3.4.7. In the figure 3.17 we compare evolution of the fraction of transmitters varying the parameter of density λ between 0.1 and 5 for the case 1, 2 and 3 respectively. In this figure, the time delays are respectively 3 and 2. Note that in this figure the equilibrium point is decreasing function in the density parameter λ. Indeed, when the density of nodes increases, the number of mobiles share a receiver increases. To avoid collision, the nodes decrease the probability of transmission. We observe also that for $\lambda = 5$, we have stability but the convergence speed is slow than for $\lambda = 0.1$.

3.4.9 Coordination mechanisms to reduce collisions

In this subsection we focus on the concept of *correlated evolutionarily stable states* (CESS) in access games. A correlated equilibrium is introduced by Aumann (32) in 1974 and can be interpreted as a distribution of actions given to the mobiles by some *referee*(which can be the base station, the receiver) before to play each local game. For more details on correlated equilibrium, we refer the reader to (32; 33; 125). As in (125), we use the concept of correlated equilibrium from the perspective of bounded rationality.

We define a probability space $\left(\Omega, 2^\Omega, \mathbb{P}^n\right)$ which generates signals on which the nodes can condition their strategic choices where $\Omega = \mathcal{A}^n$. The set Ω is partitioned as follows $I^j(a) := \{as^1s^2 \ldots s^{j-1}s^{j+1} \ldots s^n \mid s^l \in \mathcal{A}, l \neq j\}, I^j := \{I^j(a), a \in \mathcal{A}\}.$ then, I^j has exactly $|\mathcal{A}| = 2$ elements which are information sets of node $j \in \mathcal{N}$. We define an *assignment function*(called also *rule*) profile $\alpha = (\alpha^1, \ldots, \alpha^n)$ as a mapping from the set of states Ω to mixed strategies set $\Delta(\mathcal{A})$. For all j, the assignment function of the node j, α^j must satisfy :
if for some w, $\alpha^j(w) = s$, then $\alpha^j(w') = s$, for all $w' \in I^j(s)$.

That is, for each element $w \in I^j$, node j cannot distinguish states that are in the same information set. We denote the set of all pure assignment functions by $\mathcal{AF} : \{g \mid g : \mathcal{A} \longrightarrow \mathcal{A}\}$. Thus, when a node chooses an assignment function α and when he receives the signal ω from the referee, he will choose the mixed action $\alpha(\omega)$. We use $\alpha(s^j|\omega)$ to denote the probability assigned on s^j under this mixed action $\alpha(\omega)$. Then $\alpha(w) = [\alpha(s^1|\omega), \ldots, \alpha(s^n|\omega)] \in \Delta(S)$.

Given a referee $(\Omega, \mathbb{P}^n)$, we define the identity assignment function as $\alpha^{id,n}(s^j|\omega) = 1$ if $proj^j(\omega) = s^j$ where $proj^j(\omega)$ denotes the j−th element of the signal ω and $\alpha^{id,n}(s^j|\omega) = 0$ for all ω such that $proj^j(\omega) \neq s^j$.

If each node use the identity assignment then, the resulting probability distribution of actions profile actually played will be $\mathbb{P}^n$, the same as the probability distribution of actions recommended by the referee. But when nodes use other assignments, a different distribution may result.

Given an assignment profile α and a probability distribution $\mathbb{P}^n$ over $\mathcal{A}^n$, the expected payoff is given by

$$
\begin{aligned}
f^n(\alpha) \qquad &= \sum_{\omega \in \Omega} \mathbb{P}^n(\omega) \sum_{a=(a^1,\ldots,a^n) \in \mathcal{A}^n} u^n(a) \prod_{j \in \mathcal{N}} \alpha(a^j|w) \\
&= \sum_{a=(a^1,\ldots,a^n) \in \mathcal{A}^n} u^n(a) \sum_{\omega \in \Omega} \mathbb{P}^n(\omega) \prod_{j \in \mathcal{N}} \alpha(a^j|w) = \sum_{a=(a^1,\ldots,a^n) \in \mathcal{A}^n} u^n(a) \mathbb{Q}_\alpha(a)
\end{aligned}
$$

where $\mathbb{Q}_\alpha(a) = \sum_{\omega \in \Omega} \mathbb{P}^n(\omega) \prod_{j \in \mathcal{N}} \alpha(a^j|w)$. We say that two assignment functions profile α and β are equivalent if they induce the same value $\mathbb{Q}_\alpha = \mathbb{Q}_\beta$

Proposition 3.4.9.1. *Suppose that $\mathbb{P}^n(a) = \mathbb{P}^n(\sigma(a))$ for all permutation σ (the distribution $\mathbb{P}^n$ is said symmetric). If an assignment profile β is equivalent to $\alpha^{id,n}$ then $\mathbb{Q}_\beta(a) = \mathbb{P}^n(a)$, for all $a \in \mathcal{A}^n$.*

Proof. This is because $\mathbb{Q}_{\alpha^{id,n}}(a) = \mathbb{P}^n(a)$, $\forall a \in \mathcal{A}^n$. $\qquad\qquad\square$

Now, suppose a small group of mutants appears. These mutants use a mutational assignment function α'^n. which is not equivalent to the identical assignment function $\alpha^{id,n}$., but they cannot change the referee recommendation. Let ϵ be the portion of the population which are mutants (who use α',) and $1 - \epsilon$ portion of the population are non-mutants who use $\alpha^{id,n}$. At

each time, n nodes are randomly chosen to play the strategic game Γ_n. In playing the game, the nodes have the same referee $(\Omega, \mathbb{P}^n)$.

A probability distribution $\mathbb{P}^n$ over $\mathcal{A}^n$ is an CESS if non-mutants with identity assignment function perform better than mutants with assignment functions that are not equivalent to the identical assignment function.

Definition 3.4.9.2. *A CESS $\mathbb{P}^n$ is a symmetric distribution probability over $\mathcal{A}^n$ such that for every assignment function α'^n nonequivalent to the identical assignment function $\alpha^{id,n}$, there exists some $\epsilon_{\alpha'^n} > 0$ such that*

$$F^n(\alpha^{id,n}, \epsilon \alpha'^n + (1-\epsilon)\alpha^{id,n}) > F^n(\alpha'^n, \epsilon \alpha'^n + (1-\epsilon)\alpha^{id,n}) \tag{3.5}$$

for all $\epsilon \in (0, \epsilon_{\alpha'^n})$ where $F^n(\beta, \alpha) = f^n(\beta, \alpha, \dots, \alpha)$

Proposition 3.4.9.3. *If s is an ESS then the product measure $s^{\star(n)}$ given by $s^{\star(n)}(a) = \prod_{j \in \mathcal{N}} s^j(a^j)$ is a CESS.*

Proof. It is easy to see that $s^{\star(n)}$ is a probability measure on $\mathcal{A}^n$. Since s is an ESS, $(s, \dots, s) \in \Delta(\mathcal{A})^n$ is a Nash equilibrium of the $n-$player game i.e $\sum_{a \in \mathcal{A}^n} u^n(a) \prod_{j \in \mathcal{N}} s^j(a^j) \geq \sum_{a \in \mathcal{A}^n} u^n(b_k, a_{-k}) \prod_{j \in \mathcal{N} \setminus \{k\}} s^j(a^j) \geq 0, \forall b_k \in \mathcal{A}$. This means that the product probability measure $s^{\star(n)}$ is an CESS. $\square$

Our principal motivation to study CESS is that, in general the expected payoff obtained in ESS can be improved by CESS distribution (see the example depicted in Fig. 3.19). Non-ESS can be CESS as shown in the following example.

Examples 3.4.9.4. *Consider the following matrix game*

$$\Gamma_1((V=4, \Delta = \delta = \kappa = 1)): \quad \begin{array}{c|c|c} & S & T \\ \hline S & -1,-1 & 0,3 \\ \hline T & 3,0 & -2,-2 \end{array}, \quad \mathbb{P}^1: \begin{array}{c|c|c} & S & T \\ \hline S & p_{11} & p_{12} \\ \hline T & p_{21} & p_{22} \end{array}, \quad \Omega = \{S,T\}^2$$

$$I^1 = \{\{SS, ST\}, \{TS, TT\}\}, \quad I^2 = \{\{SS, TS\}, \{ST, TT\}\}.$$

A rule of the node 1 (row) is to assigns the first element and the second element of I^1 respectively. A rule of the node 2 (column) is to assigns the first element and the second element of I^2 respectively.

For the node 1,

- *SS means "always stay quiet: choose always S"(independently of the signal).*

- *TT means "always transmit: choose always T"(independently of the signal).*

- *ST means "obedient to recommendation: choose S if $\omega \in \{SS, ST\}$ and choose T if $\omega \in \{TS, TT\}$".*

- *TS means "opposite of the recommendation: choose S if $\omega \in \{TS, TT\}$ and choose T if $\omega \in \{SS, ST\}$".*

For the node 2,

- *SS means "always stay quiet: choose always S"(independently of the signal).*

- *TT means "always transmit: choose always T"(independently of the signal).*

- *ST means "obedient to recommendation: choose S if $\omega \in \{SS, TS\}$ and choose T if $\omega \in \{ST, TT\}$".*

- *TS means "opposite of the recommendation: choose S if $\omega \in \{ST, TT\}$ and choose T if $\omega \in \{SS, TS\}$".*

Thus, $\mathcal{AF} = \{SS, ST, TS, TT\}$ and the extended game becomes

	SS	ST	TS	TT
SS	-1	$-p_{11} - p_{21}$	$-p_{12} - p_{22}$	0
ST	$-p_{11} - p_{12} + 3p_{22} + 3p_{21}$	$-p_{11} + 3p_{21} - 2p_{22}$	$-p_{12} + 3p_{22} - 2p_{21}$	$-2p_{21} - 2p_{22}$
TS	$-p_{21} - p_{22} + 3p_{11} + 3p_{12}$	$-p_{21} + 3p_{11} - 2p_{21}$	$-p_{22} + 3p_{12} - 2p_{11}$	$-2p_{12} - 2p_{11}$
TT	3	$3p_{11} + 3p_{21} - 2p_{12} - 2p_{22}$	$3p_{12} + 3p_{22} - 2p_{11} - 2p_{21}$	-2

Remarks 3.4.9.5. *The mixed strategy $s = (1/3, 2/3)$ is the unique ESS of the game with payoff matrix Γ_1. The expected payoff obtained at the ESS is $-1/3$. In this example, every probability distribution $\mathbb{P}^1$ satisfying $p_{11} < \dfrac{p_{12}}{2}$, $p_{22} < 2p_{21}$, $p_{12} = p_{21}$ is a CESS. In particular,*

$$\mathbb{P}^1 : \begin{array}{c|cc} & S & T \\ \hline S & 0 & 1/2 \\ T & 1/2 & 0 \end{array}$$

is a CESS (but the mixed strategy $(1/2, 1/2)$ is not an ESS) and the assignment functions profile (ST, ST) and (TS, TS) give the payoff $3/2$ which is greater than $-1/3$.

Number of nodes which transmit: active nodes We define $numb_T : \mathcal{A}^n \longrightarrow \mathbb{N}$ as the number of transmitters of an action. Par example, $numb_T(T, \ldots, T) = n$ and $numb_T(T, S, \ldots, S) = 1$. The set $\{b \in \mathcal{A} \mid numb_T(b) = 1\}$ is exactly the situations where only one of the nodes *transmit* (and the others *stay quiet*). $numb_T(b) = 1$ if and only if there exists a unique j such that $b^j = T$ i.e all permutations of the actions profile $T \underbrace{S \ldots S}_{n-1 \text{ times}}$.

Denote by $\Theta^n := \{b \in \mathcal{A}^n \mid numb_T(b) = 1\}$.

Proposition 3.4.9.6. *The probability distribution $\mathbb{P}^n$ over $\mathcal{A}^n$ defined as*

$$\mathbb{P}^n(a) = \begin{cases} \dfrac{1}{n} & \text{if } a \in \Theta^n \\ 0 & \text{otherwise} \end{cases}$$

is a CESS with the payoff $\dfrac{V - \delta}{n} > 0$. Moreover at each slot, we have a successful transmission with probability one (if each node have some packet to transmit at each slot) and the allocation of payoffs obtained at this CESS is Pareto optimal.

Recall the payoff obtained at ESS is not Pareto optimal.

Proof. Θ^n has exactly n actions. Thus, $\mathbb{P}^n$ is a probability measure. Let $B := \{b \in \mathcal{A}^{n-1} \mid numb_T(b) = 1\}$.

$$[u^n(TS \ldots S) - u^n(SS \ldots S)]\mathbb{P}^n(TS \ldots S) = \frac{V - \delta}{n} > 0, \text{ and}$$

$$\sum_{a^{-1} \in \mathcal{A}^{n-1}} [u^n(Sa^{-1}) - u^n(Ta^{-1})]\mathbb{P}^n(Sa^{-1}) = -\sum_{a^{-1} \in \mathcal{A}^{n-1}} u^n(Ta^{-1})\mathbb{P}^n(Sa^{-1}) = (n-1)\frac{(\Delta + \delta)}{n} > 0.$$

Thus, $\mathbb{P}^n$ is a CESS and the system has a successful transmission at each slot with probability one. The total payoff obtained at the CESS is exactly $V - \delta$. Hence the allocation at this CESS is Pareto optimal. $\square$

The Figure 3.19 represents the payoff obtained at the CESS given in the proposition 3.7.4.6 and the payoff of the ESS varying the regret cost κ. We took $\Delta = 2/10, V = 1, \delta = 0$.
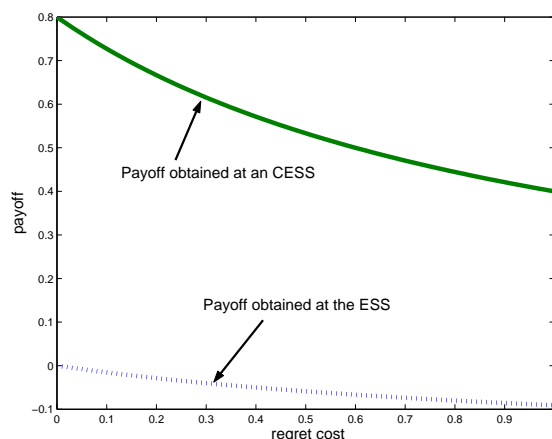
Figure 3.19: Payoffs obtained at a CESS and the ESS.

3.5 W-CDMA Wireless Networks

In this section, the random number of opponents is induced by the geographical position of the mobiles compared to the base stations.
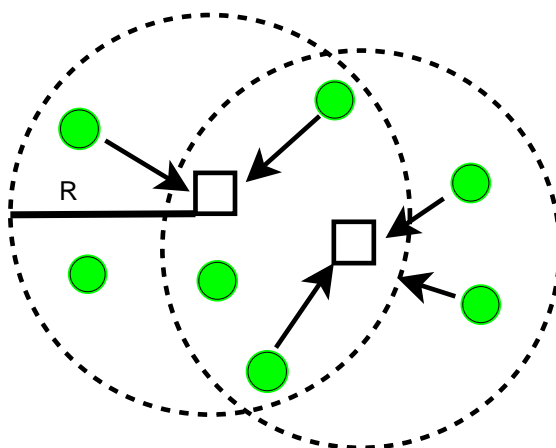
Figure 3.20: Interferences at the receiver in uplink CDMA transmissions.

We study in this section competitive decentralized power control in an wireless network where the mobiles uses, as uplink MAC protocol, the W-CDMA technique to transmit to a receiver. We assume that there is a large population of mobiles which are randomly placed over a plane following a Poisson process with density λ. We consider a random number of mobiles interacting locally. When a mobile i transmits to its receiver $R(i)$, all mobiles within a circle of radius R centered at the receiver $R(i)$ cause interference to the transmission from node i to receiver $R(i)$ as illustrated in figure 3.20. We assume that a mobile is within a circle of a receiver with probability μ. We define a random variable $\mathcal{R}$ which will be used to represent the distance between a mobile and a receiver. Let $\varsigma(r)$ be the probability density function (pdf) for $\mathcal{R}$. Then we have $\mu = \int_0^R \varsigma(r) dr$.

Remarks 3.5.0.7. *If we assume that the receivers or access points are randomly distributed following a poisson process with density v, the probability density function is expressed by $\varsigma(r) = ve^{-vr}$.*

For uplink transmissions, a mobile has to choose between High(H) power level and Low(L) power level. We denote by P_H the high power and P_L the low power levels. Let s be the population share strategy H. Hence, the signal P_r received at the receiver is given by $P_r = gP_i l(r)$, where g is the gain antenna and $\alpha > 2$ is the path loss exponent. For the attenuation, the most common function is $l(t) = \dfrac{1}{t^\alpha}$, with α ranging from 3 to 6. Note that such $l(t)$ explodes at $t = 0$, and thus in particular is not correct for a small distance r and largen intensity λ. Then, it makes sense to assume attenuation to be a bounded function in the vicinity of the antenna. Hence the last function becomes $l(t) = \max(t, r_0)^{-\alpha}$. First we note that the number of transmission within a circle of radius r_0 centered at the receiver is $\lambda \pi r_0^2$. Then the interference caused by all mobiles in that circle is $I_0(s) = \dfrac{\lambda \pi g(sP_H + (1-s)P_L)}{r_0^{\alpha-2}}$.

Now we consider a thin ring $\mathcal{A}_j$ with the inner radius $r_j = jdr$ and the outer radius $r_j = r_0 + jdr$. The signal power received at the receiver from any node in $\mathcal{A}_j$ is $P_{r_i} = \dfrac{gP_i}{r_i^\alpha}$. Hence the interference caused by all mobiles in $\mathcal{A}_j$ is given by

$$I_j(s) = \begin{cases} 2g\lambda \pi r_j dr \left(\dfrac{sP_H + (1-s)P_L}{r_j^\alpha} \right) & \text{if } r_j < R, \\[3mm] 2\mu g\lambda \pi r_j dr \left(\dfrac{sP_H + (1-s)P_L}{r_j^\alpha} \right) & \text{if } r_j \geq R. \end{cases}$$

Hence, the total interference contributed by all nodes at the receiver is

$$\begin{aligned} I(s) &= I_0(s) + 2g\lambda \pi (sP_H + (1-s)P_L) \left[\int_{r_0}^R \frac{1}{r^{\alpha-1}} dr + \mu \int_R^\infty \frac{1}{r^{\alpha-1}} dr \right], \\ &= g\lambda \pi (sP_H + (1-s)P_L)\left(\frac{\alpha}{\alpha-2} r_0^{-(\alpha-2)} - 2(1-\mu)R^{-(\alpha-2)} \right). \end{aligned}$$

Hence the signal to interference ratio $SINR_i$ is given by

$$SINR_i(P_i, s, r) = \begin{cases} \dfrac{gP_i/r_0^\alpha}{\sigma + \beta I(s)} & \text{if } r \leq r_0, \\[3mm] \dfrac{gP_i/r^\alpha}{\sigma + \beta I(s)} & \text{if } r \geq r_0, \end{cases}$$

where σ is the power of the thermal background noise and β is the inverse of the processing gain of the system. This parameter weights the effect of interference, depending on the orthogonality between codes used during simultaneous transmissions. In the sequel, we compute the mobile's utility (fitness) depending on his decision but also on the decision of his interferers. We assume the user's utility (fitness) choosing power level P_i is expressed by

$$f_{P_i}(s) = w \int_0^R \log(1 + SINR(P_i, s, r))\varsigma(r)dr - \eta P_i.$$

The pricing function P_i define the instantaneous "price" a mobile pays for using a specific amount of power that causes interference in the system and η is a parameter. This price can be the power cost consumption for sending packets.

We are now looking at the existence and uniqueness of the ESS. For this, we need the following result.

Lemma 3.5.0.8. *For all density function ς defined on $[0, R]$, the function $h : [0, 1] \rightarrow \mathbb{R}$ defined as*

$$s \longmapsto \int_0^R \log\left(\frac{1 + SINR(P_H, s, r)}{1 + SINR(P_L, s, r)}\right) \varsigma(r) \, dr \text{ is continuous and strictly monotone.}$$

Proof. The function $s \longmapsto \log\left(\frac{1 + SINR(P_H, s, r)}{1 + SINR(P_L, s, r)}\right) \varsigma(r)$ is continuous and integrable in r on the interval $[0, R]$. The function h is continuous. Using derivative properties of integral with parameter, we can see that the derivative function of h is the function $h' : [0, 1] \rightarrow \mathbb{R}$ defined as

$$s \longmapsto \int_0^R \frac{\partial}{\partial s}\left[\log\left(\frac{1 + SINR(P_H, s, r)}{1 + SINR(P_L, s, r)}\right)\right] \varsigma(r).$$

We show that the term $\frac{\partial}{\partial s}\left[\log\left(\frac{1 + SINR(P_H, s, r)}{1 + SINR(P_L, s, r)}\right)\right]$ is negative. Let $W(s) := \frac{1 + SINR(P_H, s, r)}{1 + SINR(P_L, s, r)}$.

The function W can be rewritten as $W(s) = 1 + \frac{\frac{g(P_H - P_L)}{r^\alpha}}{\sigma + \beta I(s) + \frac{gP_L}{r^\alpha}}$ where $I(s) = (s(P_H - P_L) + $

$P_L)c(r)$ and $c(r) = \lambda \pi g \left[\frac{\alpha}{\alpha - 2} r_0^{-(\alpha - 2)} - 2(1 - \mu)R^{-(\alpha - 2)}\right]$ if $r \geq r_0$ and $\frac{\lambda \pi g}{r_0^{\alpha - 2}}$ otherwise. Since

W satisfies $W(s) > 1$ and $W'(s) = -c(r)\beta(P_H - P_L)\frac{\frac{g(P_H - P_L)}{r^\alpha}}{(\sigma + \beta I(s) + \frac{gP_L}{r^\alpha})^2} < 0$. Hence,

$$\frac{\partial}{\partial s}\left[\log\left(\frac{1 + SINR(P_H, s, r)}{1 + SINR(P_L, s, r)}\right)\right] = \frac{\partial}{\partial s}(\log W(s)) = \frac{W'(s)}{W(s)} < 0$$

i.e $h'(s) < 0$. We conclude that h is strictly decreasing. $\qquad\square$

Using this lemma, we have the following proposition which gives pure strategies depending on the parameters.

Proposition 3.5.0.9. *For all density function ς, the pure strategy P_H dominates the strategy P_L if and only if $\frac{\eta}{w}(P_H - P_L) < \int_0^R \log\left(\frac{1 + SINR(P_H, P_H, r)}{1 + SINR(P_L, P_H, r)}\right) \varsigma(r) \, dr = h(1)$. For all density function ς, the pure strategy P_L dominates the strategy P_H if and only if $\frac{\eta}{w}(P_H - P_L) > \int_0^R \log\left(\frac{1 + SINR(P_H, P_L, r)}{1 + SINR(P_L, P_L, r)}\right) \varsigma(r) \, dr = h(0)$.*

Proof. We decompose the existence of the ESS in several cases.

1. P_H is preferred to P_L : The higher power level dominates the lower if and only if $f_{P_H}(P_H) > f_{P_L}(P_H)$ and $f_{P_H}(P_L) > f_{P_L}(P_L)$. These two inequalities implies that

$$\frac{\eta}{w}(P_H - P_L) < \int_0^R \log\left(\frac{1 + SINR(P_H, P_H, r)}{1 + SINR(P_L, P_H, r)}\right) \varsigma(r) \, dr.$$

2. P_L is preferred to P_H : Analogously, the lower power dominates the higher power if and only if $f_{P_L}(P_H) > f_{P_H}(P_H)$ and $f_{P_L}(P_L) > f_{P_H}(P_L)$ i.e $\frac{\eta}{w}(P_H - P_L) > \int_0^R \log\left(\frac{1 + SINR(P_H, P_L, r)}{1 + SINR(P_L, P_L, r)}\right) \varsigma(r) \, dr.$

$\qquad\square$

The following result gives a sufficient condition for the existence of the ESS.

Proposition 3.5.0.10. *For all density function ς, if $h(1) < \frac{\eta}{w}(P_H - P_L) < h(0)$, then there exists an unique ESS s^* which is given by $s^* = h^{-1}\left(\frac{\eta}{w}(P_H - P_L)\right)$.*

Proof. Suppose that the parameters w, η, P_H and P_L satisfy the following inequality $h(1) < \frac{\eta}{w}(P_H - P_L) < h(0)$. Then the game has no dominant strategy. A mixed equilibrium is characterized by $f_{P_H}(s) = f_{P_L}(s)$. It is easy to see that this last equation is equivalent to $h(s) = \frac{\eta}{w}(P_H - P_L)$. From the lemma 3.5.0.8, we have that the equation $h(s) = \frac{\eta}{w}(P_H - P_L)$ has an unique solution given by $s^* = h^{-1}\left(\frac{\eta}{w}(P_H - P_L)\right)$. We now prove that this mixed equilibrium is an ESS. To prove this result, we compare $s^* f_{P_H}(mut) + (1 - s^*)f_{P_L}(mut)$ and $mut f_{P_H}(mut) + (1 - mut)f_{P_L}(mut)$ for all $mut \neq s^*$. The difference between two values is exactly $w(s^* - mut)(h(mut) - h(s^*))$. According to lemma 3.5.0.8, h is decreasing function. Hence, $(s^* - mut)(h(mut) - h(s^*))$ is strictly positive for all strategy mut different from s^*. We conclude that the mixed equilibrium $(s^*, 1 - s^*)$ is an ESS. $\qquad\square$

From the last proposition, we can use the pricing η as a design tool for create an incentive for the user to adjust their power control. We observe that the ESS s^* decreases when η increases. That means the mobiles become less aggressive when pricing function increases and the system can limit aggressive requests for SINR.

3.5.1 Numerical examples in W-CDMA Wireless Networks:

In the numerical examples below, we show how the pricing function can optimize the overall network throughput. we first investigate the impact of the different parameters and pricing on the ESS and the convergence of the replicator dynamic. We also discuss the impact of the pricing function on the system capacity.

We first show the impact of density of nodes and pricing on the ESS and the average rate. We assume that the receivers which are randomly placed over a plane following a Poisson process with density ν, i.e, $\varsigma(r) = \nu e^{-\nu r}$. We recall that the rate of a mobile using power level P_i at the equilibrium is given by $w \int_0^R \log(1 + SINR(P_i, s, r))\varsigma(r)dr$. We took $r_0 = 0.2, w = 20, \sigma = 0.2, \alpha = 3, \beta = 0.2, R = 1$. First, we show the impact of the density of nodes λ on the ESS and the average rate. In figures 3.21, we depict the average rate obtained at the equilibrium and the ESS, respectively , as a function of the density λ. We recall that the interference for a mobile increases when λ increases. We observe that the mobiles become less aggressive when the density increases. We observe that it is important to adapt the pricing as function of the density of nodes. Indeed, we observe that for low density of nodes, the lower pricing ($\eta = 0.92$) gives better results than higher pricing ($\eta = 0.97$). When the density of nodes increases, the better performance is obtained with higher pricing.

Our second experiment in W-CDMA studies convergence to the ESS of the W-CDMA system described above under replicator dynamics. Fig. 3.22 and 3.23 represent the fraction of population using the high power level for different initial states of the population: 0.99, 0.66, 0.25 and 0.03. We observe that the choice of receiver distributions change the ESS. In the next section we study another wireless architecture where interferences are between mobiles which are located in different cell. This typical interference problem occurs in WiMAX environment.
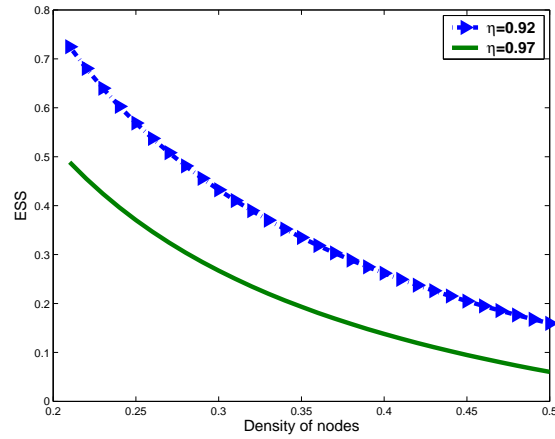
Figure 3.21: *The average rate at equilibrium versus λ for η = 0.92, 0.97.*
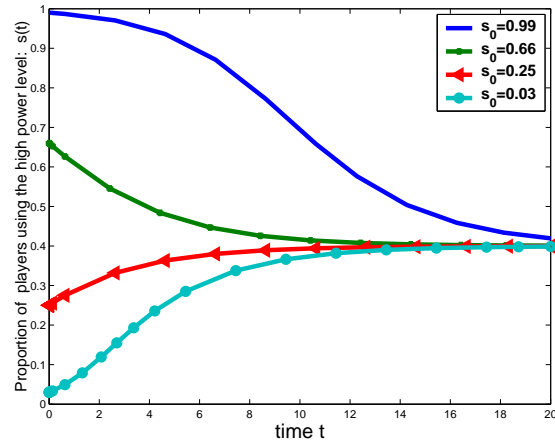
Figure 3.22: *Convergence to the ESS in W-CDMA system : uniform distribution.*
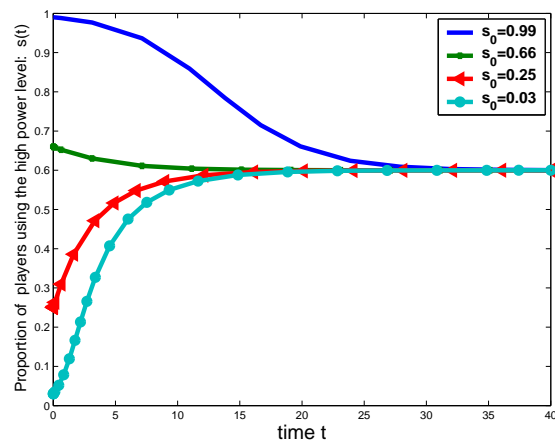
Figure 3.23: *Convergence to the ESS in W-CDMA system : quadratic distribution.*

3.6 OFDMA-based IEEE802.16 Network

OFDMA (Orthogonal Frequency Division Multiple Access) is recognized as one of the most promising multiple access technique in wireless communication system. This technique is used to improve spectral efficiency and becomes an attractive multiple access technique for 4th generation mobile communication system as WiMAX.

In OFDMA systems, each user occupies a subset of subcarriers, and each carrier is assigned exclusively to only one user at any time. This technique has the advantage of eliminating intra-cell interference (interference between subcarriers is negligible). Hence the transmission is affected by intercell interference since users in adjacent sectors may have also been assigned to the same carrier. If those users in the adjacent sectors transmitted with high power the intercell interference may severely limit the SINR achieved by the user. Some form of coordination between the different cells occupying the spectral resource are studied in ([134](#); [123](#)). The optimal resource allocation requires complete information about the network in order to decide which users in which cells should transmit simultaneously with a given power. All of these results however, rely on some form of centralized control to obtain gains at various layers of the communication stack. In a realistic network as WiMAX, centralized multicell coordination is hard to realize in practice, especially in fast-fading environments.

We consider an OFDMA system where radio resources are allocated to users on their channel measures and traffic requirements. Each carrier within a frame must be assigned to at most one user in the corresponding cell. In this way each carrier assignment can be made independently in each cell. Hence when a user is assigned to carrier, the mobile should determine the power transmission to the Base station. This power should take into account the interference experienced by the transmitted packet.

Consider the uplink of a multiple multicell system, employing the same spectral resource in each cell. Power control is used in an effort to preserve power and to limit interference and fading effects. For users located in a given cell, co-channel interference may therefore come from only few cells as illustrated in figure 3.24. Since the intra-cell interference is negligible, we focus on the users which use a specific carrier. Consider N cells, and a large number of population of
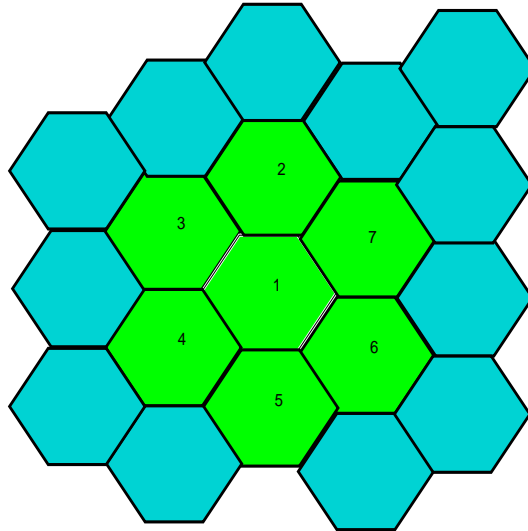
Figure 3.24: *Hexagonal cell configuration*

mobiles randomly distributed over each channel and each cell. Since in OFDMA systems, each carrier is assigned exclusively to only one mobile at any time, we assume that the interactions between mobiles are manifested through many local interactions between N' mobiles where N' is the set of neighbors of a cell. We can ignore the interaction among more than N' mobiles that transmit simultaneously and that can cause interference to each other. Hence, in each slot, interaction occurs only between the mobiles which have been assigned to the same carrier.

Let g_{ij} denote the average channel gain from user i to cell j. Hence, if a user in cell i transmits with power p_i, the received signal strength at cell i is $p_i g_{ii}$, while the interference it occurs on cell j is $p_i g_{ij}$. Hence, the interference experienced by cell i is given by $SINR_i(\mathbf{p}) = \dfrac{g_{ii}p_i}{\sigma_i + \sum_{j \neq i} g_{ij}p_j}$, where σ_i is the power of the thermal noise experienced at cell i. The rate achieved by user i is given by $r_i(\mathbf{p}) = \log(1 + SINR_i(\mathbf{p}))$, where $\mathbf{p}$ denotes the power level vector of mobiles choice which are assigned to a specific carrier. We assume that the user's utility is given by $u_i(\mathbf{p}) = r_i(\mathbf{p}) - \eta p_i$. The above utility represents the weighted difference between the throughput that can be achieved as expressed by Shannon's capacity and the power consumption cost. We assume that all mobiles perform the on-off power allocation strategy. In this strategy, each mobile transmits with full power or remains silent. Let N_i be the set of neighbors of a user in cell i. Hence the interference experienced by a user in the cell i is given by $SINR_i(\mathbf{p}) = \dfrac{g_{ii}p_i}{\sigma_i + \sum_{j \in N_i \setminus \{i\}} g_{ji}p_j}$, where $p_j \in \{0, P\}$, P is the power level of transmission.

Let s_i the proportion of transmitters in the cell i. The couple $(s_i, 1 - s_i)$ with $s_i \in (0, 1)$, represents the state of the cell i. We denote by s_{-i} the vector $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_{N_i})$. The fitness of the cell i can be defined as follows:

$$f^i(s_i, s_{-i}) = s_i f_P^i(s_{-i}) = s_i \sum_{a_{-i} \in \{0, P\}^{|N_i| - 1}} u_i(P, a_{-i}) s_{-i}(a_{-i}) \text{ where}$$

$$s_{-i}(a_{-i}) = \left(\prod_{j \in T(a_{-i})} s_j \right) \left(\prod_{j \in N_i \setminus \{T(a_{-i}) \cup \{i\}\}} (1 - s_j) \right) \text{ and}$$

$$T(a_{-i}) = \{k \in N_i \setminus \{i\}, \ p_k = P\},$$

is the set of neighbors transmitting. We have $f_0^i(s_{-i}) = 0$ for any multi-strategy s_{-i}.

A multi-strategy $s = (s_i)_{i=1,\ldots N}$ is neutrally stable if for all $s' \neq s$ there exists $\epsilon_{y'} > 0$ such that $\forall \epsilon \in (0, \epsilon_{y'})$

$$f_{s_i}^i(\epsilon s'_{-i} + (1 - \epsilon)s_{-i}) \geq f_{s'_i}^i(\epsilon s'_{-i} + (1 - \epsilon)s_{-i}) \tag{3.6}$$

for some i. The multi-strategy s is evolutionary stable if the inequality (3.6) is strict. Hence, an ESS is a neutrally stable strategy but the reciprocal is not true. See (169) for more details on neutrally stable strategy. If s is neutrally stable strategy then s is Nash equilibrium. The best response (BR) of the user i to s_{-i} is given by

$$BR_i(s_{-i}) = \begin{cases} 1 & \text{if } f_P^i(s_{-i}) > 0, \\ 0 & \text{if } f_P^i(s_{-i}) < 0, \\ [0, 1] & \text{if } f_P^i(s_{-i}) = 0. \end{cases}$$

A strategy s is a Nash equilibrium is equivalent to $s_i \in BR_i(s_{-i})$, $i = 1, \ldots, |N|$. In particular when $\eta = 0$, the strategy ON is a Nash equilibrium (ON is a dominant strategy).

The following proposition gives necessary conditions for pure strategies and strictly mixed strategies to be neutrally stable.

Proposition 3.6.0.1. *We decompose the different equilibrium strategy in the following items.*

- *If the strategy ON (P) is a neutrally stable equilibrium then*

$$\eta \leq \eta_{min} \text{ where } \eta_{min} = \frac{1}{P} \min_{i=1,\ldots,N} \log\left(1 + \frac{g_{ii}}{\frac{\sigma_i}{P} + \sum_{j\in N_i, \, j\neq i} g_{ji}}\right).$$

  Moreover, ON becomes a strictly dominant strategy if $\eta < \eta_{min}$.

- *If the strategy OFF (0) is a neutrally stable equilibrium then*

$$\eta \geq \frac{1}{P} \max_i \log\left(1 + \frac{g_{ii}P}{\sigma_i}\right) =: \eta_{max}.$$

  For $\eta > \eta_{max}$, OFF becomes a strictly dominant strategy (hence, an ESS).

- *If $\forall \, j \neq i$, $g_{ij} = g$, $g_{ii} = \bar{g}$, $\sigma_i = \sigma$ then the game becomes a symmetric game. Hence, it exists a symmetric equilibrium with the proportion $s (= s_i, \, \forall i)$ of transmitters, which must satisfy*

$$Q(s) := \sum_{k=0}^{n-1} a_k s^k (1-s)^{n-k-1} \binom{n-1}{k} = \eta P \tag{3.7}$$

  where $a_k = \log(1 + \frac{\bar{g}P}{\sigma + kgP}) > 0$ represents the fitness obtained by user i when he transmits and k of the opponents of user i decided to transmit and the $n - 1 - k$ others stay quiet.

- *Every strictly mixed equilibrium (symmetric or not) must satisfy $f^i_{s_i}(s_{-i}) = 0$ for all i.*

Proof. The two first assertions of the proposition 3.6.0.1 are obtained by Nash equilibrium conditions: the strategy ON (P) is a neutrally stable equilibrium implies that $\forall i$, $0 \leq f^i_P(P, \ldots, P)$ and the strategy OFF (0) is a neutrally stable equilibrium implies that $\forall i$, $f^i_P(0, \ldots, 0) \leq 0$. In the last assertion (symmetric case), the fitness of user i is $s(Q(s) - \eta P)$ where $(s_i, 1 - s_i)$ is state of the cell. Since $Q(1) < 0$ and $Q(0) > 0$ when $\eta_{min} < \eta < \eta_{max}$, then the equation (3.7) has a solution s in $(0, 1)$ which implies the existence of an ESS. $\square$

Remark that the strategy ON is an ESS if $\eta < \eta_{min}$. In order to prove the existence and uniqueness of the ESS in the symmetric case, we study roots of the polynomial Q in the following lemma.

Lemma 3.6.0.2. *Let $0 \leq a_{n-1} < a_{n-2} < \ldots < a_0$, η and P positive reals satisfying $\eta P \in (a_{n-1}, a_0)$. Then, the polynomial $Q(s) - \eta P$ has a unique root on $(0, 1)$.*

Proof. We show existence and uniqueness of the root of the polynomial Q on $(0, 1)$.

Existence Q is a polynomial with real coefficients. Hence, Q is continuous in $(0, 1)$. The image of the interval $(0, 1)$ contains $I = [a_{n-1} - \eta P, a_0 - \eta P]$. The inequality $\frac{a_{n-1}}{P} := \eta_{min} < \eta < \frac{a_0}{P} := \eta_{max}$, implies $0 \in I$. Thus, there exists, $a \in (0, 1)$ such that $Q(a) = 0$.

Uniqueness We show that Q is strictly monotone on $(0, 1)$. Let Q' be the derivative of Q. Q' is given by

$$Q'(s) = \sum_{k\geq 0}(k+1)\binom{n-1}{k+1}a_{k+1}s^k(1-s)^{n-2-k} - \sum_{k\geq 0}(n-k-1)\binom{n-1}{k}a_k s^k(1-s)^{n-2-k}$$

Since $(k+1)\binom{n-1}{k+1} = (n-k-1)\binom{n-1}{k}$, one has

$$Q'(s) \quad = -\sum_{k=0}^{n-2} \frac{(n-1)!}{k!(n-2-k)!}[a_k - a_{k+1}]s^k(1-s)^{n-2-k}$$

The inequalities $a_k - a_{k+1} > 0$, $k = 0, \ldots, n-2$ and $s \geq 0, 1-s \geq 0$ imply that $Q'(s) < 0$ on $(0,1)$. Hence, Q is strictly decreasing on $(0,1)$. We conclude that Q is bijective from $(0,1)$ to I and hence, Q has a unique root on $(0,1)$ □

Given this result, we have the following proposition given the existence and uniqueness of the ESS in the symmetric game. The proof is immediate from the lemma 3.6.0.2 and from proposition 3.6.0.1.

Proposition 3.6.0.3. *The symmetric power control game has a unique strictly mixed equilibrium.*

3.6.1 Numerical investigation in OFDMA-based IEEE802.16 networks

We consider below numerical examples of an OFDMA system. We shall obtain the ESS for several values of η. We assume that $g_{ii} = g$ and $g_{ij} = \bar{g}$ for all $i \neq j$ and $\sigma_i = \sigma$ for all i. We consider below a numerical example for different values of N' (see figure 3.24 in which $N' = 7$). Let the noise $\sigma = 0.1$ and the power attenuation $g = \bar{g}/4 = 0.9$. In figure 3.25 (resp. 3.26), we plot the ESS versus η (resp. power level P). In both figures, the population ratio using the strategy ON is monotone decreasing in η.
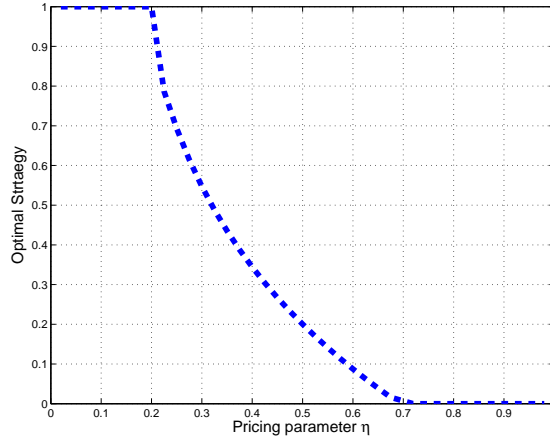
Figure 3.25: The population ratio of ON at equilibrium versus η.

We observe that the parameter η which can be interpreted as the pricing per unit of power transmitted, can determine whether the ESS is aggressive (in pure strategies or in mixed strategies). It can determine in the latter case what fraction of the population will use high power at equilibrium. Pricing can thus be used to control interference.

Notes

In this chapter we have adapted the theory of evolutionary games with a random number of players in wireless networks. This adaptation is needed in order to apply this theory for the
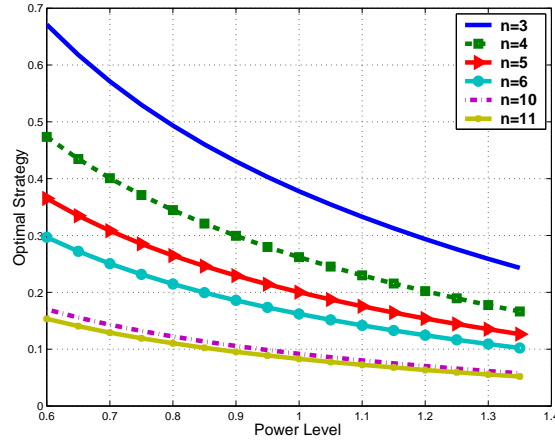
Figure 3.26: *The population ratio of strategy ON versus P.*

study of access game and particularly in wireless networks. (i) We have proposed different scenario based on the level of information for each player in the slotted Aloha model. In all cases, we have obtained the existence and uniqueness of the ESS, we have proposed optimization issues for the transmission probability of success and finally, we have studied the impact of delay in the convergence to the ESS of the replicator dynamics. (ii) We considered power control games with one or several populations of users and studied evolutionarily stable states in interaction of numerous mobiles in competition in a wireless environment. We have modeled power control for W-CDMA and OFDMA systems as a non-cooperative population game in which each user needs to decide how much power to transmit over each receiver and many local interactions between users at the same time has been considered. We have derived the conditions for existence and uniqueness of evolutionarily stable states and characterize the distribution of the users over each power level.

3.7 Correlated Evolutionarily Stable Strategies in Random Medium Access

In this section we study dynamic multiple access in distributed wireless networks with random number of users. We apply evolutionary game theoretic analysis to solve several problems: (a) We address the stability of Aloha-like systems with power levels. Specifically, we consider an arbitrary number of receivers distributed in several locations. They receive packets from random number of users accessing the resource using Aloha-like algorithms. We provide an explicit expression of equilibria, correlated evolutionarily stable strategies, and prove some asymptotic stability results. (b) We apply correlation mechanism and evaluate the performance of random medium access when saturated users interact through interference. We introduce the benefit of correlation (BoC) to measure the gap between the probability of success at correlated evolutionarily stable states and the worst probability of success of evolutionarily stable states. We show that if only two power levels are available, the correlation mechanism reduces considerably the interference and the number of collisions. Moreover, the correlation mechanism is stable in long-term under several classes of bio-inspired evolutionary game dynamics. (c) Surprisingly, when the number of strategies is at least three, the game has no pure equilibrium, and the correlation mechanisms do not improve the probability of success and BoC=0.

Random Medium Access Control (MAC) algorithms have played an increasingly important role in the development of wired and wireless networks and the performance and stability of these algorithms, such as slotted-Aloha, Carrier Sense Multiple Access (CSMA) is still an open problem (36). Distributed Medium Access Control, starting from the first version of Abramson's Aloha to the most recent algorithms used in IEEE802.11, have enabled a rapid growth of both wired and wireless networks. They aim at efficiently and fairly sharing a resource among users even though each user must decide independently (eventually after receiving some messages or listening) when and how to attempt to use the resource. MAC algorithms have generated a lot of research interest, especially recently in attempts to use multi-hop wireless networks to provide high-speed access to the internet with low-cost and low-energy consumption.

In this section we focus our attention to wireless networks, where the resources are receivers, base station or access points and where users interact because of interference, i.e., interfering users cannot transmit simultaneously. There is a collision if another user (mobile) transmits with a greater power level at the same range of the receiver. Motivated by the interest of evolving dense networks, evolutionary game theory was found to be an appropriate framework to apply in networks. It has been applied to problems such as congestion control (147), distributed cooperative sensing over cognitive radio networks (232; 160), code division multiple access (CDMA), Orthogonal Frequency-Division Multiple Access (OFDMA) based Worldwide Interoperability for Microwave Access (WiMAX) environment (11), reciprocal and non-reciprocal interference control, mobile medium access control and channel selection (213) and capacity region of Additive White Gaussian Noise (AWGN) (88), Multihoming and association problems (184). A one-shot random access game have been studied in (109). In (232; 88; 11; 109; 213) the authors did not analyze the performance and the fairness properties of these approaches and the number of strategies limited is fixed to two.

We provide a general evolutionary game theoretic framework to analyze networks where interfering users share a resource using an Aloha-type access control. We consider access control with *arbitrary number* of strategies and *variable number of mobiles* around each receiver. We study a large population of communicating terminals using a aloha-like protocol with several levels of transmission power. We examine how to choose between these power levels in order maximize their probability of success minus the cost of energy consumption. We study several solution and stability concepts: the Nash equilibrium and the Evolutionarily Stable State (ESS) and the correlated evolutionarily stable strategy. The concept of ESS were introduced in mathematical biology by Maynard Smith and Price (187; 186) in the context of Evolutionary Games, which allow to describe and to predict properties of large populations whose evolution depends on many local interactions, each involving a finite number of individuals. Evolutionary game dynamics are models of strategy change commonly used in evolutionary game theory. A strategy which does better than the average or its opponent, increases in frequency at the expense of strategies that do worse than the average or the opposed action. Many evolutionary game dynamics models are used in the literature (229; 168; 99; 65; 234). We compare the performances of these notions with the global cooperative solution. The payoffs that we consider are functions of the probability to have a successful transmission and of the cost for the power levels. We study in particular the impact of the pricing for the use of the power levels on the system performance. We analyze various solutions concepts and refinement: Nash equilibria, Pareto optimality, strong equilibria, correlated evolutionarily stable state. In order to study the interactions between mobiles and their stationary strategy in the long run, we develop some evolutionary game dynamics (99; 168; 229; 90) in order to study the convergence and the stability of equilibria.

Related works

In (232), an evolutionary game-theoretical framework with two strategies: *Cooperate or Defect* has been proposed for distributed cooperative sensing over cognitive radio networks. By employing the theory of replicator dynamics, the authors study the behavior dynamics of secondary users, and further propose a distributed learning algorithm that gradually converges to the Nash equilibrium. In (20; 212), an evolutionary access game with battery state-dependent energy management have been proposed for distributed Aloha networks with low density. The authors consider only two strategies: *high or low power* and interference between more than three mobiles is neglected. Our analysis extend (20; 213; 109) and covers general distribution (eventually with mobility) of mobiles and arbitrary number of strategies. We propose an evolutionary access control game with random number of interacting players and arbitrary number of actions. Our work extends previous evolutionary networking game theoretic works (232; 20; 11) where only two strategies is considered. We use the Price of Anarchy (PoA) concept (28; 29) in order to measure the gap between equilibrium and social optima in the population game. We show that the PoA of evolutionarily stable state is more than the PoA of Nash equilibria in any population game with arbitrary number of strategies. In the evolutionary access game with non-degenerate costs and more than three strategies, the PoA at ESS denoted by PoA_{ESS} coincide with PoA_{NE}. Moreover, $PoS_{ESS} = PoS_{NE} = 0$. In contrast, if we have only two strategies and m mobiles around the receiver, the $PoS_{NE} = 1 \geq PoA_{NE} = 0 = PoA_{ESS}$. We introduce and analyze correlated evolutionarily stable states (CESS) in access control. CESS is a refinement of correlated equilibrium. We show that the evolutionary access control game with arbitrary number of opponent around each receiver has a several CESS which are more robust in term of fairness, stability and efficiency than correlated equilibria (CE) (32). We define the *benefit of correlation* (BoC) that measures the improvement in term of probability of success by introducing correlation between mobiles in the population game. To the best to our knowledge, this work is one the first which study correlated evolutionarily stable states in networking context. For the two strategies case, we show that the BoC of the access game is 100% i.e $PoS_{CESS} = 1$ and $PoA_{NE} = 0$. As a consequence, for the model studied in (109) the global optimum is attained using CESS. We compute explicitly an equivalence class of CESS which is a social optimum and which is local asymptotically stable under the evolutionary dynamics. In particular, our contribution gives some situations where the three properties: equilibrium, stability, optimality (global cooperative solution) holds. For more than two strategies, we show that the game has no pure equilibrium (Theorem 3.7.3) and a unique mixed equilibrium which we compute explicitly (Theorem 3.7.6). We develop a general class of evolutionary game dynamics for CESS. Surprisingly, when the number of strategies is at least three, the correlation mechanism does not improve the probability of success.

In next subsection we formulate problem of access control with several actions and introduce the evolutionarily equilibrium concept, the efficiency metric "Price of Anarchy" and evolutionary game dynamics with time delays. In subsection 3.7.1 we analyze equilibria: Nash, ESS, CE, CESS and price of anarchy, price of stability and benefit of correlation in Dirac distribution. We compute explicitly the equilibria and the price of anarchy in general distribution in subsection 3.7.5. We extend our study to asymmetric random access games with several classes and random number of interacting mobiles and formulate the SINR-threshold based admission control and in subsection 5.5.

3.7.1 Access control with several power levels

We consider a wireless communication network with distributed receivers in which some mobiles contend for access on a common, wireless communication channel. We characterize this

distributed multiple access problem in terms of many random access game at each time. Random multiple access games introduce the problem of medium access. We assume that mobiles are randomly placed over an area and a distributed receivers in the corresponding area . The channels are ideal for transmission and all errors are due to collision. A mobile decides to transmit a packet with some power level or not to transmit (null power level) to a receiver when they are within transmission range of each other. Interference occurs as in the Aloha protocol where the power control is introduced: if more than one neighbors of the receiver transmit a packet with a power level which is greater than the corresponding power of the mobile at the same time slot there is a collision. The evolutionary random multiple access game is a nonzero-sum dynamic game, at each time, the mobiles in the same range have to share a common resource, the wireless medium. We denote by μ the probability that a mobile has its receiver within its range. We assume that $\mu > 0$ (if $\mu = 0$ there are no receivers, and hence no successful transmission).

Strategies

We assume that for each packet, its source can choose the transmitted power among several power levels

$$\mathcal{P} = \{p_0, p_1, p_2, \ldots, p_n\}$$

with $p_0 < p_1 < \ldots < p_n$ i.e a strategy for a mobile corresponds to the choice of a power level in $\mathcal{P}$. $p_0 = 0$ means that the mobile does not transmit, p_n is the maximum power level available to the mobiles.

Aloha-type payoff with pricing

If mobile j transmits a packet using a power p_j, it incurs a transmission cost of $c(p_j) \geq 0$. The packet transmission is successful if the other users in the range of its receiver use some power levels strictly lower than p_j in that given time slot, otherwise there is a collision. If there is no collision, user i gets a reward of V from the successful packet transmission. If $c(a) > V$ for some power a then a is dominated by 0 (not transmit). For the remainder, suppose that the reward V is greater than the cost of transmission $\max_j c(p_j) < V$. All packets of a lower power level involved in a collision are assumed to be lost and will have to be retransmitted later. In addition, if more than one packet of a higher power level is involved in a collision then all packets are lost. The power differentiation thus allows one packet of a higher power level to be successfully transmitted in collisions that do not involve other packets of the higher power level. Then, a transmission of mobile j is successful if its transmission power is strictly greater than the power levels used by the others mobiles at the same slot. When the number of mobiles which transmitting at the receiver is $m + 1$, the payoff is given by $u_{m+1}^j : \mathcal{P}^{m+1} \longrightarrow \mathbb{R}$

$$u_{m+1}^j(a^j, a^{-j}) = -c(a^j) + V \times \begin{cases} 1 & \text{if } a^j > \max_{k \neq j} a^k \\ 0 & \text{otherwise} \end{cases}$$

where a^{-j} denotes $(a^1, \ldots, a^{j-1}, a^{j+1}, \ldots, a^{m+1})$, $c : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a pricing function. We assume that

$$c(p_0) = c(0) = 0 \leq c(p_1) \leq c(p_2) \ldots \leq c(p_n)$$

Denote a population profile as the frequencies (fractions) of use of pure strategies in the population. A population profile can be represented by an element of the $n-$simplex Δ_n of the

$(n+1)-$dimensional Euclidean space $\mathbb{R}^{n+1}$. The expected payoff of a mobile j with the power level p_i when facing to m others mobiles is given by

$$
\begin{aligned}
f_{m,p_i}(x) &= \mu u_{m+1}^j(x,\ldots,p_i,x,\ldots,x) & (3.8)\\
&= \mu\left(\sum_{a\in\mathcal{P}^m} u_{m+1}^1(p_i,a^1,\ldots,a^m)\prod_{j=1}^m x_{a^j}\right) & (3.9)
\end{aligned}
$$

where x_{p_j} is the fraction of mobiles with the power level $p_j \in \mathcal{P}$, and $x = (x_{p_0}, x_{p_1}, \ldots, x_{p_n})$ is the population profile.

Lemma 3.7.1.1. $f_{m,p_0}(x) = 0$, and for $i \geq 1$,

$$
f_{m,p_i}(x) = \mu\left[-c(p_i) + V\sum_{j_0+j_1+\ldots+j_{i-1}=m} x_{p_0}^{j_0} x_{p_1}^{j_1}\ldots x_{p_{i-1}}^{j_{i-1}}\right]
$$

$$
f_{m,p_i}(x) = -\mu c(p_i) + V\mu(x_{p_0} + x_{p_1} + x_{p_2} + \ldots + x_{p_{i-1}})^m
$$

Denote by $\mathcal{K}$ the random variable representing the number of opponents of a anonymous mobile picked in a the population when the population profile is x and by $G_{\mathcal{K}}(s) = \mathbb{E}_{\mathcal{K}}(s^{\mathcal{K}})$ the generating function of M. The expected payoff of a mobile using the power level p_i can be expressed as

$$
\begin{aligned}
F_{p_i}(x) &= \mathbb{E}_{\mathcal{K}}(f_{\mathcal{K},p_i}(x)) \\
&= V\mu G_{\mathcal{K}}(x_{p_0} + x_{p_1} + x_{p_2} + \ldots + x_{p_{i-1}}) - \mu c(p_i)
\end{aligned}
$$

for all $i > 0$ and $F_{p_0}(x) = 0$.

We observe that if x is stochastically dominated by y, i.e for all $i < n$, $\sum_{j=i+1}^n x_j \leq \sum_{j=i+1}^n y_j$ then $F_b(x) \geq F_b(y)$, $\forall b \in \mathcal{P}$. The function F is extended to a more general generating function (229) G defined in a linear space that contains $\Delta_n \times \Delta_n$. The G-function satisfies

$$
G(v,x)_{|v=p_j} = F_{p_j}(x).
$$

A population profile x is an evolutionarily stable state (ESS) if for all other population profile $mut \neq x$ there exists a $\epsilon_{mut} > 0$ (which may depend on mut) such that,

$$
\sum_{b\in\mathcal{P}} (x_b - mut_b)G(v,(1-\epsilon)x + \epsilon\, mut)_{|v=b} > 0,\ \forall\epsilon \in (0,\epsilon_{mut}) \tag{3.10}
$$

When the inequality (3.10) is non-strict, and $\epsilon = 0$, we obtain that the probability distribution x is a symmetric Nash equilibrium (NE).

Bio-inspired and G-function game dynamics

We adopt a class of evolutionary game dynamics based on the generating fitness (payoff) function (G-function) and revision protocols developed respectively by Vincent (229) and Sandholm (168) in which we introduce time delays. The time delays characterize the expected delay needed to know if a transmission is successful or not. An action p_i taken at time t will have its effect at time $t + \tau_{p_i}$. Delayed evolutionary game dynamics with asymmetric time delays have

been introduced in (214). In the evolving MAC context, we construct evolutionary game dynamics from a model of individual decision making, we introduce revision of protocols, which describe how mobiles adjust their choices of strategies during the game. A revision protocol is a Lipschitz continuous map $\beta : \mathbb{R}^{n+1} \times (\Delta_n)^{n+2} \longrightarrow \mathbb{R}_+^{(n+1)\times(n+1)}$ that takes a G-function G vectors $G = (G(b,.))_{b\in\mathcal{P}}$ and population profile x as arguments, and returns nonnegative matrices with size $(n+1) \times (n+1)$ as outputs. The generating The scalar $\beta_{p_i p_j}(G; x^1, \ldots, x^{n+2})$ is called the conditional switch rate from strategy p_i to strategy p_j. If the agents receiving revision opportunities independently according to rate κ Poisson processes, then $\dfrac{\beta_{p_i p_j}(G, x^1, \ldots, x^{n+2})}{\kappa}$ represents the probability that a mobile with the power level p_i who receives a revision opportunity switches to transmit with the power level p_j. The two maps β and G together define a delayed non-linear differential equation

$$\frac{d}{dt}x_{p_i}(t) = \sum_{b\in\mathcal{P}} x_b(t)\beta_{b p_i}(G, x(t), x(t-\tau_{p_0}), \ldots, x(t-\tau_{p_n})) \tag{3.11}$$
$$-x_{p_i}(t)\sum_{b\in\mathcal{P}}\beta_{p_i b}(G, x(t), x(t-\tau_{p_0}), \ldots, x(t-\tau_{p_n}))$$

with initial condition

$$x(t) = \phi(t), \ \forall t \in (-\max_{b\in\mathcal{P}}\tau_b, 0). \tag{3.12}$$

Note that the delayed differential equation (3.11) combined the condition $x(t_0) = x_0$ can defined an infinite solutions (this class of differential equation is not covered by Cauchy-Lipschitz's theorem) . To guarantee uniqueness of solutions under locally Lipschitz property of β we need to impose a initial condition known in an interval with length at least $\max_{b\in\mathcal{P}}\tau_b =: \tau$.

Lemma 3.7.1.2. *If β is locally Lipschitz, ϕ is continuous on $(-\max_{b\in\mathcal{P}}\tau_b, 0)$ and F is generated by regular $G-$function $G : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ satisfying*

$$G(v, x)_{|v=b} = F_b(x).$$

Then the delayed evolutionary game dynamics defined by (3.11) and (3.12) has a unique solution.

The dynamic is said positively correlated (PC) if

$$\frac{d}{dt}x \neq 0 \implies \sum_b [\frac{d}{dt}x_b][G(v, x)_{|v=b}] > 0.$$

In (168), Sandholm showed that in absence of time delays, replicator dynamics (197) (or general imitation dynamics), Smith dynamics (189) (pairwise comparison dynamics), projection dynamics (168), Brown-von Neumann-Nash dynamics (55) (excess payoff dynamics) satisfy the positive correlation (PC) property. Moreover Brown-von Neumann-Nash dynamics, projection dynamics and Smith dynamics satisfy the property that every rest point of the dynamics is an equilibrium of the game. It is easy to see that the parameter $\mu > 0$ and the times does not change the set of equilibria. Hence, the following holds:

Corollary 3.7.1.3. *Delayed imitation dynamics, delayed pairwise comparison dynamics, delayed projection dynamics, excess payoff dynamics and projection dynamics satisfy the positive correlation (PC) property. Moreover Brown-von Neumann-Nash dynamics, projection dynamics and Smith dynamics satisfy the property that every rest point of the dynamics is an equilibrium of the game.*

Examples 3.7.1.4. *If*

$$\beta_{p_i, p_j} = x_{p_j}(t)\,max\left(0, G(w, x(t-\tau_w))_{|w=p_j} - G(v, x(t-\tau_v))_{|v=p_i}\right)$$

we obtain the delayed replicator dynamics

$$\frac{d}{dt}x_{p_j}(t) = x_{p_j}(t)\left[G(w,x(t-\tau_w))_{|w=p_j} - \sum_{k=0}^{n} x_{p_k}(t)G(w',x(t-\tau_{w'}))_{|w'=p_k}\right]$$
$$j = 0,\ldots,n,$$

Analogously, the delayed Brown-von Neumann-Nash dynamics is obtained for $\beta_{p_i,p_j} = \max(0,gg)$ where

$$gg = G_w(x(t-\tau_w))_{|w=p_j} - \sum_{k=0}^{n} x_{p_k}(t)G(w',x(t-\tau_{w'}))_{|w'=p_k}$$

and the delayed $\theta-$Smith dynamics for

$$\beta_{p_i,p_j} = \max\left(0, G(w,x(t-\tau_w))_{|w=p_j} - G(v,x(t-\tau_v))_{|v=p_i}\right)^\theta,$$

with $\theta \geq 1$. Note that under delayed evolutionary game dynamics, the ESS can be unstable (see Fig.3.35). If the dynamics is regular, a sufficient condition of stability of an ESS x^* is given by: (i) all roots of the Jacobian of the function H defined by

$$H_{p_j}(y) = y_{p_j}(t)\left[G(w,y(t))_{|w=p_j} - \sum_{k=0}^{n} y_{p_k}(t)G(w',y(t))_{|w'=p_k}\right]$$

have negative real value the ESS x^* and (ii)

$$\|J(x^*)\|_\infty \left(\max_{b\in\mathcal{P}} \tau_b\right) < 1$$

where $\|J(x^*)\|_\infty$ denotes the norm sup of Jacobian of H.

Dirac distribution

In this subsection, M is the Dirac distribution δ_{m-1} (the number of opponents in the same range is $m-1$). If there is no pricing for the energy consumption then, the high power level weakly dominates the others power levels. We show that the one-shot random access game between m mobiles and without cost $c(.) \equiv 0$ (degenerate case) has an infinite number of equilibria. Only $(p_n,\ldots,p_n)$ is a symmetric equilibrium.

Proposition 3.7.1.5. *If $c(.) \equiv 0$, then the one-shot random access game between m mobiles has an infinite number of Nash equilibria and unique a symmetric Nash equilibrium $(p_n,\ldots,p_n)$ which is an evolutionarily stable state. Moreover, the game has many Pareto optima[3].*

Proof. (i) **Symmetric Equilibrium** It is clear that for all j, $u^j(p_n,x^{-j}) \geq u^j(y^j,x^{-j})$, $\forall x^{-j} \in (\Delta_n)^{m-1}$ i.e p_n weakly dominates the others power levels in $\mathcal{P}\backslash\{p_n\}$. Hence, the strategy p_n is an equilibrium. Moreover the best reply to the population profile x is to play p_n if $x_{p_n} \neq 1$, and to play any strategy $z \in \Delta_n$ if $x_{p_n} = 1$. Thus, $(p_n,p_n,\ldots,p_n)$ is the unique symmetric equilibrium. At the equilibrium $(p_n,p_n,\ldots,p_n)$, the payoff of each mobile is zero. Thus, $(p_n,p_n,\ldots,p_n)$ is not Pareto optimal because the allocation obtained at $(p_n,0,\ldots,0)$ or $(p_n,p_1,p_1,\ldots,p_1)$ Pareto dominates zero.

---

[3]An allocation of payoffs is Pareto optimal or Pareto efficient if there is no other feasible allocation that makes every user at least as well off and at least one user strictly better off

(ii) **Pure Equilibria**: Fix a mobile j which uses the action p_n. Then any action profile of the others mobiles $x^{-j} \in \in \mathcal{P}^{m-1}$ leads to a Nash equilibrium (no mobile can improve its probability of success by deviating unilaterally). In particular, if $j_0, j_2, \ldots, j_{n-1}$ such that j_i of the $m-1$ mobiles p_i and $j_0 + \ldots + j_{n-1} = m - 1$ then no mobile can improve its probability of success by deviating unilaterally and this kind of configuration where only mobile use p_n is also Pareto optimal. Thus, the game has the binomial number $\binom{m}{1}$ times the cardinality of $\mathcal{P}^{m-1}$ which is mn^{m-1} pure Nash-Pareto equilibria (pure Nash equilibria which are also optimal in Pareto sense). The total number of pure Nash equilibria is given by the cardinality of the set

$$\left\{ (j_0, \ldots, j_n) \in \mathbb{N}^{n+1} | \; j_n \geq 1, \; \sum_{i=0}^{n} j_i = m, \; j_i \geq 0 \right\}$$

(iii)**Partially Mixed Equilibria:** Any situation where at least one of the mobiles use the strategy p_n, and other mobiles use an arbitrary mixed strategy, gives a mixed Nash equilibria. The allocation of payoff obtained in these partially mixed strategy need not to be Pareto optimal if at least one mobile chooses the strategy p_n with strictly positive probability.

(iv) **Evolutionarily Stable State:** The unique symmetric equilibrium p_n is evolutionarily stable. $\qquad\square$

A multi-strategy is a strong equilibria if is a configuration from which no coalition (of any size) can deviate and improve their payoff (probability minus cost) of every member of the coalition (group of the simultaneous moves), while possibly lowering the payoff of mobiles outside the coalition group. A strong equilibrium is in particular a strong equilibrium (by taking coalition of size) but also Pareto optimal (coalition of full size). The Theorem 3.7.2 below describes Nash equilibria, Pareto optimality and coalition proof of two strategies $\{p_0, p_1\}$ access game between m mobiles.

Theorem 3.7.2. *Suppose that $n = 1$ and $0 = c(p_0) < c(p_1) < V$. Then the one-shot random access game has (i) $2^m - 1$ number of Nash equilibria, (ii) m of them are Nash-Pareto equilibria, and strong equilibria, (iii) a unique fully mixed Nash equilibrium which is not Pareto optimal. (iv) a unique ESS given by $\left(1 - (\frac{c(p_1)}{V})^{\frac{1}{m-1}}, (\frac{c(p_1)}{V})^{\frac{1}{m-1}} \right)$. (v) the price of anarchy is zero for both ESS and Nash equilibria.*

Proof. (Proof the Theorem 3.7.2) We show the following results: (i) If only one node transmit with p_1 and the others stay quiet (p_0), then the mobile which transmit with p_1 gets the payoff $V - c(p_1)$ and the others receive nothing and has no cost $c(p_0) = 0$. This configuration is an equilibrium point. (ii) there are exactly m pure equilibria and all these pure equilibria are Pareto optimal (because the maximum to total payoff is attained) and strong equilibria. (iii) k ($1 \leq k < m - 1$) of the m mobiles choose to stay quiet and the $m - k$ others are active and play the optimal mixed strategy in the game with $m - k$ mobiles denote by $\Gamma^{m-k} : \left(1 - \nu^{\frac{1}{m-k-1}}, \nu^{\frac{1}{m-k-1}} \right)$ where $\nu := \dfrac{c(p_1)}{V}$. Thus, there are exactly

$$\sum_{k=1}^{m-2} \binom{m}{k} = 2^m - (m + 2)$$

partially mixed Nash equilibria. (iv) The game has a unique strictly mixed Nash equilibrium given by $\left(1 - \nu^{\frac{1}{m-1}}, \nu^{\frac{1}{m-1}} \right)$. (v) the allocation of payoff obtained in these (partially or completely) mixed strategy are not Pareto optimal.

Pure equilibria Suppose that the node i transmits and the others $\mathcal{N}\backslash\{i\}$ stay quiet. Then mobile i will obtain the maximum payoff $V - c(p_1) > 0$. Mobile i obtains 0 by deviating (stay quiet). Hence mobile i has not incentive to deviate unilaterally. A mobile $j \neq i$ will receive 0 and have no cost in this configuration. Suppose that mobile j decide to deviate unilaterally then mobile j payoff will be $-c(p_1) < 0$. Hence, no mobile has incentive to deviate unilaterally from this configuration. We change the role of mobile i and another mobile j to obtain the others pure equilibria (permutation).

Completely mixed equilibrium The payoff of node i is $f_m(s) = s^i\left[(-c(p_1)(1-\eta) + (V - c(p_1))\eta\right]$ where $\eta = \prod_{j\neq i}(1 - s^j)$. The best response ($BR^i$) of the node i to the multi-strategy s^{-i} of the others nodes is given by

$$BR^i(s^{-i}) = \begin{cases} 1 & \text{if } \eta > \dfrac{c(p_1)}{V} \\ 0 & \text{if } \eta < \dfrac{c(p_1)}{V} \\ [0,1] & \text{if } \eta = \dfrac{c(p_1)}{V} \end{cases}$$

Note that at η will be independent of i. Thus, for $i, k \in \{1, \ldots, m\}$, the quotient $\dfrac{\prod_{j\neq i}(1 - s^j)}{\prod_{j\neq k}(1 - s^m)} = $

$1 = \dfrac{1 - s^i}{1 - s^k}$. Hence, $s^1 = \ldots = s^m$. We conclude that $s^j = 1 - \sqrt[(m-1)]{v}$, $\forall j \in \mathcal{N}$ is the unique strictly mixed Nash equilibrium.

Partially mixed equilibria Suppose that $m > 2$ and fix an integer k between 1 and $m - 2$. Suppose k nodes stay quiet and $m - k$ transmit with positive probability. Without lost of generality, we can suppose that the k first nodes $\{1, 2, \ldots, k\}$ does not transmit and any node j with $j > k$ participate the game Γ^{m-k}. The optimal strategy for j ($j > k$) is $s^j = 1 - \sqrt[(m-k-1)]{v}$. For $j \leq k$, we show that mobile j have no incentive to deviate because payoff obtained by node j by deviating is

$$s^i\left[(-c(p_1))(1 - v) + (V)v\right] - 0 \times (1 - s^i)v$$

which is lower than 0. For $m = 2$, the game becomes the two-player game

1\2	p_1	p_0
p_1	$(-c(p_1), -c(p_1))$	$(V - c(p_1), 0)$
p_0	$(0, V - c(p_1))$	$(0, 0)$

It is easy to see that the Nash equilibria of the strategic game Γ^1 are (p_1, p_0), (p_0, p_1) and $(1 - v, v)$.

Pareto optimality Note that the allocation of payoff obtained in these (partially or completely) mixed strategy are not Pareto optimal because the payoff of each mobile is zero at these configuration. This completes the proof. □

The negative results (ii), (iii), (v) of Theorem 3.7.2 in term of performance can be improved by introducing the concept of correlated evolutionarily stable state (CESS). We will introduce CESS in next subsection and exhibit a class of CESS for which the system is stable and 100% efficient (the payoff at the CESS is a social welfare) and the system is stable in long-term. We first remark that the result of Theorem 3.7.2 (i),(ii),(iv) for two strategies $\sharp\mathcal{P} = 2$ does not holds for arbitrary number of strategies $\sharp\mathcal{P} \geq 3$ as shown in the following examples (see Fig.3.27) where pure equilibria may not exists.

The three strategies two-player game has no equilibrium in pure strategies (see Fig 3.28 and 3.27). The game has a unique completely mixed strategy

$$\left(\frac{V - c(p_2)}{V}, \frac{c(p_2) - c(p_1)}{V}, \frac{c(p_1)}{V}\right) \in \Delta_2$$

obtained by solving the indifference equations.

	p_2	p_1	p_0
p_2	$(-c(p_2), -c(p_2))$	$(V - c(p_2), -c(p_1))$	$(V - c(p_2), 0)$
p_1	$(-c(p_1), V - c(p_2))$	$(-c(p_1), -c(p_1))$	$(V - c(p_1), 0)$
p_0	$(0, V - c(p_2))$	$(0, V - c(p_1))$	$(0, 0)$

Figure 3.27: No pure equilibria. Three strategies

$A0 :=$

	p_2	p_1	p_0
p_2	$(-c(p_2), -c(p_2), -c(p_2))$	$(-c(p_2), -c(p_1), -c(p_2))$	$(-c(p_2), 0, -c(p_2))$
p_1	$(-c(p_1), -c(p_2), -c(p_2))$	$(-c(p_1), -c(p_1), V - c(p_2))$	$(-c(p_1), 0, V - c(p_2))$
p_0	$(0, -c(p_2), -c(p_2))$	$(0, -c(p_1), V - c(p_2))$	$(0, 0, V - c(p_2))$

$A1 :=$

	p_2	p_1	p_0
p_2	$(-c(p_2), -c(p_2), -c(p_1))$	$(V - c(p_2), -c(p_1), -c(p_1))$	$(V - c(p_2), 0, -c(p_1))$
p_1	$(-c(p_1), V - c(p_2), -c(p_1))$	$(-c(p_1), -c(p_1), V - c(p_1))$	$(-c(p_1), 0, -c(p_1))$
p_0	$(0, V - c(p_2), -c(p_1))$	$(0, -c(p_1), -c(p_1))$	$(0, 0, V - c(p_1))$

$A2 :=$

	p_2	p_1	p_0
p_2	$(-c(p_2), -c(p_2), 0)$	$(V - c(p_2), -c(p_1), 0)$	$(V - c(p_2), 0, 0)$
p_1	$(-c(p_1), V - c(p_2), 0)$	$(-c(p_1), -c(p_1), 0)$	$(V - c(p_1), 0, 0)$
p_0	$(0, V - c(p_2), 0)$	$(0, -c(p_1), V - c(p_1))$	$(0, 0, 0)$

Figure 3.28: No pure equilibrium under pricing.

For the remainder we assume that $0 = c(p_0) < c(p_1) < c(p_2) < \ldots < c(p_n) < V$. We have seen in Theorem 3.7.2 that with two strategies, the access game between m mobiles has several pure equilibria and Pareto optimal solutions. The following Theorem 3.7.3 shows that Theorem cannot be extended for three strategies and an equilibrium do not exist in pure strategies.

Theorem 3.7.3. *Let $\sharp \mathcal{P}$ be the cardinality of $\mathcal{P}$. The evolutionary random access game with more than two strategies $\sharp \mathcal{P} \geq 3$, and $m \geq 2$ mobiles has no pure equilibrium.*

Proof. Fix a strategy profile $(a^1, \ldots, a^m) \in \mathcal{P}^m$ with $\sharp \mathcal{P} \geq 3$. Let $T1 = \arg\max_j a^j$. We distinguish two cases: $\sharp T1 = 1$ or $\sharp T1 \geq 2$

- suppose that $\sharp T1 = 1$ and let $T1 = \{j^*\}$. One has in particular $p^{j^*} > p_0 = 0$. If $a^{j^*} = p_i < p_n$ then any mobile k in $\mathcal{P} \backslash T1$ which deviates and uses $a^l > p^{j^*} > p_0 = 0$ (the existence of a^l is guaranteed because there are more than three strategies) will improve its payoff from $-c(a^k)$ to $V - c(a^l)$.

  If $a^{j^*} = p_n$ then mobile j^* can improve its payoff from $V - c(p_n)$ to $V - c(b^{j^*})$ by playing a second higher power level $b^{j^*} \in T2 = \arg\max_{j \notin T1} a^j$ or one the others mobiles can save energy.

- suppose now that $\sharp T1 \geq 2$. Then if $(a^1, a^2, \ldots, a^m) \neq (p_0, \ldots, p_0)$ there is collision for all mobiles which transmit, otherwise $(a^1, a^2, \ldots, a^m) = (p_0, \ldots, p_0)$ and any mobile which deviates and uses a power greater than p_0 will have a successful transmission.

In both cases, there is at least one mobile can deviate and improve its payoff. Thus, $(a^1, \ldots, a^m)$ is not an equilibrium. Since $(a^1, \ldots, a^m)$ is an arbitrary pure strategy profile in $\mathcal{P}$ with $\sharp \mathcal{P} \geq 3$. We conclude that the game between m mobiles has no pure equilibrium if the number of pure strategies is at least three. $\qquad \square$

Theorem 3.7.4. *Then the evolutionary access game with m mobiles and at least three strategies has a unique Nash equilibrium (in mixed strategies) given by*

$$x_{p_0} = \left(\frac{c(p_1)}{V} \right)^{\frac{1}{m-1}},$$

$$1 \leq j < n, \ x_{p_j} = \left(\frac{c(p_{j+1})}{V} \right)^{\frac{1}{m-1}} - \left(\frac{c(p_j)}{V} \right)^{\frac{1}{m-1}},$$

$$x_{p_n} = 1 - \sum_{j=1}^{n-1} x_{p_j} - x_{p_0} = 1 - \left(\frac{c(p_n)}{V} \right)^{\frac{1}{m-1}}$$

Proof. The result follows by using the inverse of the function x^m in $(0,1)$ and solving the system (Cramer invertible)

$$
\begin{pmatrix}
1 & 0 & 0 & \ldots & 0 & 0 \\
1 & 1 & 0 & \ldots & 0 & 0 \\
1 & 1 & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ldots & 0 \\
1 & 1 & 1 & \ldots & 1 & 0 \\
1 & 1 & 1 & \ldots & 1 & 1
\end{pmatrix}
\begin{pmatrix}
x_{p_0} \\
x_{p_1} \\
x_{p_2} \\
\vdots \\
x_{p_{n-1}} \\
x_{p_n}
\end{pmatrix}
=
\begin{pmatrix}
\left(\frac{c(p_1)}{V} \right)^{\frac{1}{m-1}} \\
\left(\frac{c(p_2)}{V} \right)^{\frac{1}{m-1}} \\
\left(\frac{c(p_3)}{V} \right)^{\frac{1}{m-1}} \\
\vdots \\
\left(\frac{c(p_n)}{V} \right)^{\frac{1}{m-1}} \\
1
\end{pmatrix}
$$

$\qquad \square$

Proposition 3.7.4.1. *The price of anarchy and the price of stability (PoS) of Nash equilibria in the random access game are both zero for $\sharp \mathcal{P} \geq 3$.*

Correlated evolutionarily stable state

In this subsection we focus on the concept of *correlated evolutionarily stable state* (CESS) in access games. A correlated equilibrium is introduced by Aumann (32) in 1974 and can be interpreted as a distribution of actions or messages given to the mobiles by some *referee* (which can be the base station or the receiver) before to play each local game. For more details on correlated equilibrium, we refer the reader to (32; 33; 125; 58). As in (125), we use the concept of correlated equilibrium from the perspective of bounded rationality. In term of medium access control, the correlated mechanism phase (the phase of messages reception from the receiver before transmission) can be seen as the analogue of the probing or listening phase in CSMA algorithm (users

willing to use the channel must first probe -or listen to- the channel, and when it is not busy, they can decide to transmit).

Let $\mathcal{N} := \{1, 2, \ldots, m\}$. We define a probability space $\left(\Omega, 2^{\Omega}, \mathbb{P}^m\right)$ which generates signals on which the mobiles can condition their strategic power choices where $\Omega = \mathcal{P}^m$. The set Ω is partitioned as follows $I^j(b) := \{ba^1 a^2 \ldots a^{j-1} a^{j+1} \ldots a^m \mid a^l \in \mathcal{P}, l \neq j\}$, $I^j := \{I^j(b), b \in \mathcal{P}\}$. then, I^j has exactly $|\mathcal{P}| = n + 1$ elements which are information sets of mobile j. We define an *assignment function* (called also *rule*) profile $\alpha = (\alpha^1, \ldots, \alpha^m)$ as a mapping from the set of states or signals Ω to mixed strategies set $\Delta(\mathcal{P})$. For all j, the assignment function of the mobile j, α^j must satisfy : if for some w, $\alpha^j(w) = b$, then $\alpha^j(w') = b$, for all $w' \in I^j(s)$.

That is, for each element $w \in I^j$, mobile j cannot distinguish states that are in the same information set (same equivalence class). We denote the set of all pure assignment functions by $\mathcal{AF} : \{g \mid g : \mathcal{P} \longrightarrow \mathcal{P}\}$. Thus, when a mobile chooses an assignment function α and when he receives the signal $\omega \in \Omega$ from the referee, he will choose the mixed action $\alpha(\omega)$. We use $\alpha(a^j|\omega)$ to denote the probability assigned on a^j under this mixed action $\alpha(\omega)$. Then $\alpha(w) = [\alpha(p_0|\omega), \ldots, \alpha(p_n|w)] \in \Delta(\mathcal{P})$. Given a referee $(\Omega, \mathbb{P})$, we define the identity assignment function as $\alpha^{id,m}(a^j|\omega) = 1$ if $proj^j(\omega) = a^j$ where $proj^j(\omega)$ denotes the $j-$th element of the signal ω and $\alpha^{id,m}(a^j|\omega) = 0$ for all ω such that $proj^j(\omega) \neq a^j$. If each mobile use the identity assignment then, the resulting probability distribution of actions profile actually played will be $\mathbb{P}$, the same as the probability distribution of actions recommended by the referee. But when mobiles use other assignments, a different distribution may result. Given an assignment profile α and a probability distribution $\mathbb{P}$ over $\mathcal{P}^m$, the expected payoff is given by

$$
\begin{aligned}
f(\alpha) &= \sum_{\omega \in \Omega} \mathbb{P}(\omega) \sum_{a \in \mathcal{P}^m} u_m(a) \prod_{j \in \mathcal{N}} \alpha^j(a^j|w) \\
&= \sum_{a \in \mathcal{P}^m} u_m(a) \sum_{\omega \in \Omega} \mathbb{P}(\omega) \prod_{j \in \mathcal{N}} \alpha^j(a^j|w) \\
&= \sum_{a \in \mathcal{P}^m} u_m(a) \mathbb{Q}_\alpha(a)
\end{aligned}
$$

where $\mathbb{Q}_\alpha(a) = \sum_{\omega \in \Omega} \mathbb{P}(\omega) \prod_{j \in \mathcal{N}} \alpha^j(a^j|w)$. We say that two assignment functions profile α and β are equivalent if they induce the same value $\mathbb{Q}_\alpha = \mathbb{Q}_\beta$

Proposition 3.7.4.2. *Suppose that for all $a \in \mathcal{P}^m$ $\mathbb{P}(a) = \mathbb{P}(\sigma(a))$, for all permutation σ (the distribution $\mathbb{P}$ is said symmetric). If an assignment profile β is equivalent to $\alpha^{id,m}$ then $\mathbb{Q}_\beta(a) = \mathbb{P}(a)$, for all $a \in \mathcal{P}^m$.*

Proof. One has, $\mathbb{Q}_{\alpha^{id,m}}(a) = \mathbb{P}(a)$, $\forall a \in \mathcal{P}^m$. The result follows immediately. $\qquad \square$

Now, suppose a small group of mutants appears. These mutants use a mutational assignment function α'^m. which is not equivalent to the identical assignment function $\alpha^{id,m}$., but they cannot change the referee recommendation. Let ϵ be the portion of the population which are mutants (who use α',) and $1 - \epsilon$ portion of the population are non-mutants who use $\alpha^{id,n}$. At each time, m mobiles are randomly chosen to play the strategic access game. In playing the access game, the mobiles have the same referee $(\Omega, \mathbb{P})$). A probability distribution $\mathbb{P}$ over $\mathcal{P}^m$ is an CESS if non-mutants with identity assignment function perform better than mutants with assignment functions that are not equivalent to the identical assignment function.

Definition 3.7.4.3. *A CESS $\mathbb{P}$ is a symmetric distribution probability over $\mathcal{P}^m$ such that for every assignment function α'^m nonequivalent to the identical assignment function $\alpha^{id,m}$, there exists some $\epsilon_{\alpha'^m} > 0$ such that*

$$F^m(\alpha^{id,m}, \epsilon\alpha'^m + (1-\epsilon)\alpha^{id,m}) > F^m(\alpha'^m, \epsilon\alpha'^m + (1-\epsilon)\alpha^{id,m}) \tag{3.13}$$

for all $\epsilon \in (0, \epsilon_{\alpha'^m})$ where $F^m(\beta, \alpha) = \sum_{r \in \mathcal{AF}} \beta_r f(r, \alpha, \ldots, \alpha)$

When the inequality (3.13) is non-strict and $\epsilon = 0$, the probability distribution $\mathbb{P}$ is a correlated equilibrium (CE). Hence, a CESS is a CE.

Proposition 3.7.4.4. *If x is an ESS then the product measure $x^{\star(m)}$ given by $x^{\star(m)}(a) = \prod_{j \in \mathcal{N}} x^j(a^j)$ is a CESS.*

Proof. It is easy to see that $x^{\star(m)}$ is a probability measure on $\mathcal{P}^m$. Since x is an ESS, $(x, \ldots, x) \in \Delta(\mathcal{P})^m$ is a Nash equilibrium of the $m-$player game without correlated device ($\Omega = \varnothing$ or mobiles ignore the signals) i.e

$$\sum_{a \in \mathcal{P}^m} u_m(a) \prod_{j \in \mathcal{N}} x(a^j) \geq \sum_{a \in \mathcal{P}^m} u_m(b^k, a^{-k}) \prod_{j \in \mathcal{N}\backslash\{k\}} x(a^j) \geq 0, \forall b^k \in \mathcal{P}.$$

This means that the product probability measure $x^{\star(m)}$ is a CE. Moreover,

$$u_m(x_\epsilon, \ldots, x_\epsilon, x, x_\epsilon, \ldots, x_\epsilon) > u_m(x_\epsilon, \ldots, x_\epsilon, y, x_\epsilon, \ldots, x_\epsilon), \ \forall y$$

implies that

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega) \sum_{a \in \mathcal{P}^m} u_m(a) x(a^i|w) \prod_{j \in \mathcal{N}\backslash\{i\}} x_\epsilon(a^j|w) > \sum_{\omega \in \Omega} \mathbb{P}(\omega) \sum_{a \in \mathcal{P}^m} u_m(a) y(a^i|w) \prod_{j \in \mathcal{N}\backslash\{i\}} x_\epsilon(a^j|w)$$

where $\mathbb{P}(a) = \prod_{j \in \mathcal{N}} x(a^j)$, $x_\epsilon := (1-\epsilon)x + \epsilon y$. $\square$

More generally all convex combination of Nash equilibria (NE) are correlated equilibria and the set of CE is convex. This convexity properties of the CE's set is used in some algorithms based on linear or convex programming and generically, the complexity is less than in Nash equilibria in finite action games. We denote by Δ_{CESS} the set of CESS. Then,

Lemma 3.7.4.5. *The following inclusions holds:*

$$\Delta_{ESS} \subset \Delta_{NE} \subset \Delta_{CE}, \Delta_{ESS} \subset \Delta_{CESS} \subset \Delta_{CE}, \tag{3.14}$$

One of the principal interest to study CESS is that, in general CESS remains stable for small delays in various game dynamics and the expected payoff obtained at ESS can be improved by CESS distribution. Non-ESS can be CESS as shown below.

We now construct of a class CESS for arbitrary number of users with two actions.

Number of mobiles which transmit with power p_i We define $numb_{p_i} : \mathcal{P}^m \longrightarrow \mathbb{N}$ as the number of transmitters of the power p_i. For example, $numb_{p_n}(p_n, \ldots, p_n) = m$ and $numb_{p_n}(p_n, a^2, \ldots, a^m) = 1$ if $a^j \neq p_n, \forall j \geq 2$. The set $\{b \in \mathcal{P}^m \mid numb_{p_1}(b) = 1, numb_{p_0}(b) = m-1\}$ is exactly the situations where only one of the mobiles transmits with the power p_n and the others mobiles use the power level p_0 (do not transmit). $numb_T(b) = 1$ if and only if there exists a unique j such that $a^j = p_n$ i.e all permutations of the actions profile $(p_n, a^2, \ldots, a^m)$ and permutations satisfying $p^j < p_n \forall j$. There are mn^{m-1} possibilities.

Denote by $\Theta^m := \{a \in \mathcal{P}^m \mid numb_{p_1}(a) = 1, numb_{p_0}(a) = m-1\}$.

Proposition 3.7.4.6. *For two strategies, the probability distribution $\mathbb{P}^m$ over $\mathcal{P}^m$ defined as*

$$\mathbb{P}^m(a) = \left\{ \begin{array}{ll} \dfrac{1}{m} & \textit{if } a \in \Theta^m \\ 0 & \textit{otherwise} \end{array} \right.$$

is a CESS with the payoff $\dfrac{V - c(p_1)}{m} > 0$. Moreover at each slot, we have a successful transmission with probability one (if each mobile have some packet to transmit at each slot) and hence, the CESS leads to a social optimum. The price of anarchy of the class of assignment functions with the distribution $\mathbb{P}^m$ is one (i.e the proposed method is 100% efficient). This class of CESS is also energy efficient (the energy consumption is minimized).

Proof. Θ^m has exactly m actions. Thus, $\mathbb{P}^m$ is a probability measure. We show that every strategy $\alpha \in \mathcal{P}$ used in $\mathbb{P}^m$, $\sum_{a^{-j} \in \mathcal{P}^{m-1}} \mathbb{P}^m(\alpha, a^{-j}) > 0$ and any alternative strategy $b^j \in \mathcal{P} \backslash \{\alpha\}$, it holds that $\sum_{a^{-j} \in \mathcal{P}^{m-1}} \mathbb{P}^m(\alpha, a^{-j}) \left[u(\alpha, a^{-j}) - u(b^j, a^{-j}) \right] > 0$ i.e Since $n = 1$, we have verify for the inequality for two strategies: for $b^j \in \mathcal{P} \backslash \{p_1\}$ one has, $[u_m(p_1, a^{-j}) - u_m(b^j, a^{-j})]\mathbb{P}^m(p_1, a^{-j}) = \dfrac{V - c(p_1) + c(b^j)}{m} > 0$. By summing over $a^{-j} \in \mathcal{P}^{m-1}$, $\sum_{a^{-j} \in \mathcal{P}^{m-1}} [u_m(p_1, a^{-j}) - u_m(b^j, a^{-j})]\mathbb{P}^m(p_1, a^{-j}) > 0$ for all permutation on the position of j and $b^j \neq p_0$. Similarly, for $b^j \in \mathcal{P} \backslash \{p_0\}$, $a^{-j} \in \Theta^{m-1}$, one has, $[u_m(p_0, a^{-j}) - u_m(b^j, a^{-j})]\mathbb{P}^m(p_1, a^{-j}) = \dfrac{c(p_1)}{m} > 0$ Hence $\mathbb{P}^m$ is a strict CE. Thus, $\mathbb{P}^m$ is a CESS and the system has a successful transmission at each slot with probability one. The total payoff obtained at the CESS is exactly $V - c(p_1)$ which is strictly greater than the payoff obtained at the ESS (zero). The social optimum is $V - c(p_1)$ which guarantees a successful transmission with the minimum power consumption. thus, the CESS is efficient in term of energy consumption (by considering par example the ratio between the probability of success and cost of energy consumption as the energy-efficient metric). $\square$

Benefit of Correlation

One of the advantages of a correlated evolutionary stable state is that it has a potential to reduce the distance between the optimal solution and the ESS solution obtained as an outcome of users which decide independently.

Proposition 3.7.4.7. *The following inequalities holds:*

$$PoS_{CESS} \geq PoS_{ESS} \geq PoA_{ESS} \geq PoA_{NE} \geq PoA_{CE} \tag{3.15}$$

$$PoS_{CE} \geq PoS_{NE} \geq PoS_{ESS} \geq PoA_{ESS} \geq PoA_{NE} \geq PoA_{CE} \tag{3.16}$$

Proof. Apply the Lemma 3.7.4.5. $\square$

We define the benefit of correlation (BoC) as $PoS_{CESS} - PoA_{NE}$.

Corollary 3.7.4.8. *The benefit of correlation in the random access game is 100% for $\Omega = \Theta^m$.*

$$BoC = 1$$

Probability of success: Comparison at NE, ESS and CESS

In the numerical examples below, we show how the pricing function can optimize the network throughput or probability of success. we first investigate the impact of the different parameters and pricing on the NE, ESS and CESS and the convergence of the replicator dynamic with time delays. We also discuss the impact of the pricing function on the system capacity. We took the parameters: three strategies $\mathcal{P} = \{p_0, p_1, p_2\}$, $n = 2$, $V = 1$, $\mu = 0.8$ (80% of coverage). The expected probability of success of a mobile is given by $\sum_{j=1}^{n} x_{p_j} G_{\mathcal{K}}(x_{p_0} + \ldots + x_{p_{j-1}})$. We consider three examples of distribution: geometric distribution $Geo(p = 0.3)$ and Dirac distribution $\delta_m, m = 30$.
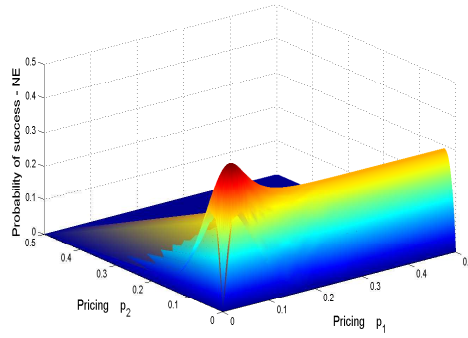
Figure 3.29: Probability of success - Dirac distribution $m = 30$
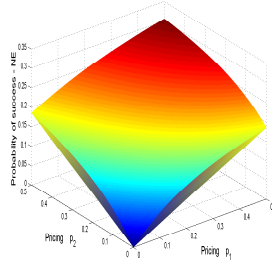
Figure 3.30: Geometric probability of success with parameter $p = 0.3$

These numerical examples (Fig.3.29,3.30, 3.31) confirm the optimality of CESS in random access game with two strategies: probability to have a successful transmission at the CESS is maximum and equal to the probability that a resource exists times the probability that a randomly selected mobile has a packet to transmit which we represent by $\mu = 0.8$.

Limitation of CESS's approach in access game

We have seen that for $\sharp\mathcal{P} \leq 2$ or for $c(.) = 0$, CESS can improve the performance of system and some class of CESS are social welfare. But for $\sharp\mathcal{P} \geq 3$, the correlated equilibria is reduced the Nash equilibria. Hence, the correlation is not needed. If we reduce the set of signal (message)
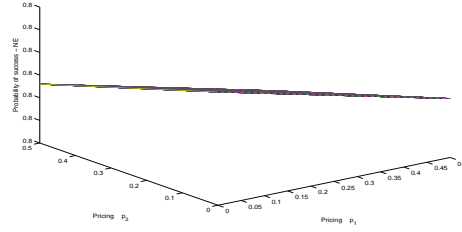
Figure 3.31: Probability of success at CESS

to $\Omega = \Theta^m$ we obtain that If $\sharp\mathcal{P} = 2$ then $PoS_{CESS} = 1 = PoS_{N}E = PoA_{CESS}$ and $0 = PoS_{ESS} = PoA_{ESS} = PoA_{NE}$. If $\sharp\mathcal{P} \geq 3$ then $PoS_{CESS} = PoA_{CESS} = PoS_{ESS} = PoA_{ESS} = PoS_{N}E = PoA_{NE} = 0$.

Game dynamics for CESS

The class of game dynamics based on rule or assignment functions $r \in \mathcal{AF}$ can be extended to correlated strategies. These dynamics are systems of non-linear delay differential with $\sharp\mathcal{AF} = m^m$ equations. The revision protocol β gives a matrix with size m^{2m}. y be a distribution of probabilities on $\mathcal{AF}$. Then y_r is the fraction of the population of mobiles with the assignment function r. The evolutionary game dynamics is then given by

$$\frac{d}{dt}y_r(t) = \sum_{\bar{r} \in \mathcal{AF}} y_{\bar{r}}(t)\beta_{\bar{r}r}(\bar{G}, y) - y_r \sum_{\bar{r} \in \mathcal{AF}} \beta_{r\bar{r}}(\bar{G}, y) \tag{3.17}$$

where $\bar{G}$ is a G-function defined on a more general space that contains the set of assignment function satisfying

$$\bar{G}_v(y)_{v=r} = F_r(y) = \sum_{\bar{r}^{-j} \in \mathcal{AF}^{m-1}} f_m(r, \bar{r}^{-j}) \prod_{i \neq j} y_{\bar{r}^i}.$$

3.7.5 General distribution

Theorem 3.7.6. *The evolutionary access game with random number of interacting mobiles around each receiver and at least three strategies has a unique fully mixed Nash equilibrium given by*

$$x_{p_0} = G_{\mathcal{K}}^{-1}\left(\frac{c(p_1)}{V}\right),$$

$$1 \leq j < n, \; x_{p_j} = G_{\mathcal{K}}^{-1}\left(\frac{c(p_{j+1})}{V}\right) - G_{\mathcal{K}}^{-1}\left(\frac{c(p_j)}{V}\right),$$

$$x_{p_n} = 1 - \sum_{j=1}^{n-1} x_{p_j} - x_{p_0} = 1 - G_{\mathcal{K}}^{-1}\left(\frac{c(p_n)}{V}\right)$$

under the condition that $P(\mathcal{K} = 0) < \dfrac{c(p_1)}{V}$

90

Proof. If $P(\mathcal{K} = 0) < \left(\dfrac{c(p_1)}{V} \right)$ all the real numbers $\dfrac{c(p_1)}{V}, \dfrac{c(p_2)}{V}, \ldots, \dfrac{c(p_n)}{V}$ are in set $G_{\mathcal{K}}(I)$ where I is the interval $(0, 1)$ and $G_{\mathcal{K}}$ is a bijection from I to $(P(\mathcal{K} = 0), 1)$. The result follows by using the inverse of the G-function of the random variable in $(0, 1)$ and solving the system (Cramer invertible)

$$
\begin{pmatrix}
1 & 0 & 0 & \ldots & 0 & 0 \\
1 & 1 & 0 & \ldots & 0 & 0 \\
1 & 1 & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ldots & 0 \\
1 & 1 & 1 & \ldots & 1 & 0 \\
1 & 1 & 1 & \ldots & 1 & 1
\end{pmatrix}
\begin{pmatrix}
x_{p_0} \\
x_{p_1} \\
x_{p_2} \\
\vdots \\
x_{p_{n-1}} \\
x_{p_n}
\end{pmatrix}
=
\begin{pmatrix}
G_{\mathcal{K}}^{-1}\left(\dfrac{c(p_1)}{V} \right) \\
G_{\mathcal{K}}^{-1}\left(\dfrac{c(p_2)}{V} \right) \\
G_{\mathcal{K}}^{-1}\left(\dfrac{c(p_3)}{V} \right) \\
\vdots \\
G_{\mathcal{K}}^{-1}\left(\dfrac{c(p_n)}{V} \right) \\
1
\end{pmatrix}
$$

$\square$

Proposition 3.7.6.1. *For any distribution of the random variable $\mathcal{K}$, and $\sharp \mathcal{P} \geq 3$, the following results holds :* $PoS_{NE} = PoS_{ESS} = 0 = PoA_{ESS} = PoA_{NE}$

Non-convergence and instability to the ESS

In figure 3.32, we plot the trajectories of replicator dynamics without time delays and three strategies. We observe the convergence to the interior rest point starting from any in the relative interior of the simplex. Moreover, the interior rest point is asymptotically stable. The parameters are $V = 1, c(p_2) = 1/2, c(p_1) = 1/4, c(p_0) = 0$.
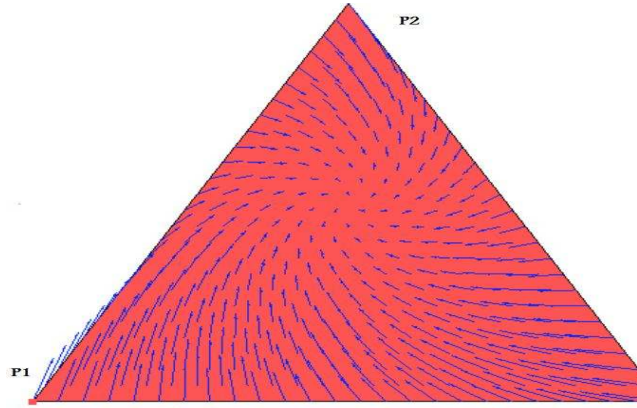
Figure 3.32: Three powers: replicator dynamics

In figures 3.33, 3.34 and 3.35, we describe numerical application of our evolutionary game model with Poisson distribution of mobiles under the delayed replicator dynamics in which

each pure strategy is associated with its own delay ($\tau_{p_0} = 0$, $\tau_{p_1} = \tau$). We took a poisson distribution with parameter $\lambda = 5$. in Poisson distribution. In the figure 3.35 we illustrate a situation where the ESS is not stable (for large delays) and in the figure 3.33 and 3.35 the dynamics converge to the ESS which is stable (for small delays).
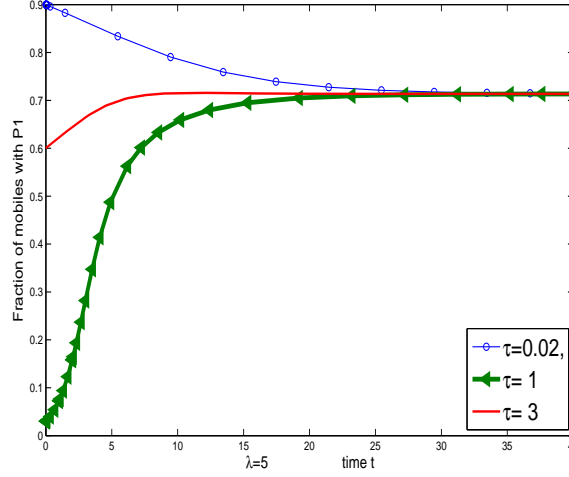
Figure 3.33: *Convergence and stability of ESS for small delays*
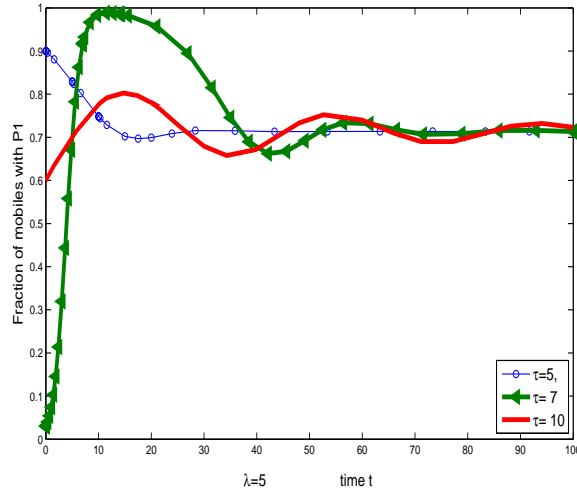
Figure 3.34: *Oscillation with decreasing amplitude.*

These figures show that the ESS can be unstable for large time delays. In the particular case of pairwise interactions $\mathcal{K} = 1$ where the payoff functions F_b are linear, there is a method to obtain the equilibrium from oscillating trajectories of the delayed replicator dynamics if its remains at the relative interior of the simplex starting from any interior function $\phi(t) \in int\Delta_n$, $t \in (-\tau, 0)$, then the time average trajectories of the delayed replicator dynamics
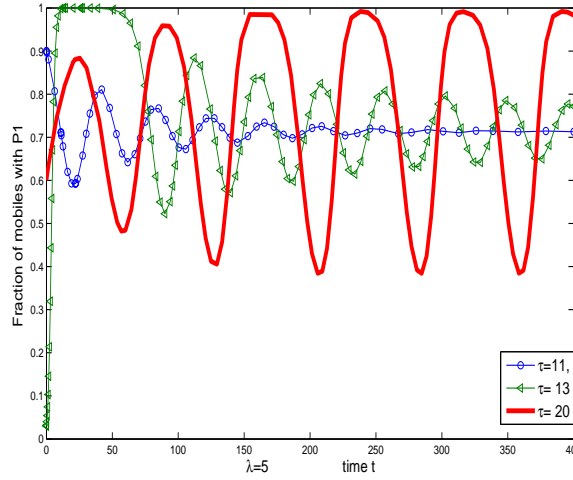
Figure 3.35: Instability of ESS: non-convergence.

converge to the ESS i.e

$$A_T = \frac{1}{T-s} \int_s^T x_{p_j}(t)\, dt \longrightarrow x_{p_j}^*.$$

Since the access game has a unique interior equilibrium, the delayed replicator dynamics has a unique interior rest point. Since the set Δ_n is compact, there exists ϕ such that the subsequence $A_{\phi(T)}$ converges (to some point x_j). Similarly, the vector $A_{\xi(T)}$ converges to x. We show that x is a rest point of the delayed replicator dynamics.

$$\frac{\log(\frac{x_{p_j}(\xi(T))}{x_{p_j}(s)})}{\xi(T) - s} = \frac{1}{\xi(T) - s} G\left(w, \int_s^{\xi(T)} x(t - \tau_w)\, dt\right)_{w=p_j} \tag{3.18}$$
$$- \frac{1}{\xi(T) - s} \sum_{b \in \mathcal{P}} \int_s^{\xi(T)} x_b(t) G(v, (x(t - \tau_b))_{v=b}\, dt$$

This implies that when T goes the infinity, one has $F_{p_j}(x) = constant$. By uniqueness of the interior equilibrium, we conclude that $x = x^*$. Note that for the Brown-von Neumann-Nash dynamics the limit of the time average trajectories can be different than x^* (see (99; 44) for details on the so-called *Shapley triangle*, Time Average of the Shapley Polygon *TASP*) and Birkoff center.

3.7.7 Extension

We describe two possible extensions of the access game studied in previous subsections. The first is the asymmetric case: multi-class of users and the cost function are not necessarily the same. We give a multi-class evolutionary game dynamics for this model. The second extension is a signal-to-noise-plus-interference ratio (SINR) based admission control in which a mobile has a successful transmission if its SINR of the mobile is greater than some fixed threshold β.

Dynamics for asymmetric random access games

We use the model of multi-class evolutionary game to study the evolution of population profile (frequencies) and the asymptotic stability of (strict) equilibria in long-term. For example, the delayed replicator dynamics becomes

$$\frac{d}{dt}x_{p_j}^e(t) = x_{p_j}^e(t)\left[G^e(w, x(t - \tau_w^e))_{|w=p_j} - \sum_{k=0}^n x_{p_k}(t)G^e(w', x(t - \tau_{w'}^e))_{|w'=p_k}\right] \quad (3.19)$$

$$j = 0, \ldots, n_e, \ e \in \mathcal{E}$$

where $\mathcal{E}$ denotes a index set for class or subpopulation, $x(t) = (x^1(t), \ldots, x^{\sharp\mathcal{E}})$, $x^e(t) = (x_{p_k}^e)_{k=0,\ldots,n_e}$ the state of class e with $n_e + 1$ actions. The number $\tau_{p_k}^e$ denotes time delays of the action p_k in class e and $G(v, x(t - \tau_v^e))_{v=p_k} = F_{p_k}^e(x(t - \tau_{p_k}^e))$ the associated expected payoff received at time t.

SINR-threshold based admission control

We now focus in the case where the value V of successful transmission may depend on the power consumption such as in the effective throughput or capacity depending on the signal to noise plus interference ratio (SINR). The system can accept several mobiles depending on their SINR. The packets of mobiles with a very low SINR (less than β) are lost. In that case the appropriate game formulation is an evolving game with variable number of players and coupled constraints. When facing to m others mobiles, the mobile is subjected to the Quality of Service (QoS) constraints

$$SINR^j(a^j, a^{-j}) = \frac{a^j h^j}{\sigma^2 + \gamma \sum_{k \neq j} a^k h^k} \geq \beta$$

where β is some threshold, σ^2 noise term, h is the gain, γ is the load. The channel decodes correctly if the $SINR^j(a) \geq \beta$. This implies that, given the power level of the mobiles a^{-j}, the mobile j has a successful transmission if its power level is in

$$\mathcal{P}^j(a^{-j}) := \left\{b \in \mathcal{P}^j \mid b \geq \frac{\beta}{h^j}(\sigma^2 + \gamma \sum_{k \neq j} a^k h^k)\right\}.$$

By doing this SINR-threshold admission control, more than one mobile can have successful transmission and the number of collisions is reduced compared the models described in previous Sections. The payoff is then

$$u_{m+1}^j(a^j, a^{-j}) = -c(a^j) + V \begin{cases} 1 & \text{if } a^j \in \mathcal{P}^j(a^{-j}) \\ 0 & \text{otherwise} \end{cases}$$

Let k be $\sup\{l, \frac{p_1 h}{\sigma^2 + l p_1 \gamma \bar{h}} \geq \beta\}$ where $h = \min h^j$, $\bar{h} = \max h^j$. Then k mobiles can have successful transmission if they use the power level p_1 and the others mobiles use p_0. The class of CESS constructed in section 3.7.1 can be formulated as follows:

$$\bar{\Theta}^m = \{b \in \mathcal{P}^m \mid numb_{p_1}(b) = k, numb_{p_0}(b) = m - k\}$$

Under the correlated strategy $\bar{\mathbb{P}}^m$ given by the distribution $\bar{\mathbb{P}}^m(a) = \frac{1}{\sharp\bar{\Theta}^m}$ if $b \in \bar{\Theta}^m$ and 0 otherwise, each user has $\frac{V - c(p_1)}{\sharp\bar{\Theta}^m}$. The expected payoff is then $k(V - c(p_1))$ and total energy consumption is kp_1 which is the minimum possible for k successful transmission.

3.7.8 Notes

Biological models and tools have inspired a growing number of studies and of designs of decentralized wireless networks. In various ways, this work falls in this category of biology-inspired algorithms and decision making theory. The evolutionary game theoretic framework proposed here incorporates several actions for each mobile for interference, and admission control. We have analyzed correlated evolutionary stable states which preserve robustness, stability and high performance compared to ESS and Nash equilibrium. An interesting extension of this work is the evolutionary *communication equilibria* (85) that extends the correlated evolutionary stable states by introducing more general signal (message) space.

Notes

In next chapter we develop a new class of evolutionary game dynamics which covers migration and mobility of players in a one-hop neighborhood. This location-dependent evolutionary game dynamics with multicomponent strategies is well adapted to hybrid systems. Example of applications to power control in heterogeneous networks is given. The chapter is based to our publications in (209; 210).

Chapter 4

Evolutionary game dynamics with migration

In this chapter, we propose a class of bio-inspired evolutionary game dynamics with migration for hybrid population games with many local interactions at the same time. Each local interaction concerns a random number of interacting players. The strategies of a player have two components. Specifically, each player chooses both (i) the region or subpopulation and (ii) a strategy among a finite set of secondary pure strategies in each region. We investigate what impact the restriction the actions, when updating a strategy, a player can change only the secondary strategies associate to the region at a time, has on the population dynamics. We apply this model to the integrated power control and base station assignment problem in a multi-cell in code division multiple access (CDMA) wireless data networks, in OFDMA-based WiMAX environment, and association between technologies in heterogeneous networks with large number of mobiles. We show that global neutrally evolutionary stable states are stationary points of *hybrid* mean dynamics called dynamics with *multicomponent strategies* under the positive correlation conditions. We give some convergence results of our hybrid model in stable population games and potential population games under some particular class of dynamics.

4.1 Introduction

Power control in wireless networks has become an important research area. Since the technology in the current state cannot provide batteries which have small weight and large energy capacity, the design of tools and algorithms for efficient power control is crucial. For a comprehensive survey of recent results on power control in wireless networks an interested reader can consult e.g., (144) and the reference therein. Power control protocols based on game theory have been designed for already ten years starting with the pioneering work (76; 115). Non-cooperative games provide a convenient framework for decentralization and distributed decision making in those applications, where as cooperative approaches of game theory have allowed to handle issues concerning fairness in power allocation (17).

Most applications of game theory to power control consider mobile terminals as players of the same type and study strategic one-shot games with a fixed number of players. Here, we consider a population game with many simultaneous local interactions, where the game is played infinitely under some self-organizing process called "hybrid dynamic" and where each inter-

action concerns a random number of players (213). We develop several class of bio-inspired game dynamics with migration (148) in a *hybrid evolutionary game model* in code division multiple access (CDMA) wireless data networks, in OFDMA-based WiMAX environment and in association problem between several types of technologies such as UMTS 3G, WiMAX, WiFi etc.

A hybrid non-cooperative game model for wireless communication has been studied in (4) as an extension of the classical non-cooperative power control game formulation in CDMA system in which mobiles were considered to be connected to the closest base station (BS).

We consider in this chapter an evolutionary game model and study a *hybrid* dynamic for updating actions along with its convergence to an equilibrium point. We apply this model to the integrated power control and base station assignment problem as in such that each mobile's action space consists of the choice of a power level and a base station in some region which corresponds to a cell. The first advantage to use evolutionary framework to model this problem is that, under suitable assumptions, evolutionary game dynamics give naturally an algorithm for converging to an equilibrium point. Evolutionary Game theory describes the evolution with some dynamic process and involves strategic interaction over time in large populations of users.

Contribution

In the first part of this chapter we develop several classes of bio-inspired game dynamics with migration for hybrid systems. We apply our *hybrid evolutionary game model* in code division multiple access (CDMA) wireless data networks, in OFDMA-based WiMAX environment.

The second part of the chapter focuses on multi-class Stackelberg approach:

- On one hand based on the leader-followers games with constraints in large population of players, general model of discrete and continuous time game dynamics are proposed.

- We apply evolutionary Stakelberg games[1] in several contexts including CDMA cognitive radio networks, OFDMA-based WIMAX environment in wireless networks and association between several technologies in hybrid networks.

- we develop a coordination mechanisms between players. To be best to our knowledge, this work is one of the first to apply multilevel evolutionary games in networking context.

Structure

The chapter is structured as follows. We first provide in the next section the model with different second strategies set in each region and we develop the *hybrid* mean dynamics. After that we study the power control game in multi-cell CDMA and we give some numerical investigations. We give some convergence results in potential population games and *stable population games* (211) with constraints. We then describe a Stackelberg approach in population games with choice constraints.

---

[1]The classical Stackelberg games model rational players with asymmetric roles. In this paper, we don't necessary consider rational players. Players update their strategies by trial and error-detection and revising their protocols of strategies process (learning).

4.2 The Hybrid Model

We consider a population game model with *multicomponent strategies*. The population game consists of

- A large number of players.

- Each member of the population has a multicomponent strategy: first strategy (region) and secondary strategy.

- The players with the same first strategy have to choose their secondary strategies in a finite set and compete in a local non-cooperative game with either a finite random number of interacting players or with all the players in the region. The set of strategies available to players in all the population is $\mathcal{S}$ and has typical elements $(r, a), (\bar{r}, b)$. Let $\mathcal{A}^r = \{p_r^1, \ldots, p_r^{n^r}\}$ the pure secondary strategies set of the region r and, $\mathcal{S}^r = \{(r, a), a \in \mathcal{A}^r\}$ the pure strategies of a $r-$player. Then $\mathcal{S} = \cup_r \mathcal{S}^r$. Denote by l^r the mass of the region r, l the total mass of the population,

$$\mathcal{X}^r = \{ x^r \in \mathbb{R}^{|\mathcal{A}^r|}, x_a^r \geq 0, \sum_a x_a^r = l^r \}$$

the set of mixed secondary strategies in the region r and,

$$\mathcal{X} = \{ x \in \mathbb{R}^{\sum_r |\mathcal{A}^r|}, x_a^r \geq 0, \sum_r \sum_a x_a^r = \sum_r l^r = l \}$$

the set of all mixed strategies. Without loss of generality, we normalize the total mass l of the population to one. We call *state of the population* an element of $\mathcal{X}$ and, *state of the region r*, an element of $\mathcal{X}^r$.

- Each player from each local interaction receives some payoff. By $F_a^r(x)$ we denote the expected payoff (reward minus loss) of the second strategy $a \in \mathcal{A}^r$ in the region r when its state is x^r and the state of the other regions is x^{-r}. If every user takes a part in some region in a non-cooperative game between $K^r + 1$ users where K^r is a random variable over $\{0, 1, 2, \ldots\}$, then the payoff $F_a^r(x)$ can be expressed as

$$F_a^r(x) = \sum_{k \geq 0} \mathbb{P}(K^r = k) F_a^{r,k}(x)$$

where $F_a^{r,k}(x)$ is the expected payoff of the strategy a obtained in the local interaction between k users in the region r when the state is x. We denote by $F : \mathcal{X} \to \mathbb{R}^{\sum_r |\mathcal{A}^r|}$, the payoff function of all the population.

In addition to the study of the equilibrium of the game we shall consider also some evolutionary dynamics and study its convergence properties. We shall assume that players revise their strategies and use the strategies with higher payoffs. The system evolves under some evolutionary game dynamic process that describes the change of strategies (incoming flow and outgoing flow) in the population.

4.2.1 Global Nash Equilibrium

We say that $x = (x_a^r)_{r,a}$ is a Global Nash equilibrium (GNE) if for all deviation multi-strategy $mut = (mut_a^r)_{r,a}$,

$$\sum_r \sum_a (x_a^r - mut_a^r) F_a^r(x) \geq 0. \tag{4.1}$$

Wardrop equilibrium When the number of opponents is very large (possibly infinity), the corresponding Nash equilibrium is sometimes referred to Wardrop equilibrium. A state x is a Wardrop equilibrium if for any region r, all strategies being used by the members of the region r yield the same payoff, the payoff that would be obtained by members which chose the region r is lower for all strategies not used by players with the first action r.

$$\begin{cases} \forall r, \ \forall a, \ x_a^r > 0 \implies F_a^r(x) = \bar{F}^r = constant, \\ x_a^r = 0 \implies F_a^r(x) \leq F_{a'}^{r'}, \ \forall a', r' \mid x_{a'}^{r'} > 0 \end{cases}$$

For the existence of such equilibria, we can transform to existence of solution the variational inequality (4.1) which is guaranteed if the function $F_a^r(.)$ is continuous.

The relation between the two definitions is given by the following equivalences (see also the interaction between several populations model in (168)):

$$x \text{ is GNE} \quad \Longleftrightarrow \quad x \in \arg\max_{mut} \left\{ \sum_r \sum_a mut_a^r F_a^r(x) \right\} \tag{4.2}$$

$$\Longleftrightarrow \quad [x_a^r > 0 \implies F_a^r(x) \geq F_{a'}^{r'}(x), \ \forall r', a'] \tag{4.3}$$

4.2.2 Global evolutionarily stable state (ESS)

We use here the notion of ESS defined by Taylor in (197) in a multi-subpopulation game models but in our case the fitness of the region r can be independent of the state of the other regions, and users migrate to the regions with higher fitnesses. We say that $x = (x_a^r)_{r,a}$ is a *global* evolutionarily stable state if for each deviation multi-strategy called "mutations" $mut = (mut_a^r)_{r,a} \neq x$, there exists some $\epsilon_{mut} > 0$ such that $\forall \epsilon \in (0, \epsilon_{mut})$,

$$\sum_{r,a} (x_a^r - mut_a^r) F_a^r(\epsilon \ mut + (1 - \epsilon)x) > 0. \tag{4.4}$$

If the inequality (4.4) is non-strict the corresponding equilibrium is called *Neutrally Stable Strategy*.

4.2.3 Choice Constrained Equilibrium

A strategy x is a *choice constrained equilibrium* (CCE) if for all (r, a) such that $x_a^r > 0$ one has

$$F_a^r(x) = \max_{\substack{\bar{r}, \\ b \in \mathcal{A}^{\bar{r}}}} F_b^{\bar{r}}(x). \tag{4.5}$$

Note that a global ESS is a *neutrally ESS* which is a GNE. But a GNE can not be a global ESS. A CCE is a collection of constrained local Nash equilibria.

4.2.4 General Game Dynamics with Migration

We study a game-theoretic dynamics of a large population of users with *migration*. Users move to a region in which the average payoff of their power control strategy is higher. If user from region $\bar{r}$ chooses a base station from region r we will say that *the user migrate to region r*. Migration between regions could help the population to evolve towards a stationary point which is related to a Wardrop equilibrium (233; 179) or an evolutionary stable state.

Strategies with migration as replicators

We introduce here the replicator dynamics with migration which describes the evolution in the population of the various strategies. In the replicator dynamics, the share of a strategy in the population grows at a rate equal to the difference between the payoff of that strategy and the average payoff of the population under the migration constraints. More precisely, let (x_t) be the state of the population at time t. Thus we have $\sum_r \sum_a x^r_{a,t} = 1$ and $x^r_{a,t} \geq 0$ where $x^r_{a,t}$ represents the fraction of players playing a strategy (r,a) in period t.

We describe by approximating from stochastic influence (198) on the change in frequency of actions, the replication dynamics. We will describe more general class of evolutionary game dynamics in next subsection.

Suppose that in every period Δt, each player learns with probability $\alpha \Delta t > 0$ the expected payoff to the other opponent players and changes to the other's strategy if he perceives that the other's payoff is higher. However the information about the difference in the expected payoffs of the strategies is imperfect, so the larger the difference in the payoffs, the more likely the player is to perceive it, and change. Specially, we assume that the probability that a player using (r,a) will shift to $(\bar{r},b)$ in some neighboring set $\mathcal{N}_{(r,a)}$ is given by

$$x^r_{a,t+\Delta t} = \begin{cases} \mu^r[F^r_{a,t}(x_t) - F^{\bar{r}}_{b,t}(x_t)] & \text{if } F^r_{a,t}(x_t) > F^{\bar{r}}_{b,t}(x_t) \\ 0 & \text{if } F^r_{a,t}(x_t) \leq F^{\bar{r}}_{b,t}(x_t) \end{cases}$$

where μ is sufficiently small so that $x^r_{a,t} \leq 1$ holds $\forall\ (\bar{r},b), (r,a)$. The expected fraction $\mathbb{E}x^r_{a,t+\Delta t}$ of the population using (r,a) in period $t + \Delta t$ is given by

$$x^r_{a,t} - \alpha \Delta t x^r_{a,t} \sum_{(\bar{r},b) \in \mathcal{I}} x^{\bar{r}}_{b,t} \mu^r [F^{\bar{r}}_{b,t}(x_t) - F^r_{a,t}(x_t)] + \sum_{(\bar{r},b) \in \mathcal{N}_{(r,a)} \setminus \mathcal{I}} \alpha \Delta t x^r_{a,t} x^{\bar{r}}_{b,t} \mu^r [F^r_{a,t}(x_t) - F^{\bar{r}}_{b,t}(x_t)]$$

$$= x^r_{a,t} + \alpha \mu^r \Delta t x^r_{a,t} \left[F^r_{a,t}(x_t) - \sum_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} x^{\bar{r}}_{b,t} F^{\bar{r}}_{b,t}(x_t) \right] \qquad (4.6)$$

where

$$\mathcal{I} = \left\{ (\bar{r},b) \in \mathcal{N}_{(r,a)},\ F^{\bar{r}}_{b,t}(x_t) > F^r_{a,t}(x_t) \right\}$$

For large population, we can replace $\mathbb{E}x^r_{a,t+\Delta t}$ by $x^r_{a,t+\Delta t}$. Taking the limit of

$$\frac{x^r_{a,t+\Delta t} - x^r_{a,t}}{\Delta t}$$

when Δt goes to zero, we then obtain

> Continuous time: $\quad \dfrac{d}{dt} x^r_{a,t} = \alpha \mu^r G_t(x_t)$

where

$$G_t = x^r_{a,t} \left[F^r_{a,t}(x_t) - \sum_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} x^{\bar{r}}_{b,t} F^{\bar{r}}_{b,t}(x_t) \right].$$

The constant $\mu^r \alpha$ changes the rate of adjustment to stationary point. In the classical replicator these parameters are fixed to one. The two parameters α and μ give us a framework for controlling game dynamics (changing or upgrading policy) through the choice of a gain parameter governing the replicator dynamics. The discrete time version of the dynamics is then given by

> Discrete time: $\quad x^r_{a,t+\Delta t} = x^r_{a,t} + \alpha \mu^r G_t(x_t)\Delta t$

Decomposable dynamics for hybrid power control games

We now describe more general decomposable dynamics with migration respectively to apply in hybrid power control games.

We assume that during any time-step Δt, each individual among a fraction $\kappa^r \Delta t$ of the region r takes part in some local interaction and receives payoffs. We allow a fraction of users to migrate to a region in which their strategies have higher payoffs. The flow in the region r is specified in terms of some functions $\rho^r_{\bar{a},a}$ which determines the rates at which a player who is considering a change in strategies opts to switch to his various alternatives in the region r, and some function $\eta^{(r,a)}_{(\bar{r},b)}$ which determine the rates at which a player who is considering a change in strategies opts to switch to his various alternatives from other regions into region r. The functions ρ and η are called *revision protocols*. The two revision protocols depend on the state of the population and the payoff functions. The inflow inside the region r is $k^r \sum_{\bar{a} \in \mathcal{A}^r} x^r_{\bar{a}} \rho^r_{\bar{a},a}$ and outflow inside the region r, $k^r x^r_a \sum_{\bar{a} \in \mathcal{A}^r} \rho^r_{a,\bar{a}}$. We denote by M^r_a the function representing increase in the number of users inside of region r due to higher fitness. M^r_a is the difference between the *intra-inflow* and the *intra-outflow*.

$$M^r_a(x^r) := k^r \left[\sum_{\bar{a} \in \mathcal{A}^r} x^r_{\bar{a}} \rho^r_{\bar{a},a} - x^r_a \sum_{\bar{a} \in \mathcal{A}^r} \rho^r_{a,\bar{a}} \right] \tag{4.7}$$

The *inter-inflow* of the region r is $k^r \gamma^r \sum_{(\bar{r},b)} x^{\bar{r}}_b \eta^{(\bar{r},b)}_{(r,a)}$ and *inter-outflow* of the region r is $\gamma^r k^r x^r_a \sum_{(\bar{r},b)} \eta^{(r,a)}_{(\bar{r},b)}$ where γ^r is a parameter which represents the migration rate of region r, and k^r is a growth parameter. By ϕ^r_a we denote a function which describes changes of the numbers of users playing the pure secondary strategy a in the region r due to migration from other region to the region r.

$$\phi^r_a(x^r, x^{-r}) = k^r \gamma^r \left[\sum_{(\bar{r},b)} x^{\bar{r}}_b \eta^{(\bar{r},b)}_{(r,a)} - x^r_a \sum_{(\bar{r},b)} \eta^{(r,a)}_{(\bar{r},b)} \right] \tag{4.8}$$

By combining the two equations (4.7), (4.8) we obtain the continuous time mean dynamics with migration or mean dynamics with *multicomponent strategies*.

$$\frac{d}{dt} x^r_a(t) := V^{r,a}_F(x) = M^r_a(x^r(t)) + \phi^r_a(x(t)) \tag{4.9}$$

We denote by β the following revision protocol

$$\beta^{(\bar{r},b)}_{(r,a)} = \begin{cases} k^r \rho^r_{ba} & \text{if } \bar{r} = r \\ k^r \gamma^r \eta^{(\bar{r},b)}_{(r,a)} & \text{if } \bar{r} \neq r \end{cases}$$

Some example of dynamics with multicomponent strategies

Here we give some class of dynamics with migration. We denote by $[d]_+$ the maximum between d and 0.

- Replicator dynamics

$$\rho^r_{\bar{a},a} = x^r_a [F^r_a - F^r_{\bar{a}}]_+, \quad \eta^{(\bar{r},b)}_{(r,a)} = x^r_a [F^r_a - F^{\bar{r}}_b]_+$$

- Smith dynamics

$$\rho^r_{\bar{a},a} = [F^r_a - F^r_{\bar{a}}]_+, \; \eta^{(\bar{r},b)}_{(r,a)} = [F^r_a - F^{\bar{r}}_b]_+$$

- BNN dynamics

$$\rho^r_{\bar{a},a} = [F^r_a - \sum_{\bar{a}} F^r_{\bar{a}} x^r_{\bar{a}}]_+, \eta^{(\bar{r},b)}_{(r,a)} = [F^r_a - \sum_{(\bar{r},b)} F^{\bar{r}}_b x^{\bar{r}}_b]_+$$

- Combined Smith and BNN dynamics (in this order)

$$\rho^r_{\bar{a},a} = [F^r_a - F^r_{\bar{a}}]_+, \eta^{(\bar{r},b)}_{(r,a)} = [F^r_a - \sum_{(\bar{r},b)} F^{\bar{r}}_b x^{\bar{r}}_b]_+$$

- Combined BNN and Smith dynamics

$$\rho^r_{\bar{a},a} = [F^r_a - \sum_{\bar{a}} F^r_{\bar{a}} x^r_{\bar{a}}]_+, \eta^{(\bar{r},b)}_{(r,a)} = [F^r_a - F^{\bar{r}}_b]_+$$

Note that the order of combination of the revision protocols is non-commutative. We next define number of properties that classify evolutionary game dynamics with migration constraints.

Constrained Nash Stationarity (CNS) property All rest points of the mean dynamic (4.9) are precisely the CCEs of the game being played i.e $(M + \phi)(x) = 0$ if and only if x is CCE.

The replicator dynamics is one of the most studied dynamics in evolutionary game theory to describe evolution of the frequencies in the population but it is known that the replicator dynamics may not lead to equilibria (see (231; 168)). To guarantee *constrained Nash stationary* (CNS) properties, the mean dynamics must satisfy some conditions.

Positive Correlation (PC) (168)

$$\sum_{r,a} F^r_a(x)(M^r_a(x^r) + \phi^r_a(x^r, x^{-r})) \geq 0.$$

Positive correlation guarantees that every Nash equilibrium of the game is a stationary point of the dynamics (4.9) as shown in Result 4.3.3.

To see why this condition is so named, observe that by the condition $\sum_{r,a} M^r_a(x^r) + \phi^r_a(x^r, x^{-r}) = 0$,

$$
\begin{aligned}
\sum_{r,a} F^r_a(x)(M^r_a(x^r) + \phi^r_a(x^r, x^{-r})) &= \sum_{r,a} \left[F^r_a(x) - \frac{1}{l^r} \sum_{(\bar{r},b)} x^{\bar{r}}_b F^{\bar{r}}_b(x) \right] [M^r_a(x^r) + \phi^r_a(x^r, x^{-r}) - 0] \\
&= \sum_r l^r Cov(M^r(x^r) + \phi^r(x), F^r(x))
\end{aligned}
$$

where *Cov* denotes the covariance between strategy growth rates and payoffs in region r. Hence, condition (PC) holds if there is a positive correlation between growth rates and payoffs in each region.

Monotonicity condition: sign preserving (SP) Monotonicity condition on ρ and η defines the following class of dynamics. The couple of functions (ρ, η) preserves the sign if for all region r

$$\rho^r_{\bar{a},a}(x) \begin{cases} > 0 & \text{if } \bar{a}, a \in \mathcal{A}^r \text{ and } F^r_a(x) > F^r_{\bar{a}}(x) \\ = 0 & \text{otherwise} \end{cases},$$

$$\eta_{r,a}^{(\bar{r},b)}(x) \begin{cases} > 0 & \text{if } a \in \mathcal{A}^r, b \in \mathcal{A}^{\bar{r}}, F_a^r(x) > F_b^{\bar{r}}(x) \\ = 0 & \text{otherwise} \end{cases}$$

The sign preserving property say that the inflow rate from the strategy a to $\bar{a}$ inside the region r is positive if and only if the payoff to a exceeds the payoff to $\bar{a}$ and the inflow rate from other regions to the region r is positive for a given strategy a if and only the payoff to r exceeds the payoff to $\bar{r}$.

4.3 Equilibrium and rest point

In (168), Sandholm showed that in absence of migration (par example if $\gamma^r = 0$, $\forall\, r$), replicator dynamics (or general imitation dynamics), Smith dynamics (pairwise comparison dynamics), projection dynamics, Brown-von Neumann-Nash dynamics (excess payoff dynamics) satisfy the positive correlation property. Moreover Brown-von Neumann-Nash dynamics and Smith dynamics satisfy (CNS) property. Here, we extend these results to evolutionary game dynamics with migrations.

Theorem 4.3.1. *Suppose that ρ and η generate one of the following dynamics: replicator, Smith, BNN dynamics. Then, the resulting dynamics with multicomponent (ρ, η) in (4.9) is (PC).*

Theorem 4.3.2. *Suppose that the functions ρ and η satisfy (SP). Then, The multicomponent dynamics (4.9) satisfies (PC). In particular, Smith dynamics is (PC).*

Proof. Sign preserving property implies that $\rho_{\bar{a},a}^r[F_a^r - F_{\bar{a}}^r] \geq 0$ and $\eta_{(r,a)}^{(\bar{r},b)}[F_a^r - F_b^{\bar{r}}] \geq 0$.

$$\sum_{(r,a)} \dot{x}_a^r F_a^r = \sum_{r,a,\bar{a}} k^r x_{\bar{a}}^r \underbrace{\rho_{\bar{a},a}^r[F_a^r - F_{\bar{a}}^r]}_{\geq 0} + \sum_{(r,a),(\bar{r},b)} k^r \gamma^r x_b^{\bar{r}} \underbrace{\eta_{(r,a)}^{(\bar{r},b)}[F_a^r - F_b^{\bar{r}}]}_{\geq 0} \tag{4.10}$$

$\square$

Theorem 4.3.3. *If V_F is positively correlated then x is CCE implies that x is a stationary point of mean dynamics (4.9).*

Proof. x is a CCE is equivalent to $\forall z$, such that $\sum_{(r,a)} z_a^r = 0$, one has $\sum_{(r,a)} F_a^r(x) z_a^r \leq 0$. Now fix a region r, and define the vector z as follows

$$z_b^{\bar{r}} = \begin{cases} V_F^{\bar{r},b}(x) & \text{if } \bar{r} = r, \ b \in \mathcal{A}^r \\ 0 & \text{if } b \notin \mathcal{A}^r \text{ or } \bar{r} \neq r \end{cases}$$

Then

$$\sum_{r,a} z_a^r = \sum_{\substack{(r,a)\\(\bar{r},b)}} x_b^{\bar{r}} \beta_{(r,a)}^{(\bar{r},b)} - \sum_{\substack{(r,a)\\(\bar{r},b)}} x_a^r \beta_{(\bar{r},b)}^{(r,a)} = 0 \tag{4.11}$$

So,

$$\sum_{(r,a)} V_F^{r,a}(x) F_a^r(x) = \sum_{(r,a)} z_a^r F_a^r(x) \leq 0$$

By (PC), this inequality implies that $V_F^{r,a}(x) = 0$.

$\square$

Theorem 4.3.4 (Characterization of stationary points)**.** *Suppose that the functions ρ and η satisfy (SP). Let x be a stationary point of the mean dynamics (4.9). Then x is a CCE.*

$$x_a^r > 0 \Rightarrow F_a^r(x) = \max_{\substack{\bar{r}, \\ b \in \mathcal{A}^{\bar{r}}}} F_b^{\bar{r}}(x) \tag{4.12}$$

Proof. We show that every stationary point of the mean dynamics is a CCE.

- x is a CCE if and only for all (r, a), $x_a^r = 0$ or there is no outflow from strategy (r, a) :
  $$\sum_{(\bar{r}, b)} \beta_{(\bar{r}, b)}^{(r, a)} = 0$$

- Suppose that $V_F(x) = 0$. If $(\bar{r}, b)$ is an optimal strategy then sign preserving assumption implies that there is no outflow from $(\bar{r}, b)$ i.e $x_b^{\bar{r}} \sum_{(r, a)} \beta_{(r, a)}^{(\bar{r}, b)} = 0$. Since $V_F(x) = 0$, one has

  $\sum_{(r, a)} x_a^r \beta_{(\bar{r}, b)}^{(r, a)} = 0$. This condition is exactly

  $$\forall (r, a), \ x_a^r = 0 \text{ or } \beta_{(\bar{r}, b)}^{(r, a)} = 0, \ \forall (\bar{r}, b), \ b \in \mathcal{A}^{\bar{r}}.$$

  This last alternative says that $F_a^r(x) \geq F_b^{\bar{r}}(x), \ \forall (\bar{r}, b), \ b \in \mathcal{A}^{\bar{r}}$.

This completes the proof. $\qquad\square$

4.4 Class of Games with Multicomponents Strategies

4.4.1 Stable Population Games

We say that F is a stable game if for all $x, y \in \mathcal{X}$,

$$\sum_r \sum_a (x_a^r - y_a^r)(F_a^r(x) - F_a^r(y)) \leq 0.$$

Proposition 4.4.1.1. *In stable population games, the set of GNE is convex and coincides with the set of neutrally ESS.*

Proof. Suppose that F satisfies

$$x, y \in \mathcal{X}, \sum_{r,a} (x_a^r - y_a^r)(F_a^r(x) - F_a^r(y)) \leq 0$$

and let x be a GNE. It is easy to see that every neutrally ESS is a GNE. Now we show that x is a neutrally ESS. So fix, an arbitrary vector y. Since F is stable and x is GNE, one has the system

$$\begin{cases} \sum_{r,a} (y_a^r - x_a^r)(F_a^r(y) - F_a^r(x)) \leq 0 \\ \sum_{r,a} (y_a^r - x_a^r) F_a^r(x) \leq 0 \end{cases}$$

Adding the two inequalities of the last system, we obtain that

$$\sum_{r,a} (y_a^r - x_a^r) F_a^r(y) \leq 0.$$

Taking $y = (1 - \epsilon)x + \epsilon\, mut$ for arbitrary $mut \neq x$, we conclude that x is a neutrally ESS. To prove the convexity, we rewrite the GNE set as an intersection of convex sets $\{GNE\} = \bigcap_y A_y$ where $A_y = \{x \in \mathcal{X}, \sum_{r,a}(y_a^r - x_a^r)F_a^r(y) \leq 0\}$. $\qquad\square$

Proposition 4.4.1.2. *Suppose that the revision protocol has the form*

$$\beta_{(r,a)}^{(\bar{r},b)}(x) = \begin{cases} \xi_a^r(F_a^r(x) - F_b^{\bar{r}}(x)) & \text{if } b \in \mathcal{A}^{\bar{r}}, a \in \mathcal{A}^r \\ 0 & \text{otherwise} \end{cases}$$

for some functions $\xi_a^r : \mathbb{R} \to \mathbb{R}_+$ then the set of CCE is globally asymptotically stable in stable games.

Proof. Let the function $B : \mathcal{X} \to \mathbb{R}_+$ be defined by

$$B(x) = \sum_{\substack{(r,a),\, a\in\mathcal{A}^r \\ (\bar{r},b),\, b\in\mathcal{A}^{\bar{r}}}} x_a^r \int_0^{-F_a^r(x)+F_b^{\bar{r}}(x)} \xi_b^{\bar{r}}(\theta)\, d\theta.$$

The function B has the following properties: The set $\{x,\ B(x) = 0\}$ is exactly the set of CCE.

(b) $\dfrac{\partial}{\partial x_{a'}^{r'}} B = \sum_{(\bar{r},b)} \int_0^{-F_{a'}^{r'}+F_b^{\bar{r}}} \xi_b^{\bar{r}}(\theta)\, d\theta + \sum_{(r,a)} \dot{x}_a^r \dfrac{\partial F_a^r}{\partial x_{a'}^{r'}}$

(c) $\dfrac{d}{dt} B(x(t)) = \sum_{(r',a')} \dfrac{\partial B}{\partial x_{a'}^{r'}} \dot{x}_{a'}^{r'} = \langle \dot{x}, DF(x)\dot{x}\rangle + \sum_{\substack{(r',a') \\ (r,a)}} \beta_{(r',a')}^{(r,a)} \sum_{(\bar{r},b)} \left[\int_0^{-F_{a'}^{r'}+F_b^{\bar{r}}} \xi_b^{\bar{r}}(\theta)\, d\theta - \int_0^{-F_a^r+F_b^{\bar{r}}} \xi_b^{\bar{r}}(\theta)\, d\theta \right]$

Since F is a stable population game, one has $\langle \dot{x}, DF(x)\dot{x}\rangle \leq 0$. In other hand, $\beta_{(r',a')}^{(r,a)} > 0 \Leftrightarrow$ $F_{a'}^{r'} > F_a^r$. Then, for any action $(\bar{r}, b)$, one has $F_b^{\bar{r}} - F_{a'}^{r'} \leq F_b^{\bar{r}} - F_a^r$. Hence, the term $\left[\int_0^{-F_{a'}^{r'}+F_b^{\bar{r}}} \xi_b^{\bar{r}}(\theta)\, d\theta - \int_0^{-F_a^r+F_b^{\bar{r}}} \xi_b^{\bar{r}}(\theta)\right.$ is negative. We conclude that $\dfrac{d}{dt} B(x) \leq 0$ and $\dfrac{d}{dt} B(x) = 0$ if and only if $x_a^r \beta_{(r',a')}^{(r,a)} = 0, \forall (r,a), (r', a')$ i.e $V_F(x) = 0$. Thus, the function satisfies Lyapunov stability criterion for the set of stationary point of V_F. $\qquad\square$

4.4.2 Potential Population Games

We say that F is a full potential game if it exists a C^1 function $f : \mathcal{X} \to \mathbb{R}$ such that

$$\frac{\partial}{\partial x_a^r} f = F_a^r(x).$$

Two strategies: if the game has only two strategies in each r, all strategy distributions lie on a line. Given any continuous payoff functions

$$F_1^r, F_2^r : \mathcal{X}^r = \{(x_1^r, x_2^r),\ x_i^r \geq 0,\ x_1^r + x_2^r = l^r\} \to \mathbb{R},$$

a potential function is given by

$$f(x) = \sum_{r=1}^N \int_0^{x_1^r} (F_1^r(t, l^r - t) - F_2^r(t, l^r - t))\, dt$$

Proposition 4.4.2.1. *Global convergence holds in potential games under (PC).*

Proof. Suppose that f is a potential function of the game and the dynamic $\dot{x} = V_F(x)$ is (PC). Then f is a Lyapunov function of this dynamic.

$$\frac{d}{dt}f(x) = \sum_r \sum_{a \in \mathcal{A}^r} V_F^{r,a}(x) F_a^r(x).$$

Moreover f satisfies $\frac{d}{dt}f(x) \geq 0$ with equally if and only if x is a CCE. It is known that if a dynamic admits a such Lyapunov function, all solution trajectories of the dynamic converge to an equilibrium. Combining with results 4.3.4 and (PC) we obtain the announced result. $\quad\square$

4.4.3 Migration with constraints

In this subsection, we assume that when updating a strategy one can change only one component of the strategies at the same time. Players with the action (r, a) can use all actions in $\mathcal{N}_{(r,a)} = \cup_{\bar{r} \neq r,(}\{(\bar{r},a), \ a \in \mathcal{A}^{\bar{r}}\} \cup \cup_{\bar{a} \in \mathcal{A}^r}\{(r,\bar{a})\}$. Taking these considerations, a strategy x is a CCE if for all (r, a) such that $x_a^r > 0$ one has

$$F_a^r(x) = \max_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} F_b^{\bar{r}}(x). \tag{4.13}$$

and the constrained dynamic becomes

$$\dot{x}_a^r = \sum_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} x_b^{\bar{r}} \beta_{(r,a)}^{(\bar{r},b)} - x_a^r \sum_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} \beta_{(\bar{r},b)}^{(r,a)}$$

$$= \sum_{\bar{r}} x_b^{\bar{r}} \eta_{(r,a)}^{(\bar{r},a)} - x_a^r \sum_{\bar{r}} \eta_{(\bar{r},a)}^{(r,a)} + \sum_{\bar{a} \in \mathcal{A}^r} x_{\bar{a}}^r \rho_{\bar{a},a}^r - x_a^r \sum_{\bar{a} \in \mathcal{A}^r} \rho_{a,\bar{a}}^r$$

Remarks 4.4.3.1. *We denote by*

$$C_{(r,a)} = \mathcal{S} \backslash \mathcal{N}_{(r,a)} = \left[\cup_{\bar{r}} \left(\{\bar{r}\} \times \mathcal{A}^{\bar{r}} \right) \right] \backslash \mathcal{N}_{(r,a)}$$

1. *Let $y \in \mathcal{X}$ be a CCE of the game with migration constraints. If y satisfies*

$$\max_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} F_b^{\bar{r}}(y) \geq \max_{(\bar{r},b) \in C_{(r,a)}} F_b^{\bar{r}}(y)$$

for all (r, a) such that $y_a^r > 0$ then y is also a CCE in the game without migration constraints. This kind of phenomena occurs in the well known Braess Paradox problem (53): increasing the number of regions (increasing the area of coverage, adding new resources such as deployment of new base stations) does not necessary lead to higher equilibria payoffs. If a CCE x of the unconstrained migration game satisfies

$$\max_{(\bar{r},b) \in C_{(r,a)}} F_b^{\bar{r}}(x) \leq F_a^r(x)$$

then x is a Braess steady state. Given such a steady state x, if there are mutations, mutations cannot spread due to the migration constraints. Hence, the mutants' strategy can be invaded strictly at the steady state x.

2. *Let x be a CCE of the game without migration constraints. If x satisfies $F_a^r(x) = \max_{(\bar{r},b) \in \mathcal{N}_{(r,a)}} F_b^{\bar{r}}(x)$*

then x is also a CCE in the game with migration constraints.

3. *Let Opt_1 (resp. Opt_2) the set of CCEs with the maximum payoff in the unconstrained migration game (constrained migration game.) Then,*

$$\forall \; x \in Opt_1, y \in Opt_2, \; x_a^r > 0 \Longrightarrow F_a^r(x) \geq F_a^r(y).$$

4. *Suppose that the unconstrained migration game has a unique CCE. Let x (resp. y) be a CCE for the game without migration constraints (resp. the constrained migration game.) Then for every action (r, a) such that $x_a^r > 0$, one has $F_a^r(x) \geq F_a^r(y)$.*

4.4.4 Inverse problem: reachable regions of a power level

In this subsection, we model the power levels as first strategies and regions as secondary strategies. Let $\mathcal{P}$ a finite set of power levels, $\mathcal{A}_p = \{r_p^1, \ldots, r_p^{n^p}\}$ the pure secondary strategies set of the power level $p \in \mathcal{P}$ and, $\mathcal{S}_p = \{(r, a), \; r \in \mathcal{A}_p\}$ the pure strategies of a $p-$player. Then $\mathcal{S} = \cup_{p \in \mathcal{P}} \mathcal{S}_p$. Given an energy (power level), the player will migrate from its location to a reachable region of this power level.

A strategy x is a CCE of this inverse problem if for all (r, a) such that $x_a^r > 0$ one has

$$F_a^r(x) = \max_{\substack{b \in \mathcal{P}, \\ \bar{r} \in \mathcal{A}_b}} F_b^{\bar{r}}(x), \tag{4.14}$$

and mean dynamic becomes

$$\dot{x}_a^r \; = \; \sum_{b \in \mathcal{P}} \sum_{\bar{r} \in \mathcal{A}_b} x_b^{\bar{r}} \beta_{(r,a)}^{(\bar{r},b)} - x_a^r \sum_{b \in \mathcal{P}} \sum_{\bar{r} \in \mathcal{A}_b} \beta_{(\bar{r},b)}^{(r,a)}$$

4.5 Global Optimization

In this section, we focus on global optimization of the payoff of all the population. The total payoff function is defined as

$$F_{total} : \; \mathcal{X} \longrightarrow \mathbb{R}$$
$$x \mapsto \sum_r \sum_{a \in \mathcal{A}^r} x_a^r F_a^r(x^r, x^{-r}).$$

The optimization problem is given by

$$\max_x \quad \sum_r \sum_{a \in \mathcal{A}^r} x_a^r F_a^r(x)$$
$$\text{subject to}$$
$$\forall \; r, \; \sum_r \sum_{a \in \mathcal{A}^r} x_a^r = l^r$$
$$\forall (r, a) \in \mathcal{S}, \; \; x_a^r \geq 0$$

The Lagrangian for this maximization problem is

$$L(x, \lambda, \mu) = \sum_{r,a} x_a^r F_a^r(x) - \sum_r \lambda^r \left(\sum_a x_a^r - l^r \right) + \sum_{r,a} x_a^r \mu_a^r$$

Thus, the Karush-Kuhn-Tucker(KKT) first order necessary conditions for maximization are $\forall r, \forall a \in \mathcal{A}^r$,

$$\frac{\partial}{\partial x_a^r} F_{total}(x) = \lambda^r - \mu_a^r, \; \mu_a^r \geq 0, \; \mu_a^r x_a^r = 0 \tag{4.15}$$

107

4.6   Hybrid power control in OFDMA-based IEEE802.16 network

OFDMA (Orthogonal Frequency Division Multiple Access) is recognized as one of the most promising multiple access technique in wireless communication system. This technique is used to improve spectral efficiency and becomes an attractive multiple access technique for 4th generation mobile communication system as WiMAX.

In OFDMA systems, each user occupies a subset of subcarriers, and each carrier is assigned exclusively to only one user at any time. This technique has the advantage of eliminating intra-cell interference (interference between subcarriers is negligible). Hence the transmission is affected by intercell interference since users in adjacent sectors may have also been assigned to the same carrier. If those users in the adjacent sectors transmitted with high power the intercell interference may severely limit the SINR achieved by the user. Some form of coordination between the different cells occupying the spectral resource are studied in (134; 123). The optimal resource allocation requires complete information about the network in order to decide which users in which cells should transmit simultaneously with a given power. All of these results however, rely on some form of centralized control to obtain gains at various layers of the communication stack. In a realistic network as WiMAX, centralized multicell coordination is hard to realize in practice, especially in fast-fading environments.

We consider an OFDMA system where radio resources are allocated to users on their channel measures and traffic requirements. Each carrier within a frame must be assigned to at most one user in the corresponding cell. In this way each carrier assignment can be made independently in each cell. Hence when a user is assigned to carrier, the mobile should determine the power transmission to the Base station. This power should take into account the interference experienced by the transmitted packet.

Power control on the uplink Consider the uplink of a multiple multicell system, employing the same spectral resource in each cell. Power control is used in an effort to preserve power and to limit interference and fading effects. For users located in a given cell, co-channel interference may therefore come from only few cells (neighbors). Since the intra-cell interference is negligible, we focus on the users which use a specific carrier.

Consider N cells, and a large number of the population of mobiles randomly distributed over each channel and each cell. Since in OFDMA systems, each carrier is assigned exclusively to only one mobile at any time, we assume that the interactions between mobiles are manifested through many local interactions between $K^r + 1$ mobiles where K^r is a random variable which represents the number of opponent mobiles in the set of neighbors of a cell r. At each slot, interaction occurs only between the mobiles which have been assigned to the same carrier. We assume that in this model that users can choose many base stations available in his neighborhood and a finite power levels as a secondary pure strategies .

Let $g_{r\bar{r}}$ denote the average channel gain from user in cell r to cell $\bar{r}$. Hence, if a user in cell r transmits with power p^r, the received signal strength at cell r is $p^r g_{rr}$, while the interference it occurs on cell $\bar{r}$ is $p^r g_{r\bar{r}}$. Hence, the interference experienced by cell r is given by

$$SINR^r_{K^r}(\mathbf{p}) = \frac{g_{rr}p^r}{\sigma_r^2 + \sum_{\substack{\bar{r} \in \mathcal{N}_r \setminus \{r\} \\ |\mathcal{N}_r| = K^r}} g_{r\bar{r}}p^{\bar{r}}},$$

where σ_r^2 is the power of the thermal noise experienced at cell r, $\mathcal{N}_r$ be the set of neighbors of a user in cell r. The rate achieved by user in cell r is given by $R^r(\mathbf{p}) = \mathcal{E}_{K^r} \log(1 + SINR^r_{K^r}(\mathbf{p}))$, where $\mathbf{p}$ denotes the power level vector of mobiles choice which are assigned to a specific carrier. We assume that the user's utility in cell r is given by $u^r(\mathbf{p}) = R^r(\mathbf{p}) - c^r(p)$. The above utility

represents the weighted difference between the throughput that can be achieved as expressed by Shannon's capacity and the power consumption cost. We show existence and uniqueness of the equilibrium point under some conditions on the cost functions. The results are given in next section in hybrid model of multicell CDMA system but the technique used there are valid for the above game.

4.7 A hybrid evolutionary game in multicell CDMA system

We consider a large population of mobile terminals and many distributed base stations in a multicell CDMA wireless network model. The system consists of N cells which are called "regions". The number of users which transmit to a base station is a random variable. Each mobile connects to a base station which it chooses from of the set of base stations $\{1, 2, \ldots, N\}$ with an uplink power level from the set $\mathcal{A}$. The action space of a mobile is given by $\{1, 2, \ldots, N\} \times \mathcal{A}$. The level of services a mobile receives is described in terms of signal-to-interference ratio (SIR). The SIR obtained by mobile m at a base station located in region r is given by

$$SIR_m^{r,k}(a^m, a^{-m}) = \frac{La^m h^{r,m}}{\sigma_r^2 + \sum_{0 \leq l \leq k, l \neq m} a^l h^{r,l}}$$

where a^{-m} denote the vector $(a^1, \ldots, a^{m-1}, a^{m+1}, \ldots, a^k)$ the power levels of the others mobiles, σ_r^2 is a constant which represents the variance of a noise power due to the factors other than the transmissions of other mobiles at the base station r. The term $a^m h^{r,m}$ represents the power level received at base station from mobile using the power level a^m and $L = W/B > 1$ is the spreading gain of the CDMA system, W is the chip rate and B is the data rate of the users.

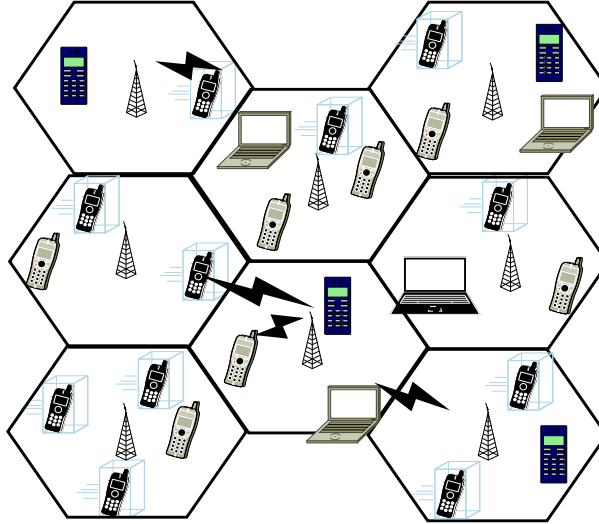
Figure 4.1: *The hybrid model*

The payoff that mobile using the power level a^m can send to the base station in region r at a given slot is given by

$$J^{r,k}(a^m, a^{-m}) = C^r \log\left(1 + SIR_m^{r,k}(a^m, a^{-m})\right) - c^r(a^m)$$

where C^r is the channel bandwidth of r and $c^r(.)$ is the cost function of region r. We assume that c^r is an increasing cost function. The expression of $J^{r,k} + c^r$ is known as Shannon capacity. The expected fitness of mobile m at the base station r when the state of the population at region r is x^r is given by

$$F_{a^m}^r(x^r) = \sum_k P(K^r = k) \sum_{a^{-m}} \left(\prod_{l \neq m} x_{a^l}^r \right) F^{r,k}(a^m, a^{-m}).$$

The total payoff of the population is

$$thp(x) := \sum_{r=1}^N \sum_{a^m \in \mathcal{S}} x_{a^m}^r F_{a^m}^r(x^r)$$

Proposition 4.7.0.1 (two power levels). *Let* $\mathcal{A} = \{0, P\}$ *and* $h^{r,m} = h^r$, $r = 1, 2$. *Let* $W_j^r = C^r \log(1 + \dfrac{h^r P}{\sigma_r^2 + jh^r P})$, $j = 0, \ldots, k.$, *and* $c^r(P)$ *positive reals satisfying* $c^r(P) \in (\sum_k p_k W_k^r, a_0^r)$ *where* $p_k = P(K = k)$. *Then, the polynomial*

$$Q^r(\xi) := -c^r(P) + \sum_k p_k \sum_{j=0}^k W_j^r \binom{k}{j} \xi^j (1 - \xi)^{k-j}$$

has unique root x_*^r *on the interval* $(0, 1)$.

Proposition 4.7.0.2. *The unique solution* $(x_*^r)_{r=1,\ldots,N}$ *given by the result 7.3.10 is a the unique interior Wardrop equilibrium.*

Proof. The fitness of the strategy P in the cell r is given by $F_P^r(x_P^r, 1 - x_P^r) = Q^r(x_P^r)$, the fitness of the strategy 0 is zero. Then the expected fitness of the cell r is $x_P^r Q^r(x_P^r)$. Hence, every interior Wardrop equilibrium $(x_*^r)_{r=1,\ldots,N}$ must satisfy

$$Q^r(x_P^r) = 0, \ \forall r.$$

From result 7.3.10, these equations have a unique solution on $(0, 1)$ under the conditions:

$$\forall r, \ Q^r(0) = -c^r(P) + W_0^r > 0,$$

$$Q^r(1) = -c^r(P) + \sum_k p_k W_k^r < 0$$

$\square$

Proposition 4.7.0.3. *The Wardrop equilibrium* $(x_*^r)_{r=1,\ldots,N}$ *is also a global ESS.*

Proof. For $x = x_*$ and every cell r, one has $(x_P^r - mut_P^r)F_P^r(\epsilon \, mut^r + (1 - \epsilon)x^r) > 0$ for small ϵ. Hence, for all r,

$$\sum_{a \in \mathcal{A}} x_a^r F_a^r(\epsilon \, mut^r + (1 - \epsilon)x^r) > \sum_{a \in \mathcal{A}} mut_a^r F_a^r(\epsilon \, mut^r + (1 - \epsilon)x^r).$$

We conclude that

$$\sum_r \sum_{a \in \mathcal{A}} x_a^r F_a^r(\epsilon \, mut^r + (1 - \epsilon)x^r) > \sum_r \sum_{a \in \mathcal{A}} mut_a^r F_a^r(\epsilon \, mut^r + (1 - \epsilon)x^r).$$

This completes the proof.

$\square$

4.8 Numerical investigation: Convergence to the equilibrium

Our numerical experiment studies the behavior of the replicator dynamics and smith dynamics with migration. We consider the following fixed parameters : we took $N = 2$ cells, $K \in \{0, 1, \ldots, 500\}$, $\mathcal{A} = \{0, P\}$, $h^{r,m} = h^r$, $r = 1, 2$, $\sigma_r^2 = r \times 10^{-3}$, $p_k = 2 \times 10^{-3}$, $k = 1, \ldots, 500$, $\gamma_1 = 0.5$, $\gamma_2 = 1$, $c^r(P) = r$, $P = 1$. An interior Wardrop equilibrium exists for these parameters, for which the fractions of the population using P is given by the result 4.7.0.2. The continuous replicator dynamics becomes

$$\frac{d}{dt}x_P^r(t) = k^r x_P^r (1 - x_P^r) Q^r(x_P^r) + \gamma^r k^r x_P^r [Q^r(x_P^r) - \sum_{\bar{r}} Q^{\bar{r}}(x_P^{\bar{r}})]$$

The continuous Smith dynamics becomes

$$\frac{d}{dt}x_P^r(t) = k^r \left((1 - x_P^r)[Q^r(x_P^r)]_+ - x_P^r[-Q^r(x_P^r)]_+ \right) +$$

$$\gamma^r k^r \sum_{\bar{r}} x_P^{\bar{r}} [Q^r(x_P^r) - Q^{\bar{r}}(x_P^{\bar{r}})]_+ - x_P^r \sum_{\bar{r}} [Q^{\bar{r}}(x_P^{\bar{r}}) - Q^r(x_P^r)]_+$$

The resulting trajectories of the population ratio each cell using the power level P, as a function of time, is given in Fig. 4.2, 4.3 and 4.4.
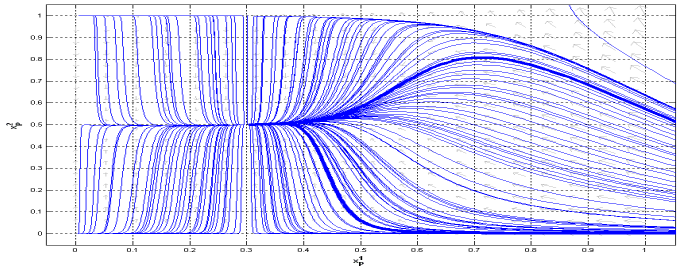
Figure 4.2: *RD: convergence to Wardrop equilibrium*

The throughput of transmitters, the total throughput and the total payoff of the whole population are represented respectively in Fig.4.5, 4.6 and 4.7. In Fig. 4.5, we can see that the throughput of players which use the strategy P decreases when the number of transmitters increases. This because the SINR decreases when the number of transmitters increases. The throughput increases with the power level P. When the number of transmitters is high, a player will increase its SINR, but will decrease the SINRs of the others players, and the total cost will decrease. Thus, the total payoff becomes negative as shown in Fig.4.7.

Notes

In the second part this manuscript we develop a new class of dynamic evolutionary games called stochastic population games which include Markov decision evolutionary games. This class is also related to the so-called anonymous sequential games. We examine both homogenous and heterogenous case and apply them to energy management and resource competition in wireless networks. The chapter is based to our publications in (217; 10; 212; 25).

Figure 4.3: RD: fraction of mobiles using the power levels P in cell 1 and cell 2

Figure 4.4: Smith dynamics

Figure 4.5: *The expected throughput of the transmitters.*

Figure 4.6: *The expected throughput of the population vs x_P.*
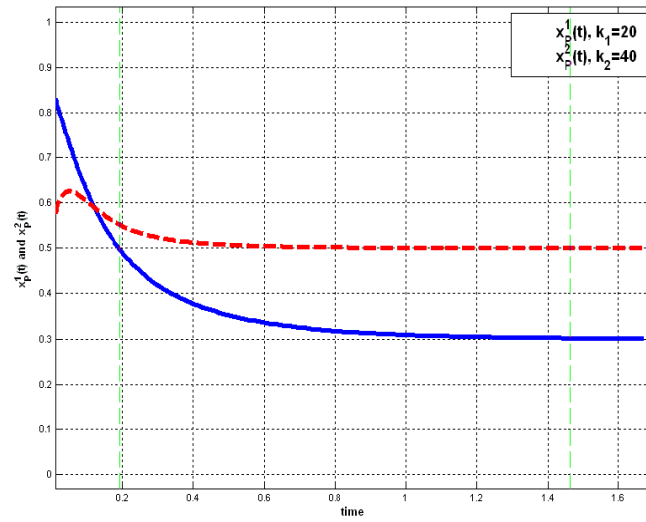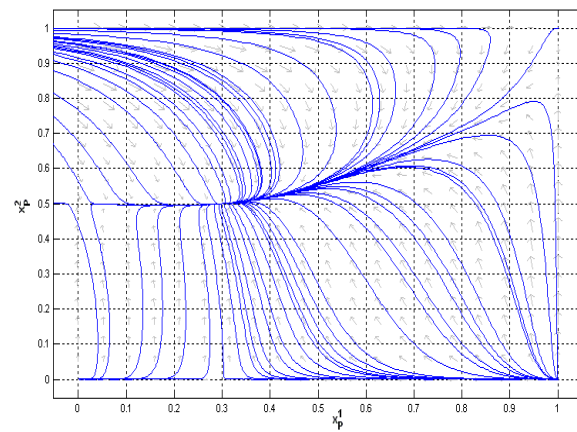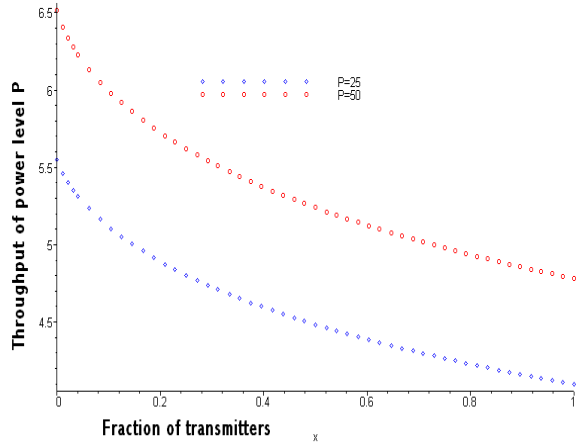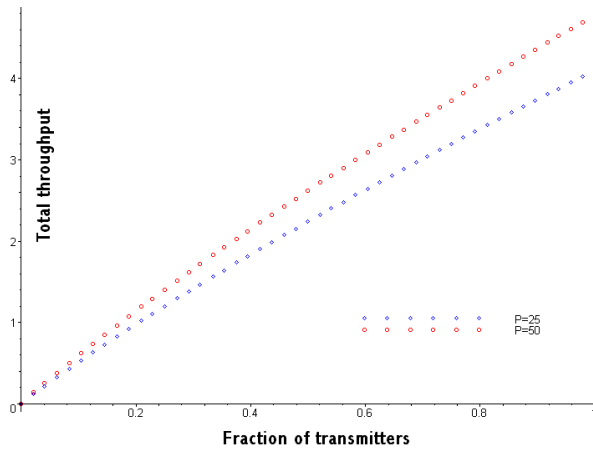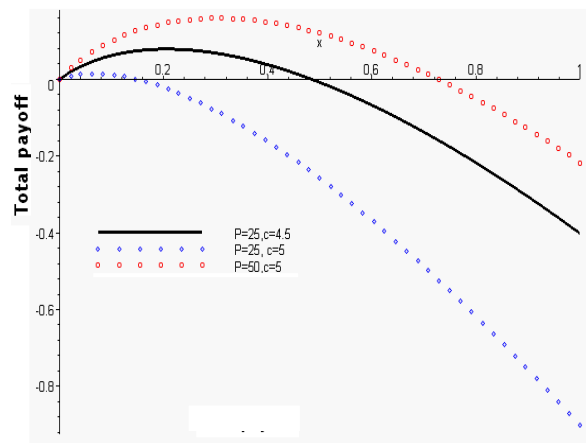
Figure 4.7: *The total payoff of all the population vs x_P.*

Part II

Stochastic Population Games

Chapter 5

Stochastic Population Games

This chapter studies non-cooperative population games with several individual states. The population consists of local resources distributed in several areas and several classes of players. Each local resource has a finite number of states. Each member of each class of the population has (i) its own state that he/she controls (ii) a finite set of actions in each pair individual state-resource state, (iii) its instantaneous reward depends on its state, the resource state and the population's profile at the given time, (iv) time-average (coupled) constraints. We apply this model to power management in ALOHA-like wireless networks, solar-powered broadband wireless networks, and battery state-dependent power control in wireless networks with several types of renewable energies.

5.1 Introduction

Randomness is implicitly hinted in the requirement of robustness against mutations, that we may view as random deviations. But the assumption of a very large population tends to hide this source of randomness since that randomness tends to average out. The evolutionary games without state may provide an interpretation in which the deterministic game is a limits of games with finitely many players who may take random actions. Such an interpretation can be found in (63). Yet, other sources of randomness have been introduced into evolutionary games. Some authors have added small noise to the replicator dynamics in order to avoid the problem of having the dynamics stuck in some local minimum, see (42; 86; 112) and references therein. The ESS can then be replaced by other related notions such as the Stochastic Stable Equilibrium (SSE) (86).

In this paper we introduce a class of stochastic population games with individual states for both players and resources. There are many local interactions among individuals belonging to large populations of players. Each individual stays permanently in the system; from time to time it moves among different individual states, and interacts with other player. The actions of the player along with those with which it interacts determine not only the immediate fitness of the player but also the transition probabilities to the next state it will have. Each individual is thus faced with an MDP in which it maximizes the expected average cost criterion. Each individual knows only the state of its own MDP, and does not know the state of the other players it interacts with. The transition probabilities of a player's MDP are only controlled by that player. The local interactions between players can be viewed as a stochastic game with

coupled constraints and variable number of interacting players. The case where the number of interacting players is constant have been studied in (15; 16).

5.2 Illustrating Examples

5.2.1 Battery state-deependent power management

Consider a large number of mobiles terminals controlling their transmission power and a distributed base stations. Each mobile has an amount of energy E when its battery is new (typically it is the case if the battery is new or if the battery is completely recharged). Each mobile implements a power control policy where the transmission power is allowed to depend on the energy level (state) of its battery. The available action (reachable base stations and powers) depends on the state of the battery. Given the remaining energy of its battery, the mobile have to choose the optimal power level. One of the important element for each mobile is its instantaneous throughput which can be characterized as a function of the signal to interference plus noise ratio (SINR) at the base station where he transmits. The battery is replaced only when it is completely empty. The cost of new battery cost is C. The new battery has the same energy of E. The mobile have to control both the power consumption as well as the time at which the batteries are changed. At each slot, each mobile is faced to a random number (213) of interacting players which transmit at the same base station. Each battery life-time game corresponds to a stochastic population game with finite horizon (absorbing state of battery when the energy is very small). The aim here is to find jointly the power levels and the base stations such that all users achieve as high payoff as possible, minimum guarantee (e.g. QoS requirement thresholds) but also to control the battery-state.

When batteries are recharged dynamically with different types of alternative energy such as renewable energies (solar, wind etc). The battery transition state becomes irreducible Markov decision process under each policy depending on an exogenous parameter which characterizes the good weather (good weather will correspond to the sun for the solar-power systems and to the wind for wind-powered systems). In this case, the interaction becomes a stochastic population game with infinite horizon and we shall consider time-average reward (discounted or not).

5.2.2 Energy management in hybrid Aloha-like networks

Consider a distributed Aloha network with large number of mobile terminals. Each mobile can choose both the channels and powers (this is in contrast to standard Aloha model in which users are associated to the closest receivers). Each terminal is faced to a random number of interacting players which transmit at the channel. A terminal attempts transmissions during a finite horizon of times depending on the state of its battery energy. At each slot, each terminal have to take a decision on the transmission power based on the battery state. At each state of the battery, there are a finite power levels. At the lowest state of battery no power is available and the mobile have to replaced the battery by a new or to recharge its battery. A transmission is successful if no other user transmit during the slot or the mobile transmits with a power which is bigger than the power of all others transmitting mobiles at the same receiver. The pairwise interactions case of this problem has been studied by Altman and Hayel in (20) as a stochastic evolutionary game. They have considered three states: *Full, Almost Empty* and *Empty,* and

simultaneous interactions with more two users are neglected[1]. Their model can be extend to more than two opponent interactions and also to finitely many states as shown in next Section. We can also extend to the case where each terminal is faced to a random number (219; 213; 11) of interacting terminals which transmit at the same range and each terminal have to control an arbitrary transition state of its energy.

5.2.3 Markov decision process

A player arrives at some random time t_0. It has a clock that dictates the times at which interactions with other players occur. It is involved in interactions that occur according to a Poisson process with rate λ. After a random number of time periods, the player leaves the system and is replaced by another one. This will be made precise below. During the player's life time, each time the timer clicks, the player interacts with another randomly selected player.

We associate with each player a Markov Decision Process (MDP) embedded at the instants of the clicks.

The parameters of the MDP are given by the tuple $\{\mathcal{S}, \mathcal{A}, Q\}$ where

- $\mathcal{S}$ is the set of possible individual states of the player

- $\mathcal{A}$ is the set of available actions. For each state s, a subset $\mathcal{A}_s$ of actions is available.

- Q is the set of transition probabilities; for each $s, s' \in \mathcal{S}$ and $a \in \mathcal{A}_s$, $Q_{s'}(s, a)$ is the probability to move from state s to state s' taking action a. $\sum_{s' \in S} Q_{s'}(s, a)$ is allowed to be smaller than 1.

Define further

- The set of policies is $\mathcal{U}$. A general policy u is a sequence $u = (u_1, u_2, \ldots)$ where u_i is a distribution over action space $\mathcal{A}$ at time i. The dependence on time is a local one: it concerns only the individual's clock; a player is not assumed to use policies that make use of some global clocks. There can only be also one global clock for all individuals. A policy is an individual decision which defines the sequences of action which will be taken by the individual at each individual's clock.

- The subset of mixed (respectively pure or deterministic) policies is $\mathcal{U}_M$ (respectively $\mathcal{U}_D$). We define also the set of stationary policies $\mathcal{U}_S$ where such policy does not depend on time.

Frequency state-action Often we encounter the notion of individual states in evolutionary games; but usually the population size at a particular state is fixed. In our case the choices of actions of an individual determine the fraction of time it would spend at each state. Hence the fraction of the whole population that will be at a given state may depend on the distribution of strategies in the population. In order to model this dependence we first need to describe the expected amount of time $f_{\eta,u}(s)$ that an individual spends at a given state s when it follows a strategy u and its initial state at time 1 is distributed according to a probability η over $\mathcal{S}$. More generally, we define $f_{\eta,u}(s, a)$ the expected number of time units during which it is at state s and it chooses action a. $f_u := \{f_{\eta,u}(s, a)\}$ is called the occupation measure corresponding to a policy u or frequency state-action.

---

[1] Note that this assumption does not holds in dense networks.

More precisely, define $p_t(\eta, u; s, a) = \mathbb{P}_{\eta,u}(X_t = s, A_s = a)$ the probability for a user to be in state s, at time t, using action a under policy u when the initial state has a probability distribution η. Further define $p_t(\eta, u; s) = \sum\limits_a p_t(\eta, u; s, a)$.

Define

$$f_{\eta,u}^t(s, a) = \frac{1}{t} \sum_{r=1}^{t} p_r(\eta, u; s, a).$$

Denote $f_\eta^t(u) := \{f_{\eta,u}^t(s, a)\}$. Define by Φ_η^u to be the set of all accumulation points of $f_\eta^t(u)$ as $t \to \infty$. Whenever Φ_η^u contains a single element, we shall denote it by $f_\eta(u)$.

5.3 Single population stochastic evolutionary games

Aim: To model and characterize evolutionary games where individuals have states that are described by controlled Markov chains. The action of an individual in a local interaction with another randomly selected individual determines not only the instantaneous fitness but also its probability to move to another state. The goal of a player is to maximize its time average fitness.

Mathematical methods: The main mathematical tool is occupation measures (expected frequencies of states and actions). This tool is a central one in the theory of Markov Decision Processes. We make use of the geometric properties of the set of achievable occupation measures.

Key assumption: Under any pure stationary policy of an individual, its Markov chain has a single ergodic class of states.

We define and characterize a new concept of Evolutionarily Stable Strategies (OMESS), based on the concept of Occupation Measures. We relate this set to the concept of ESSet (221). We present a way to transform the new type of evolutionary games into standard ones. We apply this novel framework to energy control in wireless networks.

A new rich class of evolutionary games is defined along with the corresponding new definitions of equilibrium. This combination of evolutionary games together with Markov Decision Processes introduces new dynamical features and modeling capabilities.

Keywords: evolutionary games, occupation measure evolutionarily stable strategy, Markov decision process, energy control in wireless networks.

Pairwise interaction we consider a large population of individuals. As in standard evolutionary games, there are many pairwise interactions between randomly selected pairs.

Let $r(s, a, s', b)$ be the immediate reward that a player receives when it is at state s and it uses action a while interacting with a player who is in state s' that uses action b.

Denote by $\alpha(u) = \{\alpha(u; s, a)\}$ the system state: $\alpha(u; s, a)$ is the fraction of the population at individual state s and that use action a when all the population uses strategy u. We shall add the index t to indicate a possible dependence on some time.

Consider an arbitrary tagged player and let S_t and A_t be its state and action at time t (as measured on its individual clock). Then his expected immediate reward at that time is given by

$$R_t = \sum_{s',a'} \alpha_t(u; s', a') r(S_t, A_t, s', a') := r(S_t, A_t, \alpha_t(u)).$$

Assume now that a player arrives at the system at time 1. The global expected fitness when using a policy v is then

$$F_\eta(v, u) = \liminf_{t \to \infty} \frac{1}{t} \sum_{m=1}^{t} E_{\eta,v}[R_m].$$

When η is concentrated on state s we write with some abuse of notation $F_s(v, u) = F_\eta(v, u)$. We shall often omit the index η (in case it is taken to be fixed).

Unless stated differently, we shall make throughout the following assumption.

Introduce the following assumptions.

A2(U): When the whole population uses a policy $u \in U$, then at any time t which is either fixed or is an individual time of an arbitrary player, $\alpha_t(u)$ is independent of t and is given by

$$\alpha_t(u; s, a) = f_{\eta,u}(s, a) = \pi(s)u(a|s)$$

for all s, a where $f_{\eta,u}(s, a)$ is the single limit of $f_{\eta,u}^t(s, a)$ as $t \to \infty$ and π is the stationary distribution of the chain.

A2: Assumption $A2(U)$ holds for $U = U_s$ and for $U = U_M$.

The validity of the Assumption depends on the way the infinite population model is obtained by scaling a large finite population model. This aspect is beyond the scope of this paper. Denote the set of all policies for which Φ_η^u is a singleton by U^*. For $u \in U^*$, the following holds:

$$F(v, u) = \inf_{z \in \Phi_\eta^v} \sum_{s,a} z(s, a) \sum_{s',a'} f_{\eta,u}(s', a') r(s, a, s', a'). \tag{5.1}$$

The set of occupation measures will be shown to be a polytope whose extreme points correspond to strategies in U_D. This will allow us to transform the stochastic population game or Markov decision evolutionary game to a standard evolutionary game.

Note that for any $u \in U_M$, and for any strategies v and w,

$$\Phi_\eta^v \subset \Phi_\eta^w \quad \text{implies } F(v, u) \geq F(w, u). \tag{5.2}$$

This, together with the fact that for any policy u and $z \in \Phi_\eta^u$ there exist a stationary policy $v \in U^*$ satisfying $f_\eta^v = z$, will motivate us to limit ourselves to policies in U^*.

When both u and v are in U^*, the global expected fitness simplifies to

$$F(v, u) = \sum_{t=1}^{\infty} E_{\eta,v} R_t = \sum_{s,a} f_{\eta,v}(s, a) \sum_{s',a'} f_{\eta,u}(s', a') r(s, a, s', a'). \tag{5.3}$$

Assumption A1 would not hold if the policy of a player could depend on the absolute time or on the behavior (i.e. the actions) of other players. For example, in the standard replicator dynamics, the policy of a player adapts to the instantaneous fitness which depends also on the actions of the other players in the population. Thus A1 does not hold there. On the other hand, since players of a given class are undistinguishable, and since the lifetime distribution of a player depends only on his local time, we may expect Assumption A1 to hold. Checking A1 is beyond the scope of the paper.

Definition 5.3.0.1. *We shall say that two strategies u and u' are equivalent if the corresponding occupation measures are equal. We shall write $u =_e u'$. The set of occupation measures equivalent to u is denoted by $e(u) := \{v | v =_e u\}$.*

Note that if u and u' are equivalent policies for a given player then for any v used by the rest of the population, the fitness under u and under u' are the same.

We now define the notion of evolutionary stable set in standard population games. An ESSet is a set of Nash equilibria which have the following special properties (see *Cressman, 2003*).

Definition 5.3.0.2. *A set E of symmetric Nash equilibrium is an evolutionarily stable set (ESSet) if, for all $q \in E$, we have $J(q, p) > J(p, p)$ for all $p \notin E$ and such that $J(p, q) = J(q, q)$.*

Note that for all strategies p and p' in an ESSet E, we have $J(p', p) = J(p, p)$. The concept of ESSet is stronger than Nash equilibrium and there are some simple matrix games in which such an equilibrium set does not exist (see *Weibull, 1995, page 48 example 2.7*).

In (*Thomas, 1985*), the author defines Evolutionary Stable Sets and presents an example of ESSet containing a continuum of (Nash) equilibrium strategies, none of which can be an Evolutionarily Stable Strategy (ESS). Another example is of an ESS, a special case of an ESSet restricted to one point.

The ESSet is robust against perturbation by a strategy which is outside the ESSet, but any strategy in the set need not to be robust against perturbation. For instance, every ESSet is asymptotically stable for the replicator dynamic (*Cressman, 2003*). Every ESSet is a disjoint union of equilibria.

The stronger notion of equilibrium from evolutionary game theory is the Evolutionarily Stable Strategy (ESS). The concept of ESSets generalize the standard concept of ESS as it is a one-element ESSet.

Defining the Occupation Measure ESS

With the expression (5.8) for the fitness, we observe that we are again in the framework of evolutionary games and can use definition of Theorem 5.3.1 for the Occupation Measure ESS (OMESS) in the stochastic population game:

Definition 5.3.0.3. *(i) A strategy $u \in U^*$ is an equilibrium for the stochastic population game if and only if it satisfies*

$$F(u, u) \geq F(v, u). \tag{5.4}$$

(ii) A strategy $u \in U^*$ is a Occupation-measure ESS (OMESS) for the stochastic population game if and only if

- *it is an equilibrium, and*

- *for all $v \in U^*$ such that $v \neq_e u$ that satisfy $F(u, u) = F(v, u)$, the following holds: $F(u, v) > F(v, v)$.*

We could use the following as an equivalent Definition of OMESS for stochastic population game.

Theorem 5.3.1. *A strategy u is said to be OMESS if for every $v \neq_e u$ there exists some $\overline{\epsilon}_v > 0$ such that the following holds for all $\epsilon \in (0, \overline{\epsilon}_v)$:*

$$F(u, \epsilon u + (1 - \epsilon)v) > F(u, \epsilon u + (1 - \epsilon)v) \tag{5.5}$$

In equation (5.5) we use a convex combination of two policies. We delay the definition of this to the next section (see Remark 5.3.1.3).

The following result links between the OMESS and the ESSets.

Proposition 5.3.1.1. *If u is an OMESS, then e(u) is an ESSet.*

Proof. Let u be an OMESS. We take a measure $v \notin e(u)$. By definition of equivalent class we have one of the following condition holds:

- $F(u, u) > F(v, u)$,

- $F(u, u) = F(v, u)$ and $F(u, v) > F(v, v)$.

Thus each $w \in e(u)$ is a Nash equilibrium. The second condition implies that for every $w \in e(u)$ and $v \notin e(u)$ such that $F(w, w) = F(v, w)$, we have $F(w, v) > F(v, v)$. This implies by definition that $e(u)$ is an ESSet. $\qquad\square$

In the following, we show that an ESSet is a weaker notion than OMESS: a problem with no OMESS may still have a non empty ESSet.

Consider a single state s and two actions h or l. Assume that the reward does not depend on the action. two pure stationary policies are u and v where u consists on playing always h and the policy v is to play always l. Then,

- the ESSet of the Markov game is all the feasible policies,

- u and v are not in the same equivalence class,

- $F(u, w) = F(w, w)$, $\forall w \in e(v)^c \neq \emptyset$. v is not an OMESS.

- the game has no OMESS

Computing the OMESS

Define the set of occupation measures achieved by all (individual) policies in some subset $U' \subset U$ as

$$\mathcal{L}_\eta(U') = \bigcup_{u \in U'} f_{\eta,u}(s, a).$$

It will turn out that the expected fitness of an individual (defined in next subsection) will depend on the strategy u of that individual only through $f_{\eta,u}$. We are therefore interested in the following characteristic of $\mathcal{L}_\eta(U)$ (see (119; 26; 178)):

Lemma 5.3.1.2. $\mathcal{L}_\eta(U)$ *equals to the set Q_η defined as the set of $\alpha = \{\alpha(s, a)\}$ satisfying*

$$\sum_{s' \in S} \sum_{a \in A_{s'}} \alpha(s', a)[\delta_{s'}(s) - Q_{s'}(s, a)] = \eta(s), \forall s, \quad \alpha(s, a) \geq 0, \forall s, a. \tag{5.6}$$

where $\delta_{s'}(s)$ is the Dirac distribution in state s'.
(ii) We have: $\mathcal{L}_\eta(U) = \mathcal{L}_\eta(U_S) = co\mathcal{L}_\eta(U_D)$ where $co\mathcal{L}_\eta(U_D)$ is the convex hull of $\mathcal{L}_\eta(U_D)$.
(iii) For any $\alpha \in \mathcal{L}_\eta(U)$, define the individual stationary policy $u \in \mathcal{U}_S$ by

$$u_s(a) = \begin{cases} \dfrac{\alpha(s, a)}{\sum_{a \in A_s} \alpha(s, a)} & \text{if } \sum_{a \in A_s} \alpha(s, a) > 0 \\ \text{arbitrary number in } [0, 1] & \text{if } \sum_{a \in A_s} \alpha(s, a) = 0 \end{cases}$$

Then $f_{\eta,u} = \alpha$.

Transforming stochastic population game the into a standard population game

Consider the following standard evolutionary game **EG**:

- the finite set of actions of a player is U_D,

- the fitness of a player that uses $v \in U_D$ when the other use a policy $u \in U_S$ is given by (5.8).

- Enumerate the strategies in U_D such that $U_D = (u_1, ..., u_m)$ where

$$m = \prod_{s \in \mathcal{S}} |\mathcal{A}_s|,$$

- Define $\gamma = (\gamma_1, ..., \gamma_m)$ where γ_i is the fraction of the population that uses u_i. γ can be interpreted as a mixed strategy which we denote by $\widehat{\gamma}$.

Remarks 5.3.1.3. *Here the convex combination $\epsilon\widehat{\gamma} + (1 - \epsilon)\widehat{\gamma}'$ of the two mixed strategies $\widehat{\gamma}$ and $\widehat{\gamma}'$ is simply the mixed strategy whose ith component is given by $\epsilon\gamma_i + (1 - \epsilon)\gamma_i'$, $i = 1, ..., m$.*

Combining Lemma 5.3.1.2 with eq. (5.7) we obtain:

Proposition 5.3.1.4. *(i) $\widehat{\gamma}$ is an equilibrium for the game **EG** if and only if it is an OMESS for the original stochastic population game.*
*(ii) $\widehat{\gamma}$ is an ESS for the game **EG** if and only if it is a OMESS for the original stochastic population game.*

Proof. The statements hold if we allowed for only mixed policies; indeed, they follow from Lemma 5.3.1.2 and eq. (5.7). We have to check that if a mixed policy is an equilibrium or a OMESS when restricting to U_M then it is also an equilibrium among all policies. This in turn follows from from Lemma 5.3.1.2 and eq. (5.8). $\square$

Non-pairwise interactions We now assume the number of players in local interaction is more than two and this number is not known. Let $r_k(s, a, s_2, a_2, \ldots, s_k, a_k)$ be the immediate reward that a player receives when it is at state s and it uses action a while interacting with $k - 1$ others players who is in state s_j that uses action a_j, $j = 2, \ldots k$. Then the expected immediate reward at that time is given by

$$R_t = \sum_{k \geq 1} P(K = k) \sum_{s_2, a_2, \ldots s_k, a_k} \prod_{j=2}^{k} \alpha_t(u; s_j, a_j) r_k(S_t, A_t, s_2, a_2, \ldots s_k, a_k) =: r(S_t, A_t, \alpha_t(u)).$$

Assume now that a player arrives at the system at time 1. The global expected fitness when using a policy v is then

$$F_\eta(v, u) = \liminf_{t \to \infty} \frac{1}{t} \sum_{m=1}^{t} E_{\eta, v}[R_m].$$

When η is concentrated on state s we write with some abuse of notation $F_s(v, u) = F_\eta(v, u)$. We shall often omit the index η (in case it is taken to be fixed).

For $u \in U^*$, the following holds:

$$F(v, u) = \inf_{z \in \Phi_\eta^v} \sum_{s,a} z(s, a) \sum_{s_2, a_2, \ldots s_k, a_k} r(s, a, s_2, a_2, \ldots s_k, a_k) \prod_{j=2}^{k} f_{\eta, u}(s_j, a_j). \tag{5.7}$$

When both u and v are in U^*, the global expected fitness simplifies to

$$F(v,u) = \sum_{t=1}^{\infty} E_{\eta,v} R_t = \sum_{s,a} f_{\eta,v}(s,a) \sum_{s_2,a_2,\ldots s_k,a_k} r(s,a,s_2,a_2,\ldots,s_k,a_k) \prod_{j=2}^{k} f_{\eta,u}(s_j,a_j) r(s,a,s_2,a_2,\ldots s_k,a_k).$$

(5.8)

5.4 Stochastic population games : multiclass case

Consider the following model of population game denoted by

$$\Gamma = (P, (Y^p)_{p \in P}, (A^p(y))_{p \in P, y \in Y^p}, (Q^p)_{p \in P}, (r^p)_{p \in P}),$$

where

- The population is composed of several subpopulations. Each subpopulation contains a large number of players. P denotes the set of subpopulations (we assume that P is finite).

- Each player of each subpopulation p has its own state Y^p (finite) and Markov transition structures Q^p between the states.

- For every player i from the subpopulation $p \in P$ and every state $y \in Y^p$ of i, $A^p(y)$ is the set of actions available. The action space of the subpopulation p is given by $\prod_{y \in Y^p} A^p(y)$.

  The set of all actions at all states is given by All^p where

  $$All^p = \{(y,a),\ y \in Y^p,\ a \in A^p(y)\}.$$

- We denote by $\Delta(Y^p)$ the $(|Y^p| - 1)$-dimensional simplex of $\mathbb{R}^{|Y^p|}$ and by $q^p : All^p \to \Delta(Y^p)$ a transition rule between the states. The transition probability distribution between states is defined by

  $$Q_{y,a,y'}^p := q^p(y'|y,a) = q^p(y'|y_1,a_1,\ldots,y_{t-1},a_{t-1},y,a)$$

  for each $y',y \in Y^p, a \in A^p(y)$.

- For every subpopulation $p \in P$,

  $$r^p : \prod_{p'} \prod_{y \in Y^{p'}} \mathcal{X}^{p'}(y) \to \mathbb{R}^{\sum_{y \in Y^{p'}} |A^{p'}(y)|}$$

  is the vector of all instantaneous payoff functions of a player from the class p',

  $$\mathcal{X}^{p'}(y) = \left\{ (x^{p'}(y,b))_{b \in A^{p'}(y)} \mid x^{p'}(y,b) \geq 0,\ \sum_{y \in Y^{p'}} \sum_{b \in A^{p'}(y)} x^{p'}(y,b) = m^{p'} \right\}$$

  where $m^{p'}$ is the mass associate to the subpopulation p'. Given a state y and strategy profile x^p, x^{-p}, the payoff obtained by playing the action $a \in A^p(y)$ is $r_{y,a}^p(x)$

- The game is played many times.

124

5.4.1 Histories and Strategies

Histories A history h_t at time t is a collection of states and actions $(y_1, a_1, x_1, \ldots, y_{t-1}, a_{t-1}, x_{t-1}, y_t)$. We denote by

$$H_t^p = (All^p \times \mathcal{X})^{t-1} \times Y^p$$

the set of histories of a member of the subpopulation p at time t. At $t = 1$, $H_1^p = Y^p$. Let H_∞^p be the set of all infinite histories of the subpopulation p endowed with the product $\sigma-$field and $H_\infty = \prod_{p \in P} H_\infty^p$.

Strategies

- **Pure strategy** A pure strategy of a player from subpopulation p at time t is a map σ_t^p : $H_t^p \longrightarrow A^p(y_t)$. The collection $\sigma^p = (\sigma_t^p)_{t \geq 1}$ of pure strategy at each time constitutes a pure strategy of the subpopulation p. We denote by Σ^p the set of all pure strategies of subpopulation p, by $\Sigma = \prod_p \Sigma^p$ the set of all pure strategy profiles. Note that the number of pure strategies is infinite.

- **Stationary strategy:** σ is stationary strategy if for each population p and every time t and histories,

$$h_t = (y_1, a_1, x_1, \ldots, y_{t-1}, a_{t-1}, x_{t-1}, y_t),$$
$$h_t' = (y_1', a_1', x_1', \ldots, y_{t-1}', a_{t-1}', x_{t-1}, y_t')$$

such that if $y_t = y_t'$ one has $\sigma_t(h_t) = \sigma_t(h_t')$ i.e a stationary strategy is a history and time independent strategy which depends on the state only.
Lemma 5.4.1.1. *The number of pure stationary strategies is* $\prod_{p \in P} \prod_{y \in Y^p} |A^p(y)|$.

- **Behavioral strategy** A behavioral strategy at time t is a function that assigns each finite history to a mixed action profile of the current state: $\sigma_t^p : H_t^p \longrightarrow \prod_p \Delta(A^p(y_t))$, $p \in P$.

- **Mixed strategy** A mixed strategy profile is a collection of probability distributions on Σ. Using Tychonoff's theorem, the set of all these notions of strategies is compact in the product set histories spaces in the sense of the weak-topology. A general mixed strategy is a probability distribution on the behaviorial strategies set.

For any strategy profile $\sigma = (\sigma^p)_{p \in P}$ and every initial state distribution profile $\mu = (\mu^p)_{p \in P}$, a probability measure $\mathbb{P}_{\sigma,\mu}$ is induced by σ and μ. The stochastic process $(y_t, a_t, x_t)_{t \geq 1}$ is defined on H_∞ in a canonical way, where the random variables y_t, a_t, x_t describe the individual state, the action in this state and the population profile.

5.4.2 Cesaro-limit Fitness

We examine the limit average Cesaro-type payoff

$$F_\mu^p(\sigma^p, \sigma^{-p}) = \mathbb{E}_{\sigma,\mu}\left[\liminf_{T \longrightarrow +\infty} \frac{1}{T}\left(\sum_{t=1}^T r_{y_t, a_t}^p(x_t)\right)\right]$$

where $\mathbb{E}_{\sigma,\mu}$ denotes the expectation over the probability measure $\mathbb{P}_{\sigma,\mu}$ induced by σ, μ on the set of histories endowed with the product $\sigma-$algebra.

Symbol	element	Assignments
Y^p	y	state of the subpopulation p.
$A^p(y)$	a	Set of actions of subpopulation p in state $y \in Y^p$.
q^p	$q^p(y'|y, a)$	Transition law from y to y'.
$r^p_{y,a}(x)$		Instantaneous reward.
H_t	h_t, h'_t	Set of histories for the time t.
Σ^p	σ^p	Set of pure strategies of the subpopulation p.
Σ	σ	Set of pure strategies profile.
$F^p_\mu(\sigma)$		Expected fitness function starting from μ.

***Table 5.1:** Notations*

Given a strategy σ and a initial state y, we define the expected time-average payoff. We denote by Π^p the stationary limit average matrix:

$$\Pi^p(\sigma^p) = \lim_{t \longrightarrow +\infty} \frac{1}{t} \sum_{j=1}^{t} (Q^p)^j(\sigma^p).$$

The matrix Π^p is well-defined, commutes with Q^p and satisfies the projection equation: $\Pi^p \times \Pi^p = \Pi^p$.

If F^p is the vector $(F^p_y(x))_{y \in Y^p}$, we have that $F^p(x) = \Pi^p r^p(x)$ for all stationary strategy profile x. Then, $F^p = \Pi^p F^p$. Note that the function $x \longmapsto F^p(x)$ is not necessarily continuous because the limit matrix $\Pi^p(x)$ can be discontinuous on x.

Definition 5.4.2.1. *A strategy σ is an $\epsilon-$equilibrium if for all p,*

$$F^p(\sigma) + \epsilon \geq F^p(\sigma'^p, \sigma^p), \; \forall \; \sigma'^p \in \Delta(\Sigma^p).$$

A $0-$equilibrium is called equilibrium.

Notes

- When each member of each subpopulation has a single state, we obtain a population game model which each local interaction is repeated game.

- If there exists a subpopulation p^* such that $|Y^p| = 1$ for all $p \neq p^*$. We obtain a stochastic population game with single class of controllers which is the subpopulation p.

5.5 Constrained stochastic population games

In addition to the model described in Section 5.4, we assume that players have (possibly coupled) average constraints on their actions in any state. The payoff of the subpopulation p is

$$E_{\sigma,\mu} \left(\liminf_{t \longrightarrow +\infty} \frac{1}{t} \sum_{k=1}^{t} r^p_{y_k,a_k}(x_k) \right)$$

with $\sigma^p \in \Delta(\Sigma^p)$ subject to

- Orthogonal constraints:

$$p \in P, \ P_{\sigma^p} \left(\limsup_{t \longrightarrow +\infty} \frac{1}{t} \sum_{k=1}^{t} D^p(y_k, a_k) \leq \beta^p \right) = 1,$$

where $D^p : All^p \longrightarrow \mathbb{R}$ is an individual cost function (independent of the strategies of the others players), $\beta^p \in \mathbb{R}$ is a given cost threshold.

- Coupled constraints:

$$p \in P, \ P_{\sigma,\mu} \left(\limsup_{t \longrightarrow +\infty} \frac{1}{t} \sum_{k=1}^{t} C^p_{y_k, a_k}(x_k) \leq \alpha^p \right) = 1$$

where $C^p : All^p \times \mathcal{X} \longrightarrow \mathbb{R}$ is a cost function which depends on the individual state-action but also on the population profile i.e the strategies of the others players (in the same class or not).

A strategy σ is a constrained equilibrium if for all p,

$$F^p(\sigma) \geq F^p(\sigma'^p, \sigma^p), \ \forall \ \sigma'^p \in \Delta(\Sigma^p)$$

and $\sigma^p \in \Lambda(\sigma^{-p})$ where $\Lambda(\sigma^{-p})$ is the set feasible strategies (that satisfy the orthogonal and coupled constraints) given the strategies of the others populations σ^{-p}.

Constrained stochastic population games with unknown horizon

In general, the lifetime of an individual or of the system is not known. We shall integrate this in our interaction model. In this section we develop a general formulation of a local interaction with unknown stopping time. Players does not known the length of the local interaction but have a common probability structure on the stochastic local game. At time t, they assign some probability $\mathbb{P}(T = t)$ to the event $\{T = t\}$ that the local interaction ends in time t.

$$t \geq 1, \ \mathbb{P}(T = t) \geq 0, \ \sum_{t \geq 1} \mathbb{P}(T = t) = 1.$$

Fix an anonymous member of some subpopulation p, and a sequence of state-actions σ. A player from the class p will receive

$$F^p_\mu(\sigma) = \mathbb{E}_{\sigma,\mu} \left[\liminf_{t \to +\infty} \frac{\sum_{k=1}^{t} \mathbb{P}(T = k)(\sum_{j=1}^{k} r^p_{y_j, a_j}(x_j))}{\sum_{j=1}^{t} j\mathbb{P}(T = j)} \right]$$

under the constraints: $p \in P$,

$$P_\sigma \left(\limsup_{t \to +\infty} \frac{\sum_{k=1}^{t} \mathbb{P}(T = k)(\sum_{j=1}^{k} C^p_{y_j, a_j}(x_j))}{\sum_{j=1}^{t} j\mathbb{P}(T = j)} \leq \alpha^p \right) = 1,$$

$$P_{\sigma^p} \left(\limsup_{t \to +\infty} \frac{\sum_{k=1}^{t} \mathbb{P}(T = k)(\sum_{j=1}^{k} D^p(y_j, a_j))}{\sum_{j=1}^{t} j\mathbb{P}(T = j)} \leq \beta^p \right) = 1,$$

Theorem 5.5.1.

$$F_\mu^p(\sigma) = \mathbb{E}_{\sigma,\mu} \liminf_{t \to +\infty} F_\mu^{p,t}(\sigma)$$

$$\text{where } F_\mu^{p,t}(\sigma) := \frac{1}{\sum_{j=1}^t j\mathbb{P}(T=j)} \sum_{j=1}^t \left(\sum_{k=j}^t \mathbb{P}(T=k) \right) r_{y_j,a_j}^p(x_j)$$

Proof. We apply Fubini's theorem on finite summation to change the order between k and j in the expression of $F_\mu^{p,t}(\sigma)$ where

$$F_\mu^{p,t}(\sigma) = \frac{\sum_{k=1}^t \mathbb{P}(T=k)(\sum_{j=1}^k r_{y_j,a_j}^p(x_j))}{\sum_{j=1}^t j\mathbb{P}(T=j)}.$$

□

Examples: This model generalizes the finite and infinite horizon payoff notions:

- If T is the Dirac measure concentrated on t_* i.e $\mathbb{P}(T=j) = 0$ if $j \neq t_*$ and $\mathbb{P}(T=t_*) = 1$, we obtained the arithmetic average payoff

$$\mathbb{E}_{\sigma,\mu} \frac{\sum_{j=1}^{t_*} r_{y_j,a_j}^p(x_j)}{t_*}$$

- If T is the geometric distribution $\mathbb{P}(T=t) = (1-\delta)\delta^{t-1}$, then we obtain the average discounted payoff:

$$(1-\delta) \mathbb{E}_\sigma \sum_{t=1}^{+\infty} \delta^{t-1} r_{y_t,a_t}^p(x_t)$$

- Note that when the expected horizon of local interaction is finite (for example when the lifetime of the system or of the user is finite - in expectation - but the end of the interaction is not known)[2], the average payoff can be rewritten as

$$F_\mu^p(\sigma) = \frac{\sum_{j=1}^{+\infty} \left(\sum_{k=j}^{+\infty} \mathbb{P}(T=k) \right) \mathbb{E}_{\sigma,\mu} r_{y_j,a_j}^p(x_j)}{\sum_{j=1}^{+\infty} j\mathbb{P}(T=j)} = \frac{\sum_{t \geq 1} \mathbb{P}(T \geq t)\mathbb{E}_{\sigma,\mu} r_{y_t,a_t}^p(x_t)}{\mathbb{E}(T)} \quad (5.9)$$

The following theorem generalizes the Theorem 2.1 in (22) for constrained games and also the Theorem 2.6 (ii) in (21) and the Theorem 1 in (83) for unconstrained product games.

Theorem 5.5.2. *Assume that each subpopulation has a single (aperiodic) ergodic class under each stationary strategy. Then the stochastic population game with individual independent states and unknown lifetime has an equilibrium in stationary strategies. Moreover, the constrained game has an equilibrium under Slater condition.*

For the proof we need tightness properties of the measure generated by the frequencies state-actions under the distribution of the horizon.

---

[2]Note that the expected horizon can be finite and $\mathbb{P}(T=t) > 0$. It is the case for $\mathbb{P}(T=t) = \delta^{t-1}(1-\delta)$, $\delta \in (0,1)$.$\mathbb{E}(T) = \frac{1}{1-\delta}$

Proposition 5.5.2.1. *Assume that for each subpopulation p and any stationary strategy σ^p, the state process is an irreducible Markov chain with one ergodic class then, for any strategy σ the frequencies state-actions*

$$(f_{\mu,\sigma}^{p,t}(y,a))_{p \in P, t \geq 1}$$

where

$$f_{\sigma,\mu}^{p,t}(y,a) = \frac{\sum_{j=1}^{t}\left(\sum_{k=j}^{t}\mathbb{P}(T=k)\right)P_\sigma(X_j=y,a_j=a|y_1=\mu)}{\sum_{j=1}^{t}j\mathbb{P}(T=j)}$$

are tight.

The occupation measures in this extended model are characterized by the following convergence result: $\mathbb{P}_{\sigma,\mu}$ almost surely, the random variables that give the frequencies state-action

$$\frac{\sum_{k=1}^{t}P(T=k)\sum_{j=2}^{k}\delta_{y'}(X_j^p) - \sum_{y,a}Q_{yay'}^p\delta_{(y,a)}(X_{j-1}^p,a_{j-1}^p)}{\sum_{j=1}^{t}jP(T=j)} \tag{5.10}$$

goes to zero when t goes to infinity, for all $y' \in Y^p$, $p \in P$. Hence, when $E(T) = \sum_{k \geq 0} P(T > k) = \sum_{k} kP(T=k) < +\infty$ then we obtain the equation:

$$\frac{\sum_{k=1}^{+\infty}P(T=k)\sum_{j=2}^{+\infty}\delta_{y'}(X_j^p) - \sum_{y,a}Q_{yay'}^p\delta_{(y,a)}(X_{j-1}^p,a_{j-1}^p)}{\sum_{j=2}^{\infty}P(T \geq j)} \tag{5.11}$$

5.6 Energy Control in Wireless Networks

We next illustrate the stochastic population game setting with a problem that arises in dynamic power control in mobile networks.

Users participate in local competitions for the access to a shared medium in order to transmit their packets. An individual state of each mobile represents the energy level at the user's battery which, for simplicity, we assume to take finitely many values, denoted by $\mathcal{S} = \{0, \ldots, n\}$.

Each time the battery empties (which corresponds to reaching state 0), the mobile changes the battery to a new one (this corresponds to state n), and pay a cost C. We assume that each time a mobile reaches state zero, it remains there during a period whose expected duration is τ.

In each state $s \in \mathcal{S} \setminus \{0\}$, each mobile has two available actions h and l which correspond respectively to high power p_H and low power p_L. We consider an Aloha-type game where a mobile transmits a packet with success during a slot if:

- with probability p, the mobile is the only one to transmit during this slot,

- the mobile transmits with high power and the other transmitting mobile uses low power or is in state 0.

5.6.1 Time average fitness criterion

The reward function r depends on a mobile's state as well as on the transmission powers, that is, the action of the mobile as well as that of the one it interacts with. Then we have for $s \neq 0$:

$$r(s,a,s',a') = p + (1-p)1_{(s'=0)} + (1-p)1_{\{(a=h),\,(a'=l),\,(s'\neq0)\}}.$$

129

For $s = 0$ we take $r(0, a, s', a') = C/\tau$.

For each state $s \in \mathcal{S} \setminus \{0\}$, the transition probability $Q_{s'}(s, a)$ may be non-zero (for both $a \in \{l, h\}$) only for $s' \in \{s, s - 1\}$. Then, as the two possible transitions are to remain at the same energy level or move to the next lower one, we simplify the notation and use $Q(s, a)$ to denote the probability of remaining at energy level s using action a.

To model the fact that the mobiles stays in the average τ units at state 0 and then moves to state n we set the transition probabilities from state 0 to any state other than n and 0 to be zero; the probability to move to n is $1/\tau$ and that of remaining at 0 is $1 - 1/\tau$.

The transition probabilities between energy levels which are motivated by the application of energy consumption satisfy:

- For all state $s \in \mathcal{S} \setminus \{0\}$, we have $Q(s, h) < Q(s, l)$ because using less power induces higher probability to remain in the same energy level.

- For all state $s \in \mathcal{S} \setminus \{0\}$ and for both actions $a \in \{l, h\}$, we have $Q(s, a) > Q(s - 1, a)$ because less battery energy the mobile has, less is the probability to remain at the same energy level.

We pursue the example described in Section 5.6 applying the latest proposition in order to obtain the OMESS for this stochastic population game. Indeed, we will find the OMESS for the related EG game which will be written as a matrix game with dimension 4. In order to find the equilibrium of this matrix game, we have to compute the fitness $\widetilde{F}(v, u)$ for all policies v and u. We use the renewal theorem to find the expected fitness par cycle of lifetime.

$$\widetilde{F}(v, u) = p \frac{T_{\eta,v}}{T_{\eta,v} + \tau} \frac{T_{\eta,u}}{T_{\eta,u} + \tau} + (1 - p) \frac{T_{\eta,v}(h)}{T_{\eta,v} + \tau} \frac{T_{\eta,u}(l)}{T_{\eta,u} + \tau} +$$

$$\tau \frac{1}{T_{\eta,u} + \tau} \frac{T_{\eta,v}}{T_{\eta,v} + \tau} - C \frac{1}{T_{\eta,v} + \tau}$$

where $T_{\eta,v}(a)$ is the expected number of times the action a is used under the policy v starting from initial distribution η.

In a first step, we have to compute the occupation measure f_u corresponding to each policy $u \in \{u_1, u_2, u_3, u_4\}$; for that we need the probability for a user to be in each state, at time t, using action a under policy u. At initial time $t = 0$, a mobile always starts with a battery full of energy, that is $\eta = (0, 0, 1)$. We describe the matrix game with the four following matrices:

$$\widetilde{F}_1(u_i, u_j) = \frac{T_{\eta,u_i}}{T_{\eta,u_i} + \tau} \frac{T_{\eta,u_j}}{T_{\eta,u_j} + \tau},$$

$$\widetilde{F}_2(u_i, u_j) = \frac{T_{\eta,u_i}(h)}{T_{\eta,u_i} + \tau} \frac{T_{\eta,u_j}(l)}{T_{\eta,u_j} + \tau},$$

$$\widetilde{F}_3(u_i, u_j) = \frac{1}{T_{\eta,u_j} + \tau} \frac{T_{\eta,u_i}}{T_{\eta,u_i} + \tau},$$

and

$$\widetilde{F_4} = \begin{pmatrix} \dfrac{1}{X_1 + X_3 + \tau} & \dfrac{1}{X_1 + X_3 + \tau} & \dfrac{1}{X_1 + X_3 + \tau} & \dfrac{1}{X_1 + X_3 + \tau} \\[2mm] \dfrac{1}{X_1 + X_4 + \tau} & \dfrac{1}{X_1 + X_4 + \tau} & \dfrac{1}{X_1 + X_4 + \tau} & \dfrac{1}{X_1 + X_4 + \tau} \\[2mm] \dfrac{1}{X_2 + X_3 + \tau} & \dfrac{1}{X_2 + X_3 + \tau} & \dfrac{1}{X_2 + X_3 + \tau} & \dfrac{1}{X_2 + X_3 + \tau} \\[2mm] \dfrac{1}{X_2 + X_4 + \tau} & \dfrac{1}{X_2 + X_4 + \tau} & \dfrac{1}{X_2 + X_4 + \tau} & \dfrac{1}{X_2 + X_4 + \tau} \end{pmatrix}$$

with

$$X_1 = \frac{1}{1 - Q(1,l)}, \quad X_2 = \frac{1}{1 - Q(1,h)},$$
$$X_3 = \frac{1}{1 - Q(2,l)}, \quad X_4 = \frac{1}{1 - Q(2,h)}.$$

Then we obtain the following modified fitnesses depending on the policies in the following matrix:

$$\widetilde{F} = p\widetilde{F_1} + (1 - p)\widetilde{F_2} + \tau\widetilde{F_3} - C\widetilde{F_4}.$$

The OMESS of the stochastic population game which model energy control behaviors in wireless networks is obtained by finding the OMESS of the standard EG with the matrix of fitnesses given by $\widetilde{F}$.

5.7 Energy control: absorbing state

We study in this Section an energy management non-cooperative game with an infinite number of players modeled as a stochastic population game. We consider pairwise interactions where each player has to choose between an aggressive or a non-aggressive action. The lifetime of each player depends on the level of aggressiveness of all his action during his life. The instantaneous reward of each player depends on the level of aggressiveness of his action and also of his opponent. We consider different restricted strategies and study the existence of evolutionary stable strategies.

We assume that each mobile terminal is an individual player of a global population and the interaction for transmission is interpreted as a fight for food (as in (139)). It is important to note that the Markov process of each player is controlled by himself, whereas in (139) the interaction determines the evolution of the individual process. Each player starts his life in the highest energy state or level, say state n. Starting from state n, each player will visit the different levels in the decreasing order until reaching the last state 0. The state 0 corresponds to the state *Empty* and the state n is the *Full* state of the battery. The other states $1, \ldots, n-1$ are intermediary states of the battery or energy. When the system is at the state 0 there is no energy and the player has to charge his battery or to buy a new battery. We will call *terminal absorbing fitness* the expected cost of charge or price of a new battery.

At each time slot there should be an interaction with another player and then an action should be taken by each player in the population. The actions taken by a player determine

not only the immediate reward but also the transition probabilities to its next individual state (energy level). The instantaneous reward depends also on the level of aggressiveness of the others players.

Then, the evolution of the level of energy depending on the player's action can be modeled as a discrete homogeneous time Markov process.

5.7.1 Individual sequential decision

Each player has to choose between two possible actions in each state (different from the *Empty* state): high(h) and low(l). The strategy h is a aggressive strategy and the strategy l is a passive one. Since there are only two strategies h and l, the population aggressiveness can be described by a process $(\alpha_t)_{t \geq 1}$ where α_t is the fraction of the population using the aggressive strategy, that is the action h, at time t. The stochastic process α_t may depend on the history with length t. The main issue of this paper is to study how a player will manage its energy during his lifetime in order to optimize its total reward.

5.7.2 Binary Reward

We define a binary reward of a player in an interaction as follows: The player win the resource if (i) he has no opponents for the same resource during the time slot, or (ii) he uses the most aggressive action and its opponents are passive. Otherwise, he gets zero. We extend the reward function to mixed strategies by randomized the binary reward.

We denote by $1 - p$ the probability for any player of having an opponent at the same time slot. The probability of success (probability to win to the resource) of a player using the action h when its energy level is at the state $s > 0$ is given by

$$r(s, h, \alpha) = p + (1 - p)[\alpha \times 0 + (1 - \alpha) \times 1] = p + (1 - p)(1 - \alpha), \qquad (5.12)$$

where α is the fraction of the population who use the action h at any given time. Note that assuming that α is fixed in time does not mean that the actions of each player are fixed in time. It only reflects a situation in which the system attains a stationary regime due to the averaging over a very large population, and the fact that all players choose an action in a given individual state using the same probability low.

Similarly, when he use l the probability of success in state $s \neq 0$ is

$$r(s, l, \alpha) = p. \qquad (5.13)$$

The expected reward of a player when its energy level is at the state $s > 0$ and he uses h with probability β_s is then given by

$$r(s, \beta_s, \alpha) = \beta_s r(s, h, \alpha) + (1 - \beta_s) r(s, l, \alpha) = p + \beta_s (1 - p)(1 - \alpha). \qquad (5.14)$$

Under this reward function, a player has more chance to win the resource with an aggressive strategy than a passive one, but a passive strategy save more energy for future.

Remarks 5.7.2.1. *Consider a pairwise one-shot interaction between a player i in state $s \neq 0$ and a player j in state $s' \neq 0$. Define the action set of each player as $\{h, l\}$ and the payoff function as the binary reward defined above. From the probability to win, it is not difficult to see that the instantaneous reward $r(s, \beta_s, \alpha)$ increases with β_s. Then, the one-shot game is a degenerate game with infinite Nash equilibria.*

We will see that the non-aggressive which is weakly dominated in the one-shot game is not necessarily dominated in the long-term dynamic game that we will describe later.

A pure strategy of a player at time t is a map $\alpha_t : H_t \longrightarrow \{h, l\}$, where $H_t = All^{t-1} \times S$ the set of histories at time t with $All = \{(s, a) \mid s = 1, \ldots, n, \ a \in \{h, l\}\}$. The collection $\sigma = (\sigma_t)_{t \geq 1}$ of pure strategy at each time constitutes a pure strategy of the population. We denote by Σ the set of all pure strategies.

Let σ be a strategy profile, and $h_t = (s_1, a_1, \ldots, s_{t-1}, a_{t-1}, s_t)$ be a history of length t. The continuation strategy $\overrightarrow{\sigma}_{h_t}$ from round t given the history h_t can be defined as follows

$$\overrightarrow{\sigma}_{h_t}(s_1', a_1', \ldots, a_{t'-1}', s_{t'}') = \sigma(s_1, a_1, \ldots, s_{t-1}, a_{t-1}, s_t, s_1', a_1', \ldots, a_{t'-1}', s_{t'}').$$

The fitness obtained with the policy $\overrightarrow{\sigma}$, $v = r(\overrightarrow{\sigma}_{h_t})$ is called *continuation fitness* after h_t. If σ is a stationary policy then v depends only on the state s_t.

For each state $s \neq 0$, the action space of each player becomes $[0, 1]$. We define as a player policy, the set of actions during his battery life, that is the sequence $u = (u_1, u_2, \ldots)$ where u_i is the probability of choosing action h at time slot i. We consider the energy level of energy at time t as the random process $(X_t)_{t \geq 1}$ and energy management is a Markov decision process.

For the remainder we shall consider only stationary policies where the probabilities u_i is the same for all time i and we denote it by the vector $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ where β_s is the probability of choosing action h in state s. We denote by $\overline{h} = (1, \ldots, 1)$ and $\overline{l} = (0, \ldots, 0)$ the two pure aggressive and non-aggressive policies.
Given a stationary policy β and a strategy of the global population $\alpha = (\alpha_t)_{t \geq 1}$, the transitions in energy levels of the battery is described by the Markov process (X_t) with the transition law

$$\forall s \neq 0, \forall a \in \{h, l\}, \ q(X_{t+1} = s' | X_t = s, a) = \begin{cases} Q_s(a) & \text{if } s' = s \\ 1 - Q_s(a) & \text{if } s' = s - 1 \\ 0 & \text{otherwise} \end{cases}$$

and $q(0|0, a) = 1$.

We assume two main characteristics of the transition probabilities of energy levels depending on the action taken by the player.

- First, $\forall s \neq 0$, $Q_s(h) < Q_s(l) < 1$, reflects that using action h, the probability to remain in any given energy level is lower than using action l,

- second, $\forall a$, the discrete function $s \longmapsto Q_s(a)$ is non-decreasing. This assumption means that lower is the energy level, lower is the probability to remain in the same level for each action.

Note that each player controls the transition state of its own battery: Q is independent of the decision of the other players. Let $T(\beta)$ the expected time to reach to the state 0 given the initial state *Full n* under the stationary policy β. Then the total reward of the player in all the lifetime of its battery, starting in state n, is given by

$$V_\beta(n, \alpha) = \mathbb{E}^{\beta, \alpha} \sum_{t=1}^{T(\beta)} r(s_t, \beta_{s_t}, \alpha_t) = r(n, \beta_n, \alpha) + \mathbb{E}^{\beta, \alpha} \left(\sum_{t=2}^{T(\beta)} r(s_t, \beta_{s_t}, \alpha) \, | s_{t_1} = n \right). \tag{5.15}$$

5.7.3 Computing Fitness using Dynamic Programming

We define the dynamic programming operator $Y(s, a, \alpha, v)$ to be the total expected fitness of a player starting at stat s if

- It takes action a at time 1,

- at time 2 the total sum of expected fitness from time 2 onwards is v,

- At each time slot, the probability that another player uses action h, given that there is a interaction between individual players, is α.

We denote by v the continuation fitness function which depends on the state s and the global population strategy α. We have

$$Y(s, a, \alpha, v) = r(s, a, \alpha) + Q_s(a)v(s, \alpha) + (1 - Q_s(a))v(s - 1, \alpha), \forall a \in \{h, l\}.$$

The function v is a fixed point of the operator ϕ defined by

$$v \longrightarrow (\phi v) := \beta_s Y(s, h, \alpha, v) + (1 - \beta_s)Y(s, l, \alpha, v).$$

Proposition 5.7.3.1. *The expected fitness for a player during its lifetime is given by:*

$$V_\beta(n, \alpha) = \sum_{i=1}^{n} \frac{r(i, \beta_i, \alpha)}{1 - Q_i(\beta_i)} - C.$$

Proof. We compute the fitness by using the recursive formula defined by the operator ϕ. First, we have that $v(0, \alpha) = 0$ and then

$$\begin{aligned}
v(1, \alpha) = \phi(v) &= \beta_1[r(1, h, \alpha) + Q_1(h)v(s, \alpha) + 0] + (1 - \beta_1)[r(s, l, \alpha) + Q_s(l)v(1, \alpha) + 0], \\
&= r(1, \beta_1, \alpha) + [\beta_1 Q_1(h) + (1 - \beta_1)Q_1(l)]v(1, \alpha).
\end{aligned}$$

The fixed point $v(1, \alpha)$ is given by

$$V_\beta(1, \alpha) = \frac{r(1, \beta_1, \alpha)}{1 - [\beta_1 Q_1(h) + (1 - \beta_1)Q_1(l)]} = \frac{p + \beta_1(1 - p)(1 - \alpha)}{1 - [\beta_1 Q_1(h) + (1 - \beta_1)Q_1(l)]}.$$

For all state $s > 1$ we have

$$\begin{aligned}
v(s, \alpha) &= \beta_s[r(s, h, \alpha) + Q_s(h)v(s, \alpha) + (1 - Q_s(h))v(s - 1, \alpha)], \\
&\quad + (1 - \beta_s)[r(s, l, \alpha) + Q_s(l)v(s, \alpha) + (1 - Q_s(l))v(s - 1, \alpha)] \\
&= r(s, \beta_s, \alpha) + [\beta_s Q_s(h) + (1 - \beta_s)Q_s(l)]v(s, \alpha) + [\beta_s(1 - Q_s(h)) + (1 - \beta_s)(1 - Q_s(l))]v(s - 1, \alpha), \\
&= r(s, \beta_s, \alpha) + Q_s(\beta_s)v(s, \alpha) + (1 - Q_s(\beta_s))v(s - 1, \alpha),
\end{aligned}$$

where $Q_s(\beta_s) := \beta_s Q_s(h) + (1 - \beta_s)Q_s(l)$. Since $Q_s(\beta_s) < 1$ one has

$$v(s, \alpha) = \frac{r(s, \beta_s, \alpha)}{1 - Q_s(\beta_s)} + v(s - 1, \alpha).$$

By rewriting

$$v(s, \alpha) = \sum_{i=1}^{s}[v(i, \alpha) - v(i - 1, \alpha)],$$

we obtain that

$$V_\beta(s, \alpha) = \sum_{i=1}^{s} \frac{r(i, \beta_i, \alpha)}{1 - Q_i(\beta_i)}.$$

Hence, the expected fitness obtained by starting from the *Full* state n to the next new battery is then given by

$$V_\beta(n, \alpha) - C.$$

$\square$

5.7.4 Sojourn Times

We denote by $T_s(\beta_s)$ the sojourn time of a player at the state s with the policy β_s. The sojourn time satisfies the following fixed point formula:

$$T_s(\beta_s) = 1 + \sum_{a \in \{h,l\}} Q_s(a)\beta_s(a)T_s(\beta_s).$$

Hence,

$$T_s(\beta_s) = \frac{1}{1 - Q_s(\beta_s)}$$

The fraction of times that the player chooses the aggressive h in state s is given by

$$\hat{\alpha}_s(\beta) = \beta_s \frac{T_s(\beta_s)}{\sum_{i=1}^n T_i(\beta_i)}$$

and the probability for any player to meet another player which is using action h is given by

$$\hat{\alpha}(\beta) := \sum_{s=1}^n \hat{\alpha}_s(\beta) = \frac{\sum_s \beta_s T_s(\beta_s)}{\sum_{i=1}^n T_i(\beta_i)} \leq 1. \tag{5.16}$$

The function $\hat{\alpha}$ gives a mapping between the strategy vector β of an anonymous player and the global population strategy α.

5.7.5 Reduced game

We compare the two type of attitudes: use always the high power ($\bar{h}$) or use always low power $\bar{l}$. The payoffs corresponding to these strategies are described as follows:

	$\bar{h}$	$\bar{l}$
$\bar{h}$	pT_h	T_h
$\bar{l}$	pT_l	pT_l

where $T_h = \sum_{s=1}^n \dfrac{1}{1 - Q_s(h)}$ (resp. $T_l = \sum_{s=1}^n \dfrac{1}{1 - Q_s(l)}$) is the expected lifetime of the battery under the policy $\bar{h}$ (resp. $\bar{l}$). T_h can be interpreted also as the hitting time to the state 0 starting from the state *Full*. This description gives us a two-player matrix game called *reduced* game. Generically, the reduced game has a unique ESS and global maximizer. Since the remaining energy of battery decreases with the power consumption: $T_h < T_l$ one has:

- If $p = 0$. Then $\bar{h}$ is the unique ESS and $\dfrac{1}{2}$ is the global maximizer.

- If $0 < p < \dfrac{T_h}{T_l}$ then $v = \dfrac{1}{1-p} - \dfrac{p}{1-p}\dfrac{T_l}{T_h}$ is the unique ESS and the global maximizer is $\dfrac{v}{2}$.

- If $\dfrac{T_h}{T_l} < p \leq 1$ then $\bar{l}$ is the unique ESS and global maximizer of the reduced game.

5.7.6 Deterministic strategies

The action taken in each state s is determined with a given distribution the first time the player is in this state. We are able to compute the ESS of this game by transforming our stochastic population game into a standard Evolutionary Game depicted as a matrix game.

Lemma 5.7.6.1. *The number of pure stationary strategies is 2^n.*

Proof. A pure strategy is an application from to set of states $S\backslash\{0\}$ to the actions $\{h, l\}$. Thus, the number of the pure strategies is the cardinal of the set $\{h, l\}^{S\backslash\{0\}}$ which is exactly 2^n. $\quad\square$

Denote by $U = \{\beta : y \in S\backslash\{0\} \longmapsto \{h, l\}\} = \{h, l\}^{S\backslash\{0\}}$ the set of pure stationary strategies. The cardinal of U is $2^{|S|-1}$.

Define the following finite population game $\Gamma_d = (U, G)$:

- $U = \{h, l\}^{S\backslash\{0\}}$ is the set of actions. $\Delta(U)$ the set of probability measure on U (this corresponds to be mixed strategies obtained by randomizing the pure stationary strategies).

- the payoff matrix is given by
$$G_{u,u'} = V_u(n, \alpha(u')).$$

If I denotes the set of states in which the high power is used then

$$G_{u,u'} = V_u(n, \alpha(u')) = \sum_{s=1}^{n} \frac{p + (1-p)(1 - \alpha(u'))u_s}{1 - Q_s(u_s)}$$

where $s \geq 1$, $u_s \in \{0, 1\}$ and

$$\alpha(u') = \frac{\sum_{s \in I} T_s(1)}{\sum_{s \in I} T_s(1) + \sum_{s \in S\backslash I} T_s(0)}.$$

The size of the matrix is $2^n \times 2^n$.

The strategy $w \in \Delta(U)$ is an equilibrium of Γ_d if for all $u \in U$,

$$w_u > 0 \implies G_u(w) = \max_{u' \in U} G_{u'}(w)$$

where

$$G_u(w) = \sum_{u'} G_{uu'} w_{u'}.$$

From Nash-Glicksberg's fixed point theorem, and the symmetry of the game, the following result holds:

Lemma 5.7.6.2. *The game Γ_d has at least one symmetric equilibrium (in $\Delta(U)$).*

Any symmetric equilibrium of the matrix game Γ_d induced an equilibrium of the population game in stationary strategies (this is because the set of symmetric mixed strategies in the matrix game contains the set of stationary strategies). Any strategy $u \in U$ satisfying

$$G_{uu} > G_{u'u}, \ \forall \, u' \in U\backslash u, \tag{5.17}$$

is an ESS of the evolutionary game. Note that the non-strict inequalities $G_{uu} \geq G_{u'u}, \ \forall \, u' \in U$, are necessary condition for u to be an equilibrium in the matrix game Γ_d.

5.7.7 Pure stationary strategies

Preliminaries

As already mentioned, our game is different and has a more complex structure than a standard evolutionary game. In particular, the fitness that is maximized is not the outcome of a single interaction but of the sum of fitnesses obtained during all the opportunities in the player's lifetime.

We obtain the following characterization of an Evolutionary Stable Strategy (ESS).

Lemma 5.7.7.1. *A necessary condition for β^* to be an ESS is*

$$\text{for all } \beta' \neq \beta^*, \quad V_{\beta^*}(n, \widehat{\alpha}(\beta^*)) \geq V_{\beta'}(n, \widehat{\alpha}(\beta^*)) \tag{5.18}$$

A sufficiency condition for β^* to be an ESS is that (5.18) holds with strict inequality for all $\beta' \neq \beta^*$.

Using this corollary and this simple necessary condition to be an ESS, we are now able to determine the existence of an ESS. First we considering the restricted game to independent state actions, i.e. the probability of choosing the aggressive action h does not depend on the individual state. Second, we generalize to dependent state actions which is more complicated. However we need latter the following lemma in numerous proofs.

Lemma 5.7.7.2. *Given $\hat{\alpha}(\beta')$, the function $Z : \beta_s \longmapsto \dfrac{p + (1-p)(1 - \hat{\alpha}(\beta'))\beta_s}{1 - Q_s(\beta_s)}$ is monotone.*

Proof. Since Z is a continuously differentiable function, Z is strictly decreasing if

$$(1-p)(1 - \hat{\alpha}(\beta'))(1 - Q_s(l)) < (Q_s(l) - Q_s(h))p,$$

constant if

$$(1-p)(1 - \hat{\alpha}(\beta'))(1 - Q_s(l)) = (Q_s(l) - Q_s(h))p$$

and strictly increasing if

$$(1-p)(1 - \hat{\alpha}(\beta'))(1 - Q_s(l)) > (Q_s(l) - Q_s(h))p.$$

Thus, given p, Q, the function Z is monotone. $\qquad\square$

5.7.8 State-independent strategies

We examine the case where the probability of choosing action h does not depend on the level of energy of the battery, that is $\forall s\ \beta_s = \rho$. We first consider only $\bar{h}$ the fully aggressive strategy (the player uses always the high energy level available in the state) or $\bar{l}$ the fully passive strategy (the player uses always weak action level when it is possible) policies.

Proposition 5.7.8.1. *If $p \neq 0$, then the strategy $\bar{h}$ which consists to play always h ($\rho = 1$) cannot be an ESS. Otherwise, if $p = 0$ then $\bar{h}$ is an evolutionary stable strategy.*

Proof. Suppose that $p \neq 0$. Then, for all stationary mixed policy β'

$$V_{\bar{0}}(n, \alpha(\beta')) = p \sum_s \frac{1}{1 - Q_s(l)}, \quad \text{and}$$

$$V_{\bar{1}}(n, \alpha(\bar{1})) = p \sum_s \frac{1}{1 - Q_s(h)}.$$

Thus $V_1(n, \alpha(1)) < V_0(n, \alpha(1))$. It is better to use the low power l if the opponent use high power h (the payoff is p and the lifetime is increased). Hence, the strategy $\bar{h}$ is not is a best reply to itself. This means that $\bar{h}$ is not a Nash equilibrium and then $\bar{h}$ is not an ESS.

Suppose now that $p = 0$. Then the strategy l is weakly dominated, and $V_{\bar{h}}(n, \alpha(\bar{h})) = 0 = V_{\bar{l}}(n, \alpha(\beta'))$, $\forall \beta'$. Moreover, for all $\beta' \neq \bar{h}$ we have

$$V_{\beta'}(n, \alpha(\bar{1})) = 0 = V_{\bar{h}}(n, \alpha(\bar{h})), \text{ and}$$

$$V_{\bar{h}}(n, \alpha(\beta')) = (1 - \alpha(\beta')) \sum_s \frac{1}{1 - Q_s(h)} > V_{\beta'}(n, \alpha(\beta')) = (1 - \alpha(\beta')) \sum_s \frac{\beta'_s}{1 - Q_s(\beta'_s)}.$$

This completes the proof. □

This result is somehow logical as when, at each time slot, any player is in interaction with another player, the only strategy which permits to win the fight is to be aggressive. We have the same kind of result for the non-aggressive strategy.

Proposition 5.7.8.2. *The strategy $\bar{l}$ which consists to play always l ($\rho = 0$) is an ESS if and only if the subset of states $I_1 := \{s, \ p < \frac{1 - Q_s(l)}{1 - Q_s(h)}\}$ is empty.*

Proof. We have

$$V_0(n, \alpha(\beta')) = p \sum_{s=1}^n \frac{1}{1 - Q_s(l)}, \quad \text{and} \quad V_\beta(n, \alpha(0)) = \sum_{s=1}^n \frac{p + (1 - p)\beta_s}{1 - Q_s(\beta)}.$$

Thus

$$\max_\beta V_\beta(n, \alpha(0)) = \sum_{s \in I_1} \frac{1}{1 - Q_s(h)} + p \sum_{s \in I_0} \frac{1}{1 - Q_s(l)} + \sum_{s \in I_*} \frac{1}{1 - Q_s(h)} < V_0(n, \alpha(0)) \quad (5.19)$$

where

$$I_1 := \{s, \ p < \frac{1 - Q_s(l)}{1 - Q_s(h)}\},$$

$$I_0 := \{s, \ p > \frac{1 - Q_s(l)}{1 - Q_s(h)}\},$$

$$I_* := \{s, \ p = \frac{1 - Q_s(l)}{1 - Q_s(h)}\}.$$

The subsets I_0, I_1, I_* constitute a partition of S. I_* has at most one element. Using lemma 5.7.7.2 and the equation (5.19), $\bar{l}$ is a best response to itself if

$$\sum_{s \in I_1} \frac{1}{1 - Q_s(h)} < p \sum_{s \in I_1} \frac{1}{1 - Q_s(l)}.$$

This inequality does not holds if I_1 is not empty. □

We consider now stationary mixed strategies where each player determine the probability ρ of choosing the aggressive strategy at each energy level of his battery.

Proposition 5.7.8.3. *If the same level of aggressiveness $\beta_s = \rho \in (0, 1)$, $\forall s$ is an ESS then it must satisfy*

$$\frac{p \sum_{s=1}^n \frac{(Q_s(l) - Q_s(h))}{(1 - Q_s(\rho))^2}}{(1 - p) \sum_{s=1}^n \frac{(1 - Q_s(l))}{(1 - Q_s(\rho))^2}} = 1 - \rho. \quad (5.20)$$

If p, Q and $0 \leq \rho \leq 1$ satisfy (5.20), then ρ is an ESS.

Proof. The strategy ρ is a best reply to the strategy β' if it satisfies $V_\rho(n, \alpha(\beta')) \geq V_\mu(n, \alpha(\beta'))$, $\forall \mu \in [0,1]^n$. Then, ρ is a best response to itself if

$$V_\rho(n, \alpha(\rho)) \geq V_\mu(n, \alpha(\rho)), \forall \mu \in [0,1]^n.$$

A necessary condition for an interior extremum of the function $h : x \longmapsto \sum_{s=1}^n \frac{p + (1-\rho)(1-p)x}{1 - Q_s(x)}$
is to satisfy

$$h'(x) = 0 \iff 1 - \rho = \frac{p \sum_{s=1}^n \frac{(Q_s(l) - Q_s(h))}{(1 - Q_s(x))^2}}{(1-p) \sum_s \frac{(1 - Q_s(l))}{(1 - Q_s(x))^2}}.$$

Note that $\alpha(\rho) = \rho$.

We now show that for every p, Q such that $\dfrac{p \sum_{s=1}^n \frac{(Q_s(l) - Q_s(h))}{(1 - Q_s(\rho))^2}}{(1-p) \sum_s \frac{(1 - Q_s(l))}{(1 - Q_s(\rho))^2}} \leq 1$, every strategy $(\rho, 1 - \rho)$

satisfying

$$1 - \rho = \frac{p \sum_{s=1}^n \frac{(Q_s(l) - Q_s(h))}{(1 - Q_s(\rho))^2}}{(1-p) \sum_s \frac{(1 - Q_s(l))}{(1 - Q_s(\rho))^2}}$$

is an ESS. Let β' be a strategy such that $V_\rho(n, \rho) = V_{\beta'}(n, \rho)$. Denote by $C := V_\rho(n, \alpha(\beta')) - V_{\beta'}(n, \alpha(\beta'))$. We now prove that C is strictly positive for all $\beta' \neq \rho$.

The condition $C > 0$ is satisfied if

$$V_\rho(n, \alpha(\beta')) > V_{\beta'}(n, \alpha(\beta')),$$
$$\sum_{s=1}^n \frac{p + (1-p)(1 - \alpha(\beta'))\rho}{1 - Q_s(\rho)} > \sum_{s=1}^n \frac{p + (1-p)(1 - \alpha(\beta'))\beta'}{1 - Q_s(\beta')}.$$

Since $V_\rho(n, \rho) = V_{\beta'}(n, \rho)$, one has

$$[p + (1-p)(1-\rho)\rho] \sum_s \frac{1}{1 - Q_s(\rho)} = [p + (1-p)(1-\rho)\beta'] \sum_s \frac{1}{1 - Q_s(\beta')}.$$

This implies that

$$\sum_s \frac{1}{1 - Q_s(\beta')} = \frac{p + (1-p)(1-\rho)\rho}{p + (1-p)(1-\rho)\beta'} \sum_s \frac{1}{1 - Q_s(\rho)}.$$

By restituting this value in C, we obtain $C > 0$ if

$$\sum_{s=1}^n \frac{p + (1-p)(1 - \alpha(\beta'))\rho}{1 - Q_s(\rho)} > [p + (1-p)(1 - \alpha(\beta'))\beta'] \frac{p + (1-p)(1-\rho)\rho}{p + (1-p)(1-\rho)\beta'} \sum_s \frac{1}{1 - Q_s(\rho)},$$

$$(p + (1-p)(1 - \alpha(\beta'))\rho) > [p + (1-p)(1 - \alpha(\beta'))\beta'] \frac{p + (1-p)(1-\rho)\rho}{p + (1-p)(1-\rho)\beta'},$$

$$(p + (1-p)(1 - \alpha(\beta'))\rho)(p + (1-p)(1-\rho)\beta') > [p + (1-p)(1 - \alpha(\beta'))\beta'](p + (1-p)(1-\rho)\rho).$$

Developing and simplifying the last expression, we obtain that $D > 0$ is equivalent to $[\beta' - \rho][\alpha(\beta') - \rho] = [\beta' - \rho]^2 > 0$. This completes the proof. $\square$

We are now able to determine explicitly, if it exists, an ESS in stationary mixed strategies of our stochastic population game for energy management. Then, it is interesting to provide conditions of existence of such stationary mixed ESS.

Proposition 5.7.8.4. *A sufficiency condition of existence of an interior state-independent evolutionary stable strategy is given by*

$$p < p_0 := \frac{\sum_{s=1}^{n} \frac{1}{1-Q_s(l)}}{\sum_{s=1}^{n} \frac{(1-Q_s(h))}{(1-Q_s(l))^2}}. \tag{5.21}$$

Proof. The function

$$\xi(p,\rho) = p \sum_{s=1}^{n} \frac{(Q_s(l) - Q_s(h))}{(1-Q_s(\rho))^2} - (1-p)(1-\rho) \sum_{s=1}^{n} \frac{(1-Q_s(l))}{(1-Q_s(\rho))^2},$$

is continuous on $(0,1)$ and $\xi(p,1) > 0$, $\forall p \in (0,1)$. Thus, if (5.21) satisfied then $\xi(p,0) < 0$ and 0 is in the image of $(0,1)$ by $\xi(p,.)$. This implies that there exists $\rho \in (0,1)$ such that $\xi(p,\rho) = 0$. □

Proposition 5.7.8.5. *The game has at most one ESS in state independent strategies.*

Proof. (i) Uniqueness of pure ESS: We have to examine the two strategies: $\bar{h}$ and $\bar{l}$. From Proposition 5.7.8.2 $\bar{l}$ is an ESS if and only $p > \max_{s \neq 0} \left(\frac{1 - Q_s(l)}{1 - Q_s(h)} \right)$ and from Proposition 5.7.8.1, $\bar{h}$ is an ESS if and only $p = 0$. Since, $\max_{s \neq 0} \left(\frac{1 - Q_s(l)}{1 - Q_s(h)} \right) > 0$, we conclude that only one of two strategies can be an ESS.

(ii) Strictly mixed ESS (or interior ESS): A necessary condition to interior ESS is given by Equation (5.20). Let ρ and ρ' in $(0,1)$ be two solutions of (5.20) i.e

$$\xi(p,\rho) = \xi(p,\rho') = 0.$$

From Proposition 5.7.8.3, one has

$$V_{\rho'}(n, \alpha(\rho')) \geq V_\rho(n, \alpha(\rho')) > V_{\rho'}(n, \alpha(\rho'))$$

which is a contradiction. We conclude that if an ESS exists in state-independent strategies, it is unique. □

Note that the inequality (5.21) is satisfied if

$$0 < p < \min_{s \neq 0} \left(\frac{1 - Q_s(l)}{1 - Q_s(h)} \right). \tag{5.22}$$

Moreover the following result holds:

Lemma 5.7.8.6.

$$\min_{s \neq 0} \left(\frac{1 - Q_s(l)}{1 - Q_s(h)} \right) < p_0 < \max_{s \neq 0} \left(\frac{1 - Q_s(l)}{1 - Q_s(h)} \right). \tag{5.23}$$

Proof. Let $a = \min\limits_{s \neq 0} \left(\dfrac{1 - Q_s(l)}{1 - Q_s(h)} \right)$ and $b = \max\limits_{s \neq 0} \left(\dfrac{1 - Q_s(l)}{1 - Q_s(h)} \right)$. Then,

$$\forall s, \qquad\qquad\qquad b > \frac{1 - Q_s(l)}{1 - Q_s(h)} > a \qquad\qquad\qquad \Longleftrightarrow \qquad (5.24)$$

$$\forall s, \qquad b(1 - Q_s(h)) > 1 - Q_s(l) > a(1 - Q_s(h)) \qquad \Longleftrightarrow \qquad (5.25)$$

$$\forall s, \quad b(1 - Q_s(h)) - (1 - Q_s(l)) > 0 > a(1 - Q_s(h)) - (1 - Q_s(l)) \quad \Longleftrightarrow \qquad (5.26)$$

$$\forall s, \quad \frac{b(1 - Q_s(h)) - (1 - Q_s(l))}{(1 - Q_s(l))^2} > 0 > \frac{a(1 - Q_s(h)) - (1 - Q_s(l))}{(1 - Q_s(l))^2} \qquad (5.27)$$

By taking the sum over s from one to n, one has,

$$b \sum_{s=1}^{n} \frac{1 - Q_s(h)}{(1 - Q_s(l))^2} - \sum_{s=1}^{n} \frac{1}{1 - Q_s(l)} > 0 > a \sum_{s=1}^{n} \frac{1 - Q_s(h)}{(1 - Q_s(l))^2} - \sum_{s=1}^{n} \frac{1}{1 - Q_s(l)}.$$

This means that $\xi(b, 0) > \xi(p_0, 0) > \xi(a, 0)$. Since, the function $p \longmapsto \xi(p, \rho)$ is strictly increasing, the last inequality implies that $b > p_0 > a$. This completes the proof. $\qquad\square$

The relation (5.22) gives another sufficient condition of existence of a stationary mixed ESS. By proposition 5.7.8.2, there is a pure non-aggressive ESS $(\bar{l})$ if and only if $p > \max\limits_{s \neq 0} (\dfrac{1 - Q_s(l)}{1 - Q_s(h)})$. Then in some particular cases of parameters p and transition probabilities, there is no pure independent stationary ESS nor mixed, when

$$p_0 < p < \max_{s \neq 0} (\frac{1 - Q_s(l)}{1 - Q_s(h)}).$$

The cooperative optimal strategy for players is denoted by ρ^* that is defined by:

$$\widetilde{\rho} = \arg\max_{\beta} V_\beta(n, \beta).$$

This strategy gives the global optimum solution of the centralized system of our energy management model. We are interested in comparing the aggressiveness of the ESS to the global optimum solution.

Proposition 5.7.8.7. *Let ρ^* be the ESS in stationary strategies of the stochastic population game and $\widetilde{\rho}$ the global optimum solution. Then we have*

$$\widetilde{\rho} \leq min\{\frac{1}{2}, \rho^*\}.$$

Proof. The function

$$\beta \longmapsto V_\beta(n, \beta) = \sum_{s=1}^{n} \frac{p + (1 - p)(1 - \beta)\beta}{1 - Q_s(l) + \beta(Q_s(l) - Q_s(h))}$$

is continuous and strictly decreasing on $(\dfrac{1}{2}, 1)$. Thus, the function has a global maximizer on $[0, 1]$ and the global maximizer is lower than $\dfrac{1}{2}$.

Let ρ^* be an ESS. Suppose that $\widetilde{\rho} > \rho^*$. Since ρ^* is an ESS, ρ^* satisfies

$$V_{\rho^*}(n, \rho^*) \geq V_{\widetilde{\rho}}(n, \rho)$$

and $V_{\rho^*}(n,\widetilde{\rho}) > V_{\widetilde{\rho}}(n,\widetilde{\rho})$ if $V_{\rho^*}(n,\rho^*) = V_{\widetilde{\rho}}(n,\rho^*)$. Given a strategy β, the function $\alpha \longmapsto V_\beta(n,\alpha)$ is strictly decreasing. Hence,

$$V_{\rho^*}(n,\rho^*) > V_{\rho^*}(n,\widetilde{\rho}) > V_{\widetilde{\rho}}(n,\widetilde{\rho}).$$

The inequality

$$V_{\rho^*}(n,\rho^*) > V_{\widetilde{\rho}}(n,\widetilde{\rho})$$

is a contradiction with the definition of $\widetilde{\rho}$. Hence, $\widetilde{\rho} \le \rho^*$. We conclude that a global maximizer is lower than $\dfrac{1}{2}$ and $\widetilde{\rho}$ coincides with ρ^* (it is the case if the ESS is a pure strategy) or $\widetilde{\rho} < \rho^*$. $\qquad\square$

The main important result here is that the ESS is more aggressive than the global optimum solution, i.e. $\rho^* \ge \widetilde{\rho}$. This seems relatively intuitive because in a context of an evolutionary game, every player is somehow afraid to meet another player using an aggressive strategy, then the aggressive strategy is more used in the population.

5.7.9   State-dependent actions

We consider now that the action taken by each player depends on his energy level or state. We first compute explicitly the best response correspondence BR $:$ $S\backslash\{0\} \longrightarrow 2^I$ where I is the compact set $[0,1]$.

Proposition 5.7.9.1. *In stationary strategies, the best response to the strategy $\alpha(\beta')$ is determined by $(\beta_1,\ldots,\beta_n)$ such that for all $s = 1,\ldots,n$,*

$$\beta_s(\alpha(\beta')) = \begin{cases} 1 & \text{if } \alpha(\beta') < 1 - \dfrac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \\[2mm] 0 & \text{if } \alpha(\beta') > 1 - \dfrac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \\[2mm] \text{any strategy } \eta_s \in [0,1] & \text{if } \alpha(\beta') = 1 - \dfrac{p(Q_s(l)-Q_s(h))}{(1-p)(1-Q_s(l))} \end{cases}$$

Thus, the optimal stationary strategy $\beta = (\beta_1,\ldots,\beta_n)$ is characterized by

$$\beta_s = \begin{cases} 1 & \text{if } s \in I_\alpha := \{j, \ \dfrac{Q_j(l)-Q_j(h)}{(1-Q_j(l))} < (1-\alpha)(-1+\dfrac{1}{p})\} \\[2mm] 0 & \text{if } s \in J_\alpha := \{j, \ \dfrac{Q_j(l)-Q_j(h)}{(1-Q_j(l))} > (1-\alpha)(-1+\dfrac{1}{p})\} \end{cases}$$

Proof. The best reply β_s to $\alpha(\beta')$ maximizes the function $Z : \ \beta_s \longmapsto \dfrac{p+(1-p)(1-\hat\alpha(\beta'))\beta_s}{1-Q_s(\beta_s)}$ defined in Lemma 5.7.7.2. From Lemma 5.7.7.2, Z is monotone. The maximizer is one if

$$(1-p)(1-\hat\alpha(\beta'))(1-Q_s(l)) > (Q_s(l)-Q_s(h))p,$$

zero if

$$(1-p)(1-\hat\alpha(\beta'))(1-Q_s(l)) < (Q_s(l)-Q_s(h))p,$$

and any strategy in $[0,1]$ is maximizer of Z if the equality

$$(1-p)(1-\hat\alpha(\beta'))(1-Q_s(l)) = (Q_s(l)-Q_s(h))p$$

holds. We conclude that

$$
\beta_s(\alpha(\beta')) = \begin{cases}
1 & \text{if } \alpha(\beta') < 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1-p)(1 - Q_s(l))} \\[2ex]
0 & \text{if } \alpha(\beta') > 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1-p)(1 - Q_s(l))} \\[2ex]
\text{any strategy } \eta_s \in [0,1] & \text{if } \alpha(\beta') = 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1-p)(1 - Q_s(l))}
\end{cases}
$$

$\square$

Let

$$
f(s) = \frac{T_s(l)}{T_s(h)} = \frac{1 - Q_s(h)}{1 - Q_s(l)}, \tag{5.28}
$$

be the ratio between the sojourn time with the low power l and the sojourn time with h in each state s. By taking assumption on the transition probabilities, we have that for all energy level s, $f(s) > 1$. This ratio can be interpreted also as the proportion of time obtained taken the non-aggressive action l in state s compared to the aggressive one h.

Definition 5.7.9.2 (Threshold Policies). *We define the two following pure threshold policies u_1 and u_2:*

- *Risky Policy (RP) : there exists a state s such that $\forall s' > s$, $u_1(s') = L$ and $\forall s'' < s$, $u_1(s'') = H$. This policy is called a control limit policy in (139).*

- *Carefully Policy (CP) : there exists a state s such that $\forall s' > s$, $u_2(s') = H$ and $\forall s'' < s$, $u_2(s'') = L$.*

That kind of threshold policies, keeping the same action until the level of energy is low, has been also obtained in (139) in a context of a dynamic Hawk and Dove game. The inequality $\alpha(\beta') < 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1-p)(1 - Q_s(l))} = 1 - \dfrac{p}{1-p}(f(s) - 1)$ is equivalent to

$$
(1 - \alpha(\beta'))\frac{1-p}{p} > f(s) - 1, \tag{5.29}
$$

where $f(s) - 1$ is exactly the gain in mean sojourn time in state s using action l instead of action h. Equation 5.29 can be rewritten:

$$
\frac{p}{1 - Q_s(l)} < \frac{1 - (1 - p(1 - \alpha(\beta')))}{1 - Q_s(h)},
$$

where the left side term is the average reward for a player during the time it is in state s and it uses action l, and the right term is the average reward during the time it is in state s and it uses action h. Then, the threshold policies are based on the comparison between the average instantaneous reward in each state depending on the action taken by each player which determines the instantaneous reward and also the remaining time in this state.

We have the following result showing that the best response strategies are u_1 or u_2 depending on the structure of the Markov Decision Process.

Proposition 5.7.9.3. *The best response is u_1 (resp. u_2) if $f(s)$ is increasing (resp. decreasing).*

Proof. First we assume that the function f is decreasing, meaning that more the level of energy is high, more is the gain of surviving time in each energy level. If $\dfrac{1 - Q_n(h)}{1 - Q_n(l)}) < (1 - \alpha)(-1 + \dfrac{1}{p})$

then
$$\beta_s = \begin{cases} \text{plays } h & \text{if } s \geq j(\alpha, p) \\ \text{plays } l & \text{if } s \leq j(\alpha, p) - 2 \end{cases}$$

else if $\dfrac{1 - Q_n(h)}{1 - Q_n(l)} > (1 - \alpha)(-1 + \dfrac{1}{p})$ then β_s consists to play always l where

$$j(\alpha, p) := \min\{j, \ \frac{1 - Q_j(h)}{1 - Q_j(l)} < (1 - \alpha)(-1 + \frac{1}{p})\}.$$

Note that one has at most one state s such that $\dfrac{1 - Q_j(s)}{1 - Q_s(l)} = (1 - \alpha)(-1 + \dfrac{1}{p})$. We have the inverse relations if $f(s)$ is increasing. $\qquad\square$

Theorem 5.7.10 (Partially mixed equilibrium). *If there exists a level s such that $\beta_s = \dfrac{\alpha^* + (1 - Q_s(l))(\kappa_2 \alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2 \alpha^* - \kappa_1)} \in$
(0, 1) where*

$$\kappa_1 = \sum_{j \in I} \frac{1}{1 - Q_j(h)}, \ \kappa_2 = \kappa_1 + \sum_{j \notin I, \ j \neq s} \frac{1}{1 - Q_j(l)}, \ \alpha^* = 1 - \frac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))},$$

$$I = \{j \neq s, \ f(s) > f(j)\}$$

and β_j, $j \neq s$ given by

$$\beta_j = \left\{ \begin{array}{ll} \text{plays } h & \text{if } \dfrac{(Q_s(l) - Q_s(h))}{(1 - Q_s(l))} > \dfrac{(Q_j(l) - Q_j(h))}{(1 - Q_j(l))} \ \text{i.e } f(s) > f(j) \\ \text{plays } l & \text{otherwise} \end{array} \right\},$$

then $\beta = (\beta_1, \ldots, \beta_n)$ is an equilibrium.

Proof. By Proposition 5.7.9.1, the best reply has the form

$$BR(\beta_s) = \left\{ \begin{array}{ll} \text{play action } h & \text{if } \alpha(\beta_s) < 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))}, \\ \text{play action } l & \text{if } \alpha(\beta_s) > 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))}, \\ \text{any strategy } h, l \text{ or mixed} & \text{if } \alpha(\beta_s) = 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))}. \end{array} \right.$$

The fixed point equation gives $\beta \in BR(\beta)$ with

$$\alpha(\beta) = \frac{\sum_{s=1}^{n} \frac{\beta_s}{1 - Q_s(\beta_s)}}{\sum_{s=1}^{n} \frac{1}{1 - Q_s(\beta_s)}}.$$

Since f is monotone and $Q_s(l) < Q_s(h)$, there exist at most one state s_0 such that

$$\frac{1 - Q_{s_0}(h)}{1 - Q_{s_0}(l)} = (1 - \alpha(\beta))(-1 + \frac{1}{p}) \tag{5.30}$$

and a mixed equilibrium is characterized by

$$\beta'_s = \begin{cases} \text{play action } h & \text{if } s' > s_0, \\ \text{play action } l & \text{if } s' < s_0, \\ \text{any strategy } h, l \text{ or mixed} & \text{if } s' = s_0. \end{cases}$$

where β_{s_0} satisfies equation 5.30. After some basic calculations, equation 5.30 has a unique solution given by

$$\beta_{s_0} = \frac{\alpha^* + (1 - Q_s(l))(\kappa_2\alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2\alpha^* - \kappa_1)} \in (0, 1)$$

with

$$\kappa_1 = \sum_{j \in I} \frac{1}{1 - Q_j(h)}, \ \kappa_2 = \kappa_1 + \sum_{j \notin I, \ j \neq s} \frac{1}{1 - Q_j(l)}, \ \alpha^* = 1 - \frac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))},$$

$$I = \{j \neq s, \ f(s) > f(j)\}.$$

$\square$

Note that a necessary condition of existence such s is that

$$p > \frac{(1 - c)(1 - Q_s(l))}{(1 - Q_s(h) - c(1 - Q_s(l)))}$$

with $c = \dfrac{1 + \kappa_1(1 - Q_s(h))}{1 + \kappa_2(1 - Qs(h))} \in (0, 1)$. We then obtain several equilibria by permutation on the state satisfying $\beta_s = \dfrac{\alpha^* + (1 - Q_s(l))(\kappa_2\alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2\alpha^* - \kappa_1)} \in (0, 1)$.

Lemma 5.7.10.1. *The partially mixed equilibrium is decreasing in p.*

Proof. It suffices to proof in the unique state s in which $\beta_s \in (0, 1)$. The fraction of players with the aggressive action in state s is then given by

$$\beta_s = \frac{\alpha^* + (1 - Q_s(l))(\kappa_2\alpha^* - \kappa_1)}{1 - (Q_s(l) - Q_s(h))(\kappa_2\alpha^* - \kappa_1)} \in (0, 1)$$

where

$$\kappa_1 = \sum_{j \in I} \frac{1}{1 - Q_j(h)}, \ \kappa_2 = \kappa_1 + \sum_{j \notin I, \ j \neq s} \frac{1}{1 - Q_j(l)},$$

$$\alpha^* = 1 - \frac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))}, \ I = \{j \neq s, \ f(s) > f(j)\}.$$

Since the function $p \longmapsto \alpha^*(p) = 1 - \dfrac{p(Q_s(l) - Q_s(h))}{(1 - p)(1 - Q_s(l))}$ is strictly decreasing, the denominator is non-decreasing in p and the numerator is strictly decreasing in p. Thus β_s is decreasing in p. This completes the proof.

$\square$

5.7.11 Dynamics

By proposition 5.7.9.3 we know that, when the function f is monotone, the best response strategies have a given structure defined as threshold policies RP or CP. In order to numerically observe the equilibrium policies in the population, we construct the replicator dynamics which have properties of convergence to equilibrium and ESS in evolutionary games. A replicator dynamic is a differential equation that describes the way strategies change in time as a function of the fitness. Roughly speaking they are based on the idea that the average growth rate per player that uses a given action is proportional to the excess of fitness of that action with respect to the average fitness.

When f is monotone increasing (resp. decreasing) the set of policies are $u_1^1, \ldots, u_1^{n+1}$ (resp. $u_2^1, \ldots, u_2^{n+1}$) where policy u_1^i (resp. u_2^i) consists of taking action h (resp. l) in the $i-1$ first states. For all $i = 1, \ldots, n+1$ the proportion of the population playing u_1^i is denoted by σ_i. The replicator dynamics describes the evolution of the different policies in the population in time and is given by the following differential equation:

$$\dot{\sigma}_i(t) = \sigma_i(t) \left[\sum_{j=1}^{n+1} \sigma_j(t) G_{u_1^i, u_1^j} - \sum_{k=1}^{n+1} \sum_{j=1}^{n+1} \sigma_k(t) \sigma_j(t) G_{u_1^k, u_1^j} \right] =: f_i(\sigma(t)). \tag{5.31}$$

The replicator dynamics satisfies the so-called *positive correlation* condition:

$$f(\sigma) \neq 0 \implies \sum_j G_{u_1^j}(\sigma) f_j(\sigma) > 0.$$

It follows from (**?**) that any equilibrium σ^* (in the $n-$simplex of the $(n+1)-$dimensional Euclidean space $\mathbb{R}^{n+1}$) of the evolutionary game is a rest point of the replicator dynamics. In order to describe the long-term behavior of the dynamics, we shall say that a stationary point (or rest point) σ^* is *stable* under (5.31) if for every neighborhood $\mathcal{N}$ of σ^* there exists a neighborhood $\mathcal{N}' \subseteq \mathcal{N}$ of σ^* such that $\sigma^* \in \mathcal{N}' \implies \sigma(t) \in \mathcal{N}$, $\forall t \geq 0$. If σ^* is a stable rest point of (5.31) then σ^* is an equilibrium.

5.7.12 Numerical Illustrations

We consider an evolutionary game where each player has $n = 5$ levels of energy. Each player starts his lifetime full of energy (level 5) and has to decide between an aggressive or a non-aggressive action at each pairwise interaction with another player. We assume that all players are in the same population that means that all players have the same transition probabilities Q.

We consider the following transition probabilities

$$Q = \begin{pmatrix} 0.1 & 0.05 \\ 0.2 & 0.1 \\ 0.5 & 0.4 \\ 0.7 & 0.6 \\ 0.8 & 0.7 \end{pmatrix},$$

where for $i = 1, \ldots, 5$, Q_{i1} (resp. Q_{i2}) is the probability to remain in energy level i using action l (resp. h).

5.7.13 State-Independent action

Each player chooses action h with a probability that does not depend on their level of energy. We denote this probability by ρ. First, using this transition probabilities, the threshold p_0 given by equation 5.21 induces the existence of an ESS is

$$p_0 = 0.75$$

Then using result from proposition 5.7.8.4, if the probability p is less than 0.75, there exists an ESS which is the unique solution of the system defined in equation 5.20.

Second, we have $M = \max_s \dfrac{1 - Q_s(l)}{1 - Q_s(h)} = 0.9474$, then using proposition 5.7.8.2 the subset of states I_1 is empty. Then, we have that if $p > 0.9474$ the unique ESS is to take always action l, that is $\rho = 0$.

We observe also on figure 5.1 that taking the pure aggressive action h is only an ESS when $p = 0$. It is somehow intuitive because if the system is such that every player meets always another player, the best strategy to survive is to be aggressive. Related to this comment, we observe also that the probability to choose an aggressive action is monotone decreasing with p.
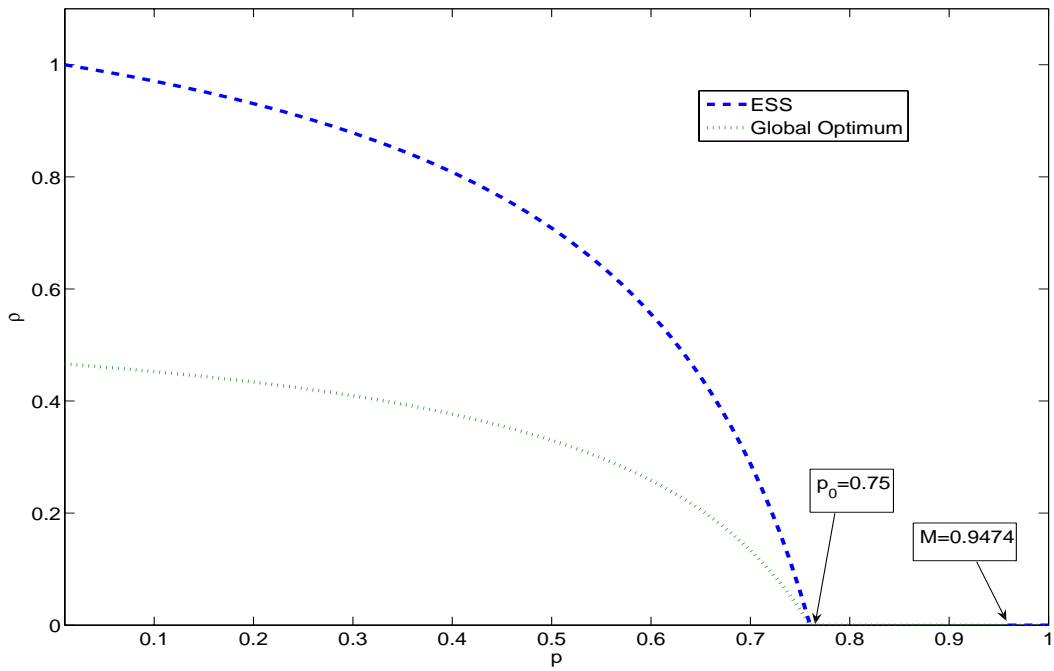
Figure 5.1: *ESS ρ with state-independent actions*

Finally, we also verify the result of the proposition 5.7.8.7 comparing the ESS with the global optimum solution.

5.7.14 Dependent state action

We consider that each player decide to take an aggressive action depending on their level of energy. Using the transition probabilities Q the function f defined by equation 5.28 is strictly

increasing, that means that the best response strategy is a risky policy as u_1. We describe the set of policies $u_1^1, u_1^2, \ldots, u_1^6$ where the policy u_1^i has 5 components which are the action in each energy level from 1 to 5. For example, $u_1^3 = (hhhll)$ means action h in states 1, 2 and 3; and action l in states 4 and 5.
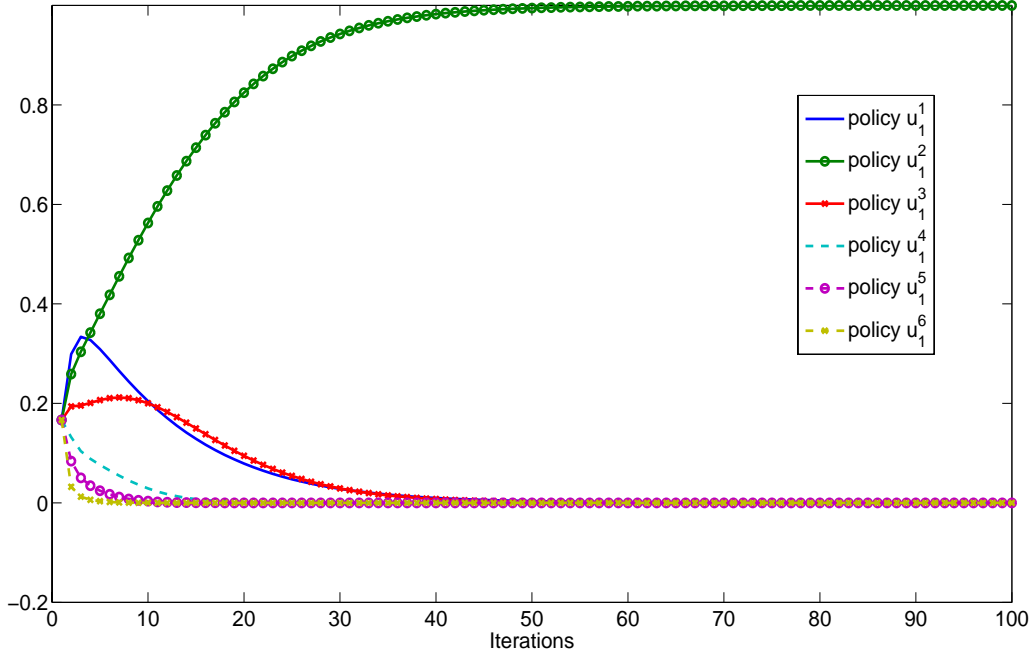

Figure 5.2: *Replicator dynamic with $p = 0.5$*

On figure 5.2, the replicator dynamics converge when $p = 0.5$ to the pure policy u_1^2, i.e. each player takes action *l* only in the full energy level (state 5). On figure 5.3 the replicator dynamics converge to a mixed strategy between policies u_1^1 and u_1^2 where 87.41% of the population uses policy u_1^1 and 12.59% policy u_1^2. When $p = 0.9$, we observe on figure 5.4 the replicator dynamics converge to the pure policy u_1^5 which consists to take action *h* only in state 1. Moreover, we observe that every rest points of the replicator dynamics in the three cases are stable and then there are equilibrium.

Finally, with those examples, we have observed that when the probability to meet another player decreases (means p increases), players become less aggressive and the equilibrium tends to non-aggressive policies.

5.7.15 Notes

We have studied an energy management non-cooperative population game using evolutionary Markov game framework. We have presented a problem considering the stochastic evolutionary games where each player can be in different state during his life, and has the possibility to take several actions in each individual state of each player. Those actions have an impact not only on the instantaneous fitness but also on the future individual's state of the player. Restricting the game to stationary mixed policies, we have determined explicitly the ESS of the stochastic population game. Considering more general dependent state policies, we have obtained
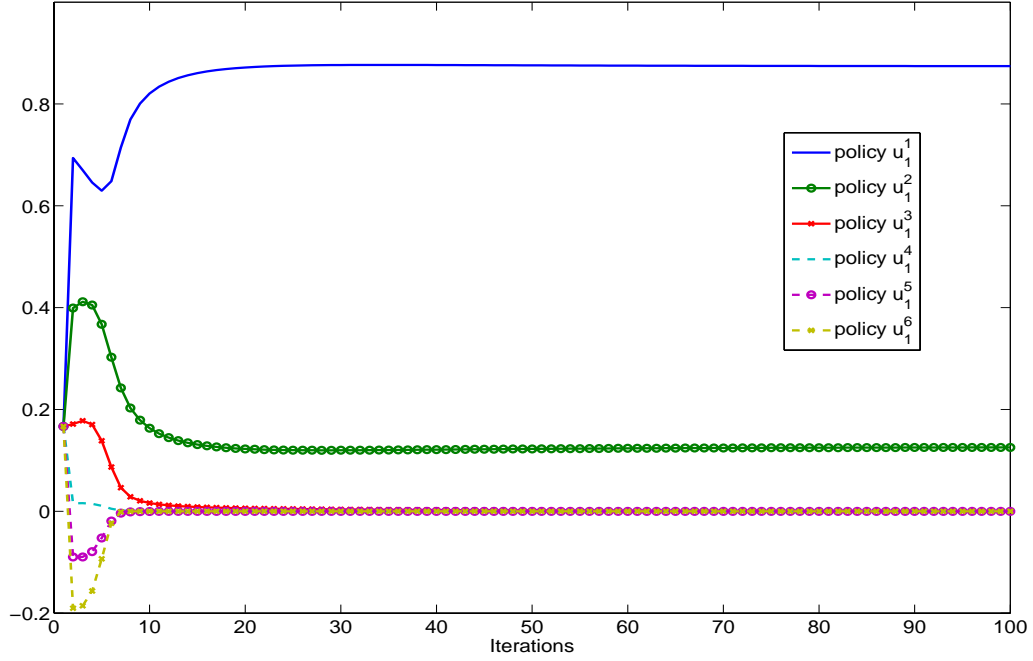
Figure 5.3: Replicator dynamic with $p = 0.1$

the threshold structure of the best response policies and then we have studied numerically the convergence of the replicator dynamics to an ESS.

5.8 Access Control in Solar-Powered Broadband Networks

This Section studies both power control and multiple access control in solar-powered broadband wireless networks. We assume that the mobiles use power storage element, such as rechargeable Solar-powered batteries, in order to have energy available for later use. By modeling the energy-level of Solar-powered batteries as a stochastic process, we study noncooperative interactions within large population of mobiles that interfere with each other through many local interactions. Each local interaction involves a random number of mobiles. The actions taken by a mobile determine not only the immediate payoff but also the state transition probabilities of its battery. We define and characterize the evolutionary stable strategies (ESS) of the stochastic evolutionary game.

Environmental energy is becoming a feasible alternative for many low-power systems, such as wireless sensor and mesh networks. However, this provides an unpredictable and limited amount of energy over time. The power storage elements, such as rechargeable batteries or super-capacitors, become very useful to increase the system lifetime and the system availability. In particular, solar power is made possible with the use of Photovoltaic cells (see Fig. 5.5). Comprised of several layers of material, these cells are able to produce electrical power from exposure to sunlight. Since in many geographic areas, nice weather is not guaranteed and is unpredictable, the nodes should be able to recover from blackout periods caused by the unavailability of energy. In (158), a stochastic model for a solar powered wireless sensor/mesh networks is used to analyze the following QoS measures for several stochastic policies pro-
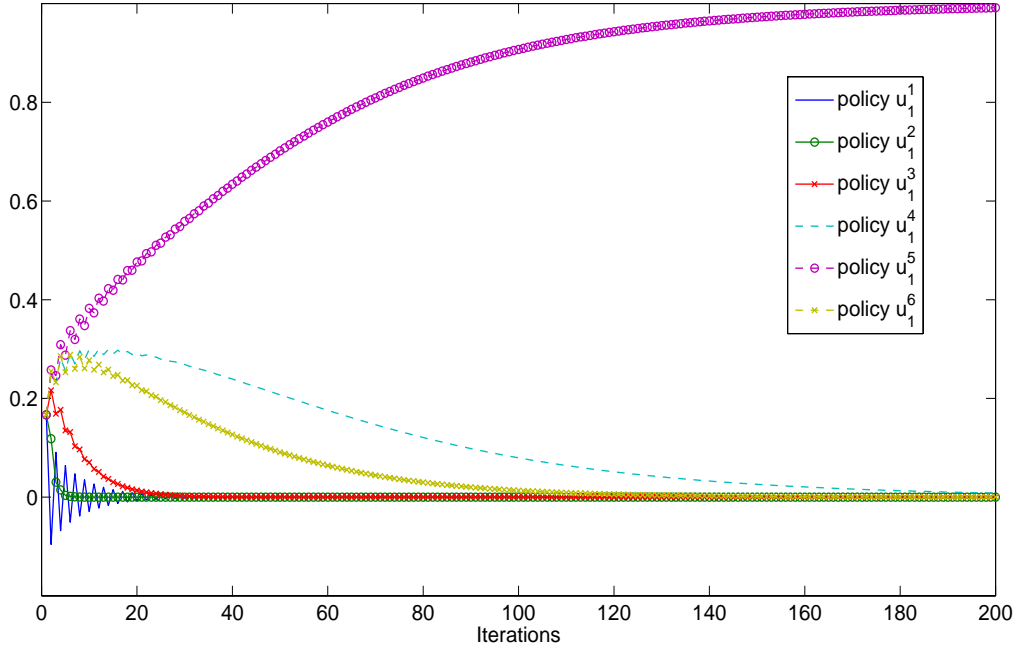
149

Figure 5.4: *Replicator dynamic with $p = 0.9$*

posed: the average battery capacity, the sleeping probability and the average delay.

In this Section, we consider an evolutionary game approach with dynamic rechargeable battery depending on the weather (solar energy). We make use of stochastic evolutionary games in solar-powered broadband wireless networks. There are many local interactions among individuals belonging to large populations of mobiles. The result of the interaction between mobiles depends on their current individual state. From time to time the individual state of a mobile varies. The action choice of mobiles involved in a local interactions as well as their individual states determine the not only the result of the interactions but also the transition probabilities to the other possible individual states. Each individual is thus faced with an Markov Decision Process (MDP) in which it maximizes the expected average cost criterion. Each individual knows only the state of its own MDP, and does not know the state of the other mobiles it interacts with.

The destination of some transmission occasionally may receive simultaneously a transmission from another terminal which results in a collision. It is assumed however that even when packets collide, one of the packets can be received correctly if transmitted at a higher power. As state of the MDP of a user we take its energy level. The immediate fitness (rewards) is the number of successful transmissions. By allowing the mobiles to be equipped with rechargeable solar powered batteries, the mobiles may have infinite life time and the criteria that is maximizing is the limit average Cesaro-type payoff.

Consider the following setting of evolutionary games: there is a large populations of mobiles; each mobile has a finite number of transmission power level available. There are many local interactions at the same time. At each slot, some of the terminals have to take a decision on their transmission power based on their own battery state. At the lowest state of the battery no power is available and the mobile has to wait the time to have good weather for regain some energy. Each player has its individual states set $S = \{0, 1, 2, \ldots, n\}$. Each mobile of the population has a finite action set in each state $s : A(s)$. We assume that there are a random
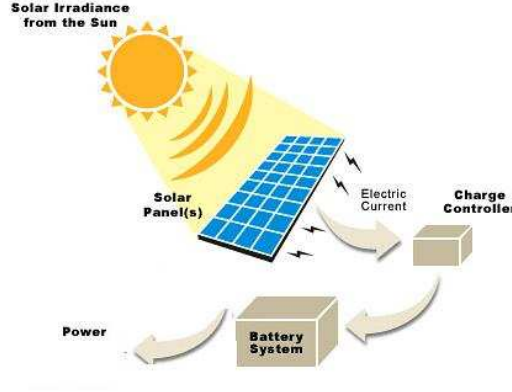
Figure 5.5: Solar-power mechanism

number of interacting mobiles in each local interaction. At time t, each mobile knows its own state s_t and selects an action $a_t \in A(s_t)$. The mobile receives some payoff $r(s_t, a_t, \alpha_t)$ where α_t is the composition of the population at time t (the $j-$th element of $\alpha_t^{S_t}$ represents the fraction of mobiles choosing the action j) and the state of mobile goes to the state s_{t+1} with the probability $q(s_{t+1}|s_t, a_t)$. Each individual is thus faced with an MDP in which it maximizes the expected average fitness criterion. Each individual knows only the state of its own MDP, and does not know the state of the other players it interacts with. The transition probabilities of a player's MDP are only controlled by that player. Let $\Delta(S)$ be the $(|S| - 1)$-dimensional simplex of the Euclidean space $\mathbb{R}^{|S|}$. The set of all action profiles at all states is given by $All = \{(s, a), s \in S, a \in A(s)\}$, then $q : \ All \times \Theta \to \Delta(S)$ is a transition rule between the states where $\Theta = \prod_{s \in S} \Delta(A(s))$, $\Delta(A(s))$ is the set of probability distribution on $A(s)$. The vector $[q(s_{t+1} = 0|s_t, a_t), \ldots, q(s_{t+1} = n|s_t, a_t)]$ satisfies $\forall \, j \in S$, $q(s_{t+1} = j|s_t, a_t) \geq 0$, $\sum_{j \in S} q(s_{t+1} = j|s_t, a_t) = 1$. A state s is absorbing state if $q(s|s, a) = 1$, $\forall \, a \in A(s), \forall \alpha \in \Theta$. We examine the limit average Cesaro-type payoff given a population profile σ and an individual trajectory u,

$$F_n(u, \sigma) = \liminf_{T \longrightarrow +\infty} \mathbb{E}_{u, \sigma} \left(\frac{1}{T} \sum_{t=1}^{T} r(s_t, a_t, \alpha_t) \right)$$

where $\mathbb{E}_{u, \sigma, n}$ denotes the expectation over the probability measure $\mathbb{P}_{u, \sigma, n}$ induced by u, σ on the set of histories endowed with the product $\sigma-$algebra (initial state of the battery is n). Define further

- The subset $\mathcal{U}_S$ of stationary policies; a stationary policy u is a policy in which the probability to choose an action a depends only on the current state s; it is denoted by $u(a|s)$.

- The subset $\mathcal{U}_D \subset \mathcal{U}_S$ of pure or deterministic stationary policies $\mathcal{U}_D$. A policy of this type can be viewed as a function from the states to the actions.

- The set $\mathcal{U}_M$ of mixed strategies: A mixed strategy is identified with a probability γ over the set of pure stationary strategies. It can be considered as first choosing a pure stationary policy u with probability $\gamma(u)$ and then keeping choosing forever the actions according to u. A general mixed strategy is a mixture of behaviorial strategy.

Occupation measure Often we encounter the notion of individual states in evolutionary games; but usually the population size at a particular state is fixed. In our case the choices of

151

actions of an individual determine the fraction of time it would spend at each state. Hence the fraction of the whole population that will be at a given state may depend on the distribution of strategies in the population. In order to model this dependence we first need to introduce the expected average frequency $f_{n,u}(s)$ that an individual spends at a given state s when it follows a strategy u and its initial state at time 1 is n. Moreover, we define $\bar{f}_{n,u}(s,a)$ the expected average frequency during which it is at state s and it chooses action a, $f_{u,n} = \lim_{T_o \infty} \frac{1}{T} \sum_{t=1}^{T} Q^t(u)$ where $Q(u) := [q(s'|s,a,u)]_{s',s,a}$.

We now define the equilibrium concept in the context of evolutionary games. A behaviorial stationary strategy $u = (u_1, \ldots, u_n)$ with $u_i = (x_i, y_i)$ is an evolutionary stable strategy (ESS) if for all strategy mut such that $f_{mut} \neq f_u$ there exists $\epsilon_{mut} > 0$ such that

$$\sum_{s,a}(f_u(s,a) - f_{mut}(s,a))r(s,a,\alpha^\epsilon) > 0 \text{ where } \alpha^\epsilon = (1-\epsilon)\alpha(u) + \epsilon\alpha(mut).$$

Interpretation: Suppose that, initially, the population profile is $\alpha(u)$. Now suppose that a small group of mutants enters the population playing according to a different profile $\alpha(mut)$. If we call $\epsilon \in (0,1)$ the size of the subpopulation of mutants after normalization, then the population profile after mutation will be $\epsilon\alpha(mut) + (1-\epsilon)\alpha(u)$. After mutation, the average payoff of non-mutants will be given by $\sum_{s,a} f_u(s,a)r(s,a,\alpha^\epsilon)$. Analogously, the average payoff a mutant is $\sum_{s,a} f_{mut}(s,a)r(s,a,\alpha^\epsilon)$. That is, u is ESS if, after mutation, non-mutants are more successful than mutants. In other words, mutants cannot invade the population and will eventually get extinct. In case where the payoff function r is linear in last variable α (it is not the case in this model) then this definition is equivalent to $\alpha(u)$ is a strict symmetric Nash equilibrium or $\alpha(u)$ is a strictly better response than $\alpha(mut)$ given that the others mobiles mut. Note that if α is not linear, we can have $\alpha(\epsilon mut + (1-\epsilon)u) \neq \epsilon\alpha(mut) + (1-\epsilon)\alpha(u)$.

5.8.1 Stochastic modeling of the energy levels of Solar-powered battery

we assume that a battery has $n+1$ energy states $S = \{0, 1, \ldots, n\}$. The state 0 corresponds to the state *Empty* and the state n is the *Full* state of the battery. The other states $1, \ldots, n-1$ are intermediary states of the battery. We associate with each mobile a Markov Decision Process (MDP) which represents the transition probabilities between energy levels. Let X_t^i be the energy level of battery at time t. Given a stationary policy σ and a strategy of all the population $\alpha = (\alpha_t)_{t \geq 1}$, the transition probability of the energy level of battery is is described by the (first order, time-homogeneous) Markov process (X_t) where the transition probability law q which is given by

$$q_{sas'} = \begin{cases} 1 - R_{\gamma,s}(a) - Q_{\gamma,s}(a) & \text{if } s' = s-1 \\ R_{\gamma,s}(a) & \text{if } s' = s+1 \\ Q_{\gamma,s}(a) & \text{if } s' = s \\ 0 & \text{otherwise} \end{cases}, \forall 1 \leq s \leq n-1, \forall a_s \in A(s),$$

$$q_{nas'} = \begin{cases} 1 - Q_{\gamma,n}(a) & \text{if } s' = n-1 \\ Q_{\gamma,n}(a) & \text{if } s' = n \\ 0 & \text{otherwise} \end{cases}, \quad q_{0as'} = \begin{cases} \gamma & \text{if } s' = 1 \\ (1-\gamma) & \text{if } s' = 0 \\ 0 & \text{otherwise} \end{cases},$$

where $\gamma \longmapsto R_{\gamma,s}(a) \in [0,1]$ is an increasing function $\forall s, a$ with $R_{0,s}(a) = 0$, $0 \leq R_{\gamma,s}(a) + Q_{\gamma,s}(a) \leq 1$. The factor γ represents the probability to have a "good weather". If γ is zero, the

state 0 is unique absorbing sate and expected lifetime of the battery is finite. For $\gamma \neq 0$, the Markov chain is irreducible i.e., there is only one class, that is, all states communicate with each other for stationary policy.

5.8.2 Battery-state dependent access control in solar-powered system

This subsection studies is a generalization of the *random access game* with unknown number of mobiles, finite state, and three strategies.The channel is ideal for transmission and all errors are due to collision. A mobile can transmit a packet using a power level among three available levels: transmit with high power P_h, transmit with low power P_l or does not transmit 0. We consider a general capture model where a packet transmitted by a mobile is received successfully when if and only if that mobile uses a transmission power which is greater than the power used by the others transmitters at that time slot. Given a population profile $\alpha = (x, y, 1 - x - y)$ with $0 \leq x, 0 \leq y$, $x + y \leq 1$, the expected probability to have a successful transmission :

(i) When the mobile chooses P_l, the reward is the probability that the others mobiles choose zero i.e no others mobiles transmit, i.e.,

$$r(s, P_l, \alpha) = \sum_{k \geq 0} \mathbb{P}(K = k)\,(1 - x - y)^k = \mathbb{E}((1 - x - y)^K) = G_K(1 - x - y)$$

where G_K is the generating function of K.

(ii) When the mobile chooses P_h, the reward is the probability that no other mobiles transmit with the high power i.e.,

$$r(s, P_h, \alpha) = \mathbb{E}_K \sum_{l=0}^{K} \binom{K}{l}(1 - x - y)^l y^{K-l} = \mathbb{E}_K(1 - x)^K = G_K(1 - x). \qquad (5.32)$$

where $\binom{K}{l}$ is the binomial coefficient of K and l.

(iii) When the terminal chooses 0, the reward is zero, i.e., $r(s, 0, \alpha) = 0$

Since there are three strategies P_h, P_l and 0, the population aggressiveness can be described by a process $(x_t, y_t)_{t \geq 1}$ where x_t (resp. y_t) is the fraction of the population using the high power level (low power) at time t. However, for each state $s \neq 0$, the action space becomes $M := \{(x, y),\ x \geq 0,\ y \geq 0,\ x + y \leq 1\}$. A stationary policy of an user is a map $\beta : S \longrightarrow M$. The expected reward of a user when its battery is at the state $s \neq 0$ is then given by

$$\tilde{r}(s, \beta, \alpha) = x_s' r(s, P_l, \alpha) + y_s' r(s, P_l, \alpha),$$

where $\beta(s) = (x_s', y_s')\ \forall s \in S$.

5.8.3 Computing Equilibria and ESS

Nash Equilibria and Pareto optimality

In this subsection, we study the existence and uniqueness of Nash equilibrium in different scenarios :

Two mobiles $K = \delta_2$ The interaction in each slot and each non-empty individual state is described in the following tabular. Mobile 1 chooses a row, mobile 2 chooses a column. The payoff of mobile 1 (resp. mobile 2) is the first (resp. the second) component of the vector payoff.

$1\backslash 2$	”0”	P_l	P_h
”0”	$(0,0)$	$(0,1)^\diamond$	$(0,1)^{\diamond\star}$
P_l	$(1,0)^\diamond$	$(0,0)$	$(0,1)^{\diamond\star}$
P_h	$(1,0)^{\diamond\star}$	$(1,0)^{\diamond\star}$	$(0,0)^\star$

The Nash equilibria are represented by $\star$ and Pareto optimal allocation [3] are represented by $\diamond$.

Three mobiles $K = \delta_3$ The instantaneous reward is presented in the matrix game. Mobile 1 chooses a row, mobile 2 chooses a column and mobile 3 chooses an array M_i.

$$M_1 := \begin{array}{|c|c|c|} \hline (0,0,0) & (0,1,0)^\diamond & (0,1,0)^\star \\ \hline (1,0,0)^\diamond & (0,0,0) & (0,1,0)^\star \\ \hline (1,0,0)^{\diamond\star} & (1,0,0)^{\diamond\star} & (0,0,0)^\star \\ \hline \end{array} \quad M_2 := \begin{array}{|c|c|c|} \hline (0,0,1)^\diamond & (0,0,0) & (0,1,0)^{\diamond\star} \\ \hline (0,0,0) & (0,0,0) & (0,1,0)^{\diamond\star} \\ \hline (1,0,0)^{\diamond\star} & (1,0,0)^{\diamond\star} & (0,0,0)^\star \\ \hline \end{array}$$

$$M_3 := \begin{array}{|c|c|c|} \hline (0,0,1)^{\diamond\star} & (0,0,1)^{\diamond\star} & (0,0,0)^\star \\ \hline (0,0,1)^{\diamond\star} & (0,0,1)^{\diamond\star} & (0,0,0)^\star \\ \hline (0,0,0)^\star & (0,0,0)^\star & (0,0,0)^\star \\ \hline \end{array}$$

Proposition 5.8.3.1. *In any state $s \neq 0$, the one-shot local interaction between $p \geq 2$ mobiles has an infinite (Nash) equilibria, $\binom{p}{1} 2^{p-1} = p2^{p-1}$ of them are Pareto optimal, and a unique symmetric (Nash) which is the strategy h (independently of the state).*

Note that this one-shot game has a unique evolutionary stable state (see (11)).

Proof. The Nash equilibria and Pareto optimality of the local interaction in state $s \neq 0$ can be described as follows:

(i) **Symmetric Equilibrium** It is clear that $r(s, P_h, \alpha) \geq r(s, P_l, \alpha) \geq r(s, 0, \alpha)$, $\forall \alpha$ i.e P_h weakly dominates P_l which weakly dominates 0. Hence, the strategy P_h is an equilibrium. Moreover the best reply to the population profile $\alpha = (x, y, 1 - x - y)$ is to play P_h if $x \neq 1$, and to play any strategy $z \in [0,1]$ if $x = 1$. Thus, $(P_h, P_h, \ldots, P_h)$ is the unique symmetric equilibrium. At the equilibrium $(P_h, P_h, \ldots, P_h)$, the reward of each mobile is zero. Thus, $(P_h, P_h, \ldots, P_h)$ is not Pareto optimal because the allocation obtained at $(P_h, P_l, \ldots, P_l)$ or $(P_h, 0, 0, \ldots, 0)$ Pareto dominates zero.

(ii)**Pure Equilibria**: Fix a mobile m which uses the action P_h. Then any action profile of the others mobiles $b^{-m} \in\in \{0, P_l, P_h\}^{p-1}$ leads to a Nash equilibrium (no mobile can improve its probability of success by deviating unilaterally). In particular, if k $(0 \leq k \leq p-1)$ of the $p-1$ mobiles choose P_l and the $p - k - 1$ others use 0 then no mobile can improve its probability of success by deviating unilaterally. The mobile m has exactly $\sum_{k=0}^{p-1} \binom{p-1}{k} = 2^{p-1}$ pure equilibria in which he/she has successful transmission. By changing the role of m, we get $\binom{p}{1} 2^{p-1} = p2^{p-1}$ pure equilibria with successful transmission. All these pure equilibria are Pareto optimal: if only one terminal uses the high power P_h and the others mobiles use 0 or P_l,, then the mobile with the high power P_h gets the payoff 1 and the others gets the payoff 0.

---

[3]An allocation of payoffs is Pareto optimal or Pareto efficient if there is no other allocation that makes every node at least as well off and at least one node strictly better off.

(iii)**Mixed Equilibria:** Any situation where at least one of the mobiles use the strategy P_h, and other mobiles use an arbitrary mixed strategies, gives a mixed Nash equilibria. The allocation of payoff obtained in these mixed strategy are not Pareto optimal if at least one mobile chooses the strategy P_h with positive probability. $\square$

In the figure 5.6, we plot the best response (power level) of a mobile for a given trajectory of a profile population, with $P_h = 10$ and $P_l = 2$. We observe that the best response is is to use a low power when the population is very aggressive ($x = 1$), and to use the high power in any state $s \neq 0$ when the population is less aggressive $x < 1$. In the figure 5.7, we plot the sojourn time of a mobile to stay in a state as function of its stationary strategy $\beta = (x, y)$ with $s \neq 0, Q_s(0) = 0.95, Q_s(P_l) = 0.65$ and $Q_s(P_h) = 0.25$. We observe that the sojourn time is decreasing function in aggressiveness of that mobile.
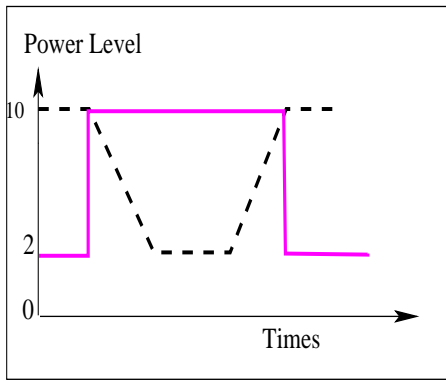
Figure 5.6: *best reply (line) during several slots.*
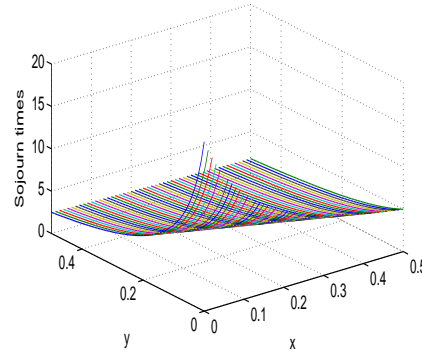
Figure 5.7: *Sojourn time*

Fig.5.8 represents the probability of success in uniform distribution between three mobiles and Fig.5.9 in Poisson distribution with intensity 0.2. The probability to have a successful transmission in any state s decreases with the aggressiveness of the population.
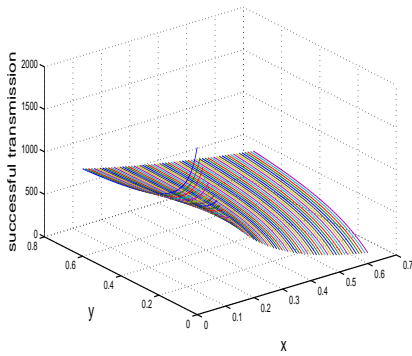

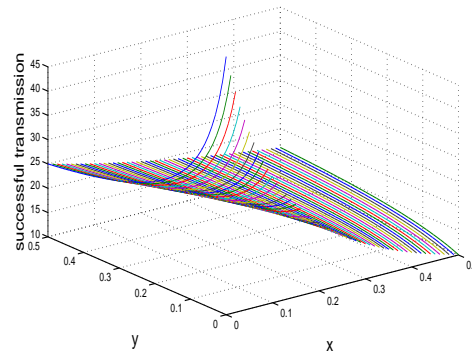**Figure 5.8:** *Payoff – uniform distribution*

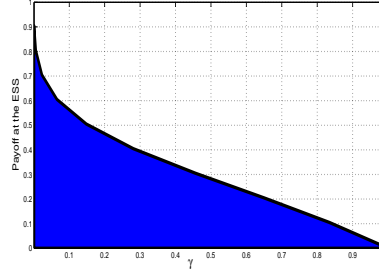**Figure 5.9:** *Payoff – Poisson distribution*

Figure 5.10: *Expected average payoff versus γ.*

Evolutionarily Stable Strategy

Under the condition $\gamma > 0$, we have the following result:

Proposition 5.8.3.2. *The payoff function has a representation in term of occupation measure, i.e.,*
$F_{n,u}(\sigma) = \sum_{s,a} f_u(s,a) r(s,a,\alpha)$ *where* $\alpha = (x,y)$ *is the profile of population under stationary strategy* σ.

Proposition 5.8.3.3. *The fully aggressive strategy* $\bar{P}_h$ *which consists to transmit with the maximum power* (P_h) *in each state* $s \neq 0$ *is an Evolutionarily Stable Strategy.*

Proof. Let mut such that $\alpha'(mut) \neq 1$. For $u = \bar{P}_h$ the time average payoff is $F_u(\alpha^\epsilon) = G_K\left(\epsilon(1 - \alpha'_{P_h}(mut))\right) \sum_{s=1}^n f_{\bar{P}_h}(P_h)$,

and

$$F_{mut}(\alpha^\epsilon) = \sum_{s=1}^n f_{mut}(s, mut_s) \times \left[mut_{P_h} G_K(\epsilon(1 - \alpha'_{P_h}(mut))) + mut_{P_l} G_K(\epsilon(1 - \alpha'_{P_h}(mut) - \alpha'_{P_l}(mut))) \right].$$

Since the generating is increasing, $G_K(\epsilon(1 - \alpha'_{P_h}(mut) - \alpha'_{P_l}(mut))) \leq G_K(\epsilon(1 - \alpha'_{P_h}(mut))$ with equality if and only $\alpha'_{P_l}(mut) = 0$. We deduce that $F_{mut}(\alpha^\epsilon)$ is strictly greater $F_{P_h}(\alpha^\epsilon)$ for all mut such that $\alpha'_{P_h}(mut) < 1$. $\qquad\square$

The Figure 5.10 illustrates the expected long-term payoff versus the parameter γ. We took four states and $R_{s,\gamma} = \gamma$, $R_{s,\gamma} = 0.3 * (1 - \gamma)$.

5.8.4 Power control in clouded weather

We assume now that the solar-powered system is in a clouded weather ($\gamma = 0$, "no sunlight") during a long period (clouded sky, raining time or due to the season). Power storage elements, such as supercapacitors (but finite in practice), in order to have energy available for later use has been proposed. Because of limited capacity (hence energy of the battery), the aggressive terminals (which use the high power) will be rapidly in the state 0 which becomes an absorbing state. The terminals with empty battery need an alternative solution such as external recharge or to buy a new battery. So there is an additional cost to survive in this situation. If the maximum lifetime of the battery (for example using the power "0") is finite then the power control in clouded weather can be modeled by the stochastic evolutionary game with total successful transmission before to reach to the absorbing state 0.

Lemma 5.8.4.1. *The total expected successful transmission during the lifetime of the battery under the strategy u is given by*

$$\frac{1}{1-\gamma} \times 0 + \sum_{s=1}^{n} \frac{u_s(P_l)G_K(1-\alpha_{P_l}-\alpha_{P_h}) + u_s(P_h)G_K(1-\alpha_{P_h})}{1-Q_{0,s}(u_s)}$$

where $Q_{0,s}(u_s) := u_s(P_h)Q_{0,s}(P_h) + u_s(P_l)Q_{0,s}(P_l)$

Proof. This says that the total reward total is the sum over the states of the expected successful transmission times the expected sojourn times spent in this state. The sojourn time in state s under the policy u_s is $\dfrac{1}{1-Q_{0,s}(u_s)}$. This completes the proof. $\qquad\square$

The following Proposition holds in the stochastic evolutionary game with total reward.

Proposition 5.8.4.2. *(i) The strategy "stay quiet" (plays "0" in each state) cannot be an ESS. (ii) A necessary condition for the full aggressive strategy to be an ESS is $G_K(0) = P(K=0) = 0$. Moreover if*

$$G_K(1-\alpha_{P_h}) \sum_{s=1}^{n} \frac{1}{1-Q_{0,s}(P_h)} > \max_{mut \in \{P_h,P_l\}^n} \left\{ \sum_{s=1}^{n} \frac{r(s, mut_s, \alpha)}{1-Q_{0,s}(mut_s)} \right\} \text{ then } \bar{P}_h \text{ is an ESS. (iii) Similarly,}$$

if $G_K(1-\alpha_{P_l}) \sum_{s=1}^{n} \dfrac{1}{1-Q_{0,s}(P_l)} > \max_{mut \in \{P_h,P_l\}^n} \left\{ \sum_{s=1}^{n} \dfrac{r(s, mut_s, \alpha)}{1-Q_{0,s}(mut_s)} \right\}$ *then* $\bar{P}_l$ *is an ESS.*

Proof. The strategy "stay quiet" (play "0" in each state) cannot be an ESS because it is best reply to itself. Hence, the strategy *not transmit* can be invaded by mutations. If $\bar{P}_h$ is an ESS then $G_K(0) \sum_{s=1}^{n} \dfrac{1}{1-Q_{0,s}(P_h)}$ must be greater than $G_K(0) \sum_{s=1}^{n} \dfrac{1}{1-Q_{0,s}(P_l)}$. Since the power consumption is greater with P_h than P_l ($P_h > P_l$), one has,

$$\sum_{s=1}^{n} \frac{1}{1-Q_{0,s}(P_h)} < \sum_{s=1}^{n} \frac{1}{1-Q_{0,s}(P_l)}.$$

This implies that $G_K(0) = P(K=0) = 0$. The other results are immediate by best response conditions. $\qquad\square$

5.8.5 Notes

Thanks to the renewable energy techniques, designing autonomous mobile terminal and consumer embedded electronics that exploit the energy coming from the environment is becoming a feasible option. However, the design of such devices requires the careful selection of the components, such as power consumption and the energy storage elements, according to the working environment and the features of the application. In this paper we have investigated power control interaction based on stochastic modeling of the remaining energy of the battery for each user in each local interaction. We have showed existence of equilibria and conditions for evolutionary stable strategies.

5.9 Wardrop equilibria in nonatomic stochastic power control games

Consider a large number of mobiles terminals controlling their transmission power in a base station. Each mobile has an amount of energy E when its battery is new (typically it is the case

if the battery is new or if the battery is completely recharged). Each mobile implements a power control policy where the transmission power is allowed to depend on the energy level (state) of its battery. The available action depends on the state of the battery. Given the remaining energy of its battery, the mobile have to choose the optimal power level in terms of reward and cost. One of the important element for each mobile is its instantaneous throughput which can be characterized as a function of the signal to interference plus noise ratio (SINR) at the base station where he transmits. The battery is replaced or recharged only when it is completely empty. The cost of new battery or recharge cost is C. The new battery has the same energy of E. The mobile have to control both the power consumption as well as the time at which the batteries are changed or recharged. At each time slot, each mobile is faced to a large number of interacting users which transmit at the same base station. Each battery life-time game corresponds to a stochastic power control game with finite horizon in expectation (absorbing state of battery when the energy is very small). Our aim here is to find the power consumption policy at each state such that all users achieve as high throughput as possible, minimum guarantee (e.g. QoS requirement thresholds) but also to control the battery-state.

Potential Function and Optimization Problems

Lemma 5.9.0.1. *The function* $V : \mathbb{R}_+ \longrightarrow \mathbb{R}$ *defined by* $V(x) = (C^l - C^h)x_s + \log\left(1 + \dfrac{\rho^h x_s + (1 - x_s)\rho^l}{\sigma^2}\right)$

is a potential function.

Proof. $V(x) = -(C^h - C^l)x_s + \displaystyle\int_0^{x_s} \left(\rho^h - \rho^l\right) \dfrac{1}{\sigma^2 + \rho^h t + (1 - t)\rho^l} dt$ where $\rho^h = \dfrac{P_h}{1 - Q_s(h)}$, $\rho_s^l = \dfrac{P_l}{1 - Q_s(l)}$. $\qquad\square$

Interpretation of the Potential Function

The above can be seen as the expression of capacity minus energy consumption.

Lemma 5.9.0.2. *The potential function is strictly concave.*

Proof. The potential function is continuously twice differentiable, and the second derivative is given by $V''(x) = -\dfrac{(\rho^h - \rho^l)^2}{(\sigma^2 + \rho^h x + (1 - x)\rho^l)^2} < 0$. We conclude V is strictly concave (it can be seen also as a sum of strict concave function and the concave function). $\qquad\square$

Proposition 5.9.0.3. *The stochastic power control game has a unique Wardrop equilibrium.*

Proof. The existence of Wardrop equilibrium is guaranteed by maximization of the potential function. Since V is a C^1 concave function, V has a maximum in compact convex set $[0, 1]$ and maximizers are Wardrop equilibria. The uniqueness of the equilibrium is immediate from the strict concavity of the potential function given by Lemma 5.9.0.2. $\qquad\square$

Convergence to Wardrop equilibrium

Proposition 5.9.0.4. *The unique equilibrium is asymptotically stable under the* $\theta-$*Smith dynamics. Moreover, for any initial strategy distribution, the dynamics to converge to the Wardrop equilibrium.*

5.10 Different types of renewable energy

Power control in wireless networks has became an important research area. Since the technology in the current state cannot provide batteries which have small weight and large energy capacity, the design of tools and algorithms for efficient power control is crucial.

Thanks to the renewable energy techniques, designing autonomous mobile terminal and consumer embedded electronics that exploit the energy coming from the environment is becoming a feasible option. However, the design of such devices requires the careful selection of the components, such as power consumption and the energy storage elements, according to the working environment and the features of the application.

Menache and Altman have studied in (142) a battery-energy dependent power control with finite number of mobiles as a dynamic non-cooperative game with power cost assumption. In this model we consider a stochastic population game approach with dynamic rechargeable battery based on renewable energy. Environmental energy is becoming a feasible alternative for many low-power systems, such as wireless sensor/mesh networks. Nevertheless, environmental energy is an exciting challenge. Because of the limited amount of energy over time, the power provided is unpredictable. Power storage elements, such as rechargeable batteries or supercapacitors, in order to have energy available for later use has been proposed. Alternative energy as solar, wind, or nuclear energy, that can replace or supplement traditional fossil-fuel sources, as oil, and natural gas is needed. We refer the reader to (158) for advantageous to use renewable energy in broadband wireless networks such as Wi-Fi, WiMax or mesh networks.

We consider several class of large number of mobiles terminals controlling their transmission power and a distributed base stations. The mobiles with the same type of renewable energy (wind, solar, hydro) are in the same class or subpopulation. Each mobile of the subpopulation p has an amount of energy E^p when its battery is at the full state. Each mobile implements a power control policy where the transmission power is allowed to depend on the energy level (state) of its battery. The available action (reachable base stations) depends on the state of the battery. Given the remaining energy of its battery, the mobile have to choose the optimal power level. One of the important element for each mobile is its instantaneous throughput which can be characterized as a function of the signal to interference plus noise ratio (SINR) at the base station where he/she transmits. The battery is recharged by different techniques of renewable energy (solar-power, wind-power etc). The mobile have to control both the power consumption as well as the level of its battery and its throughput. Each mobile is faced to a non-cooperative stochastic game with individual states with many others mobiles which transmit at the same base station or at the same range. The goal of a terminal is to find jointly the power levels and the base stations such that the terminal achieves as high payoff as possible, minimum guarantee (e.g. QoS requirement thresholds) but also to control the battery-state.

Battery-state transition

We consider the energy reserve of the battery type p, $(X_t^p)_{t \geq 1}$ and power level management as a Markov decision process. For each state $y \neq 0$, the action space is $A^p(y)$ with at least two elements, and $A^p(0)$ has at most one element (empty or singleton). Given a stationary policy σ and a strategy of all the populations the change in energy reserves of the battery type p is described by the (first order, time-homogeneous) Markov process (X_t^p) with the transition law

q^p. $\forall y \neq 0, n^p, \forall a$, the probability of transition $q^p(X_{t+1}^p = y' | X_t^p = y, a)$ is expressed as

$$
\begin{cases}
1 - R_{\gamma^p, y}^p(a) - Q_{\gamma^p, y}^p(a) & \text{if } y' = y - 1 \\
R_{\gamma^p, y}^p(a) & \text{if } y' = y + 1 \\
Q_{\gamma^p, y}^p(a) & \text{if } y' = y \\
0 & \text{otherwise}
\end{cases},
$$

$$
q^p(X_{t+1}^p = y' | X_t^p = n^p, a) = \begin{cases}
Q_{\gamma^p, n^p}^p(a) & \text{if } y' = n^p - 1 \\
1 - Q_{\gamma^p, n^p}^p(a) & \text{if } y' = n^p \\
0 & \text{otherwise}
\end{cases},
$$

and

$$
q^p(X_{t+1}^p = y' | X_t^p = 0, 0) = \begin{cases}
\gamma^p & \text{if } y' = 1 \\
(1 - \gamma^p) & \text{if } y' = 0 \\
0 & \text{otherwise}
\end{cases},
$$

where $\forall y, a, \gamma^p \longmapsto R_{\gamma^p, y}^p(a) \in [0, 1]$ is an increasing function with $R_{0, y}^p(a) = 0, 0 \leq R_{\gamma^p, y}^p(a) + Q_{\gamma^p, y}^p(a) \leq 1$. The factor γ^p represents a function of the probability to have a "good weather"(for example, the sun for the solar-power battery, the wind for wind-power battery) and the probability for battery of type p to go from state 0 to state 1. If γ^p is zero, the state 0 is absorbing. For $\gamma^p \neq 0$ is the chain is communicating.
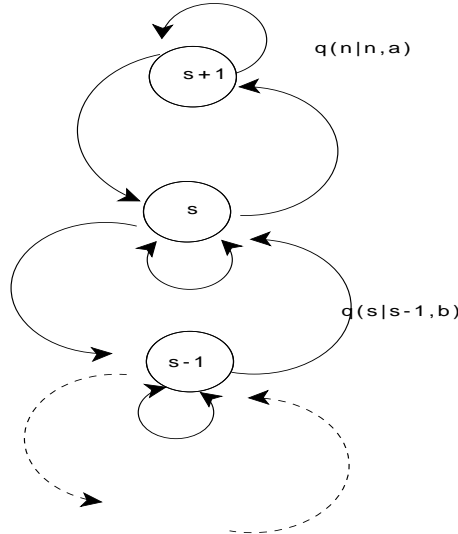
Figure 5.11: *Generic battery state transition rule.*

Note that each user controls the transition state of its battery: q^p is independent of the decision of the other mobiles.

Reward

We focus on utility function based on a simplified version of the signal to noise plus interference ratio (SINR). The battery-state have the property that more energy is available in high state.

Hence, that set of powers in $y+1$ contains the set of power available in y. For example,

$$\varnothing \subset A^p(0) = \{pow_0^p\} \subset A^p(1) = \{pow_0^p, pow_1^p\} \subset$$

$$A^p(2) = \{pow_0^p, pow_1^p, pow_2^p\} \subset$$

$$\dots A^p(n^p) = \{pow_0^p, pow_1^p, \dots, pow_{n^p}^p\}$$

The signal to noise plus interference ratio of a mobile with the battery type p in state y at the position $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ is

$$\text{SINR}_y^p(a, x; \lambda, BS) = \frac{\frac{ag^{pp}}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \kappa I_{own}(x^p) + \kappa I_{other}(x^{-p})},$$

$$a \in A^p(y), p \in P, y \in Y^p = \{0, 1, 2, \dots, n^p\}$$

where

$$I_{own}(x^p) = \sum_{y,b} bg^{pp}x^p(y,b)h^{pp}$$

$$I_{other}(x^{-p}) = \sum_{k \neq p} \sum_{y,b} g^{kp}bx^k(y,b)h^{kp},$$

where

$$h^{kp} = \int_{\lambda \in D} \frac{d\mu^{k,BS}}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}$$

$x^j(y,a)$ is the fraction of the sub-population j in state y with the power level a, N_0 is the power of the thermal background noise, $\mu^{p,BS}$ is the distribution of mobiles (in the 3-dimensional space) with the battery type p around the base station BS, $D \subseteq \mathbb{R}^3$ is the domain (geographical placement of base stations and mobiles) and α is the path-loss and κ is the inverse of the processing gain of the system, it weights the effect of interferences, depending on the orthogonality between codes used during simultaneous transmissions. The coefficient κ is equal to 1 in a narrow band system, and is smaller than 1 in a broadband system that uses CDMA. The instantaneous expected reward $r_{y,a}^p(x)$ of an user in state y is expressed as

$$\int_{\lambda \in D} f\left(\text{SINR}_y^p(a, x, \lambda, BS)\right) d\mu^p(\lambda)$$

where f is a non-decreasing function with $f(0) = 0$. ϵ is a positive parameter (to eliminate of continuity problem at zero) and the g^{ij} are positive gain parameters. The $3-$dimensional vector (x_0, y_0, z_0) describes the position of the base station BS in $\mathbb{R}^3$

Computing the interference term in presence of continuum of users

In order to compute explicitly the SINR term, we first need the following lemma:
Lemma 5.10.0.5.

$$\nu \geq 1, \; b_\nu = \int_0^{+\infty} \frac{1}{(1+x^2)^\nu} dx = \begin{cases} \dfrac{\pi}{2} & \text{if } \nu = 1 \\ \dfrac{\sqrt{\pi}}{2} \dfrac{\Gamma(\nu - \frac{1}{2})}{\Gamma(\nu)} & \text{if } \nu > 1. \end{cases}$$

where Γ is the Euler function $\Gamma(x) = \int_0^{+\infty} e^{-t}t^{x-1}\, dt$.

Proof. For $\nu = n$ a positive integer, the polynomial $(1 + z^2)^n$ has two zeros $z = \pm i$ each zero with the order n. Consider the circuit $C_{R'} = [-R', R'] \cup \{R'e^{i\theta}, 0 \leq \theta \leq \pi\}$. Since the complex function $z \in \mathbb{C} \longrightarrow \dfrac{1}{(1 + z^2)^n}$ has no zero on the circuit $C_{R'}$, Using residue's theorem of complex analysis, we obtain the following result: $\displaystyle\int_{x \geq 0} \dfrac{1}{(1 + x^2)^n} dx = \pi i Res(\xi(z), i)$. The residue of ξ around $z = i$, a pole of order n, can be found by the formula:

$$Res(\xi, i) = \frac{1}{(n-1)!} \lim_{z \longrightarrow i} \left(\frac{d}{dz}\right)^{n-1} [(z - i)^n \xi(z)]$$

Thus, $b_n = \dfrac{\sqrt{\pi}}{2} \dfrac{\Gamma(n - \frac{1}{2})}{\Gamma(n)}$. We then use the extension of the Euler function Γ on the positive real axis. $\qquad\square$

From the lemma 5.10.0.5, we derive immediately that, $n \geq 2$, $\displaystyle\int_0^{+\infty} \dfrac{x^2}{(1 + x^2)^n} dx = b_{n-1} - b_n$

Proposition 5.10.0.6. $h^{jp} = \dfrac{4\pi^2}{\epsilon^{2\alpha - 3}} (b_{\frac{\alpha}{2} - 1} - b_{\frac{\alpha}{2}})$

Proof. Using spherical coordinates from cartesian coordinates by the transformation

$$\begin{cases} \lambda_1 = r \sin\theta \cos\phi \\ \lambda_2 = r \sin\theta \sin\phi \\ \lambda_3 = r \cos\theta \end{cases}$$

and the volume element $r^2 dr \sin\theta d\theta d\phi$, one has,

$$h^{jp} = 4\pi^2 \int_0^{+\infty} \frac{r^2}{(\epsilon^2 + r^2)^{\frac{\alpha}{2}}} dr = \frac{4\pi^2}{\epsilon^{2\alpha - 3}} \int_0^{+\infty} \frac{r^2}{(1 + r^2)^{\frac{\alpha}{2}}} dr$$

i.e $h^{jp} = \dfrac{4\pi^2}{\epsilon^{2\alpha - 3}} (b_{\frac{\alpha}{2} - 1} - b_{\frac{\alpha}{2}})$ $\qquad\square$

Proposition 5.10.0.7. *The highest payoff that a mobile with the battery type p can obtain against any strategies of others mobiles in the one-shot power control game is given by*

$$\bar{v}_y^p = \int_D f \left(\frac{\frac{u_y^p g^{pp}}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \sum_j \sum_y h^{jp} u_y^j m_y^j g^{jp}} \right) d\mu^p$$

where u_y^p is the maximum power level available in the battery-type p in state y.

Proof. Since the payoff decreases when the others players increases their power levels (in average), the minmax point is obtained when they uses their high powers. The maximum payoff that a mobile with the battery type p can obtain against any strategies of others mobiles is then given by

$$\bar{v}_y^p = \max_{a \in A^p(y)} \int_D f \left(\frac{\frac{a g^{pp}}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \sum_j \sum_y h^{jp} u_y^j m_y^j g^{jp}} \right) d\mu^p$$

$$= \int_D \max_{a \in A^p(y)} f \left(\frac{a g^{pp} \frac{}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \sum_j \sum_y h^{jp} u_y^j m_y^j g^{jp}} \right) d\mu^p$$

$$= \int_D f \left(\max_{a \in A^p(y)} \frac{a g^{pp} \frac{}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \sum_j \sum_y h^{jp} u_y^j m_y^j g^{jp}} \right) d\mu^p .$$

This completes the proof. $\qquad\square$

Theorem 5.10.1. *Each mobile with the battery type p can guarantee the payoff*

$$\sum_{y \neq 0} \Pi_y^p \bar{v}_y^p$$

for all $\gamma^p > 0$, where $\Pi_y^p = \lim_{t \longrightarrow \infty} \mathbb{P}(X_t^p = y)$ is the probability to be in state y under the maximum power strategy.

Proof. Π_y^p is the frequency of visit of the battery state in y. From Proposition 5.10.0.7, each mobile with the battery type p can obtain at least $\bar{v}_y^p$ against any strategies of others mobiles. Each mobile of subpopulation can then obtain at least $\sum_y \Pi_y^p \bar{v}_y^p$ which is an equilibrium payoff. This completes the proof. $\qquad\square$

Notes

Our stochastic population game model with multi-class of of users can be extended to the local common resources states case. The model including the state of local resources allows us to take into account the local resources for which users are interacting cooperatively or non-cooperatively.

In the third part of this manuscript we develop a new class of mean field games. Using stochastic processes results we will be able to drop the assumptions of *stationary regime* and derive weak convergence of the population profile process when the size of population goes to infinity. The two chapters show that in general, the validity of standard analysis such as fixed point analysis, stationarity and ergodicity conditions need to be justified. The chapters are based to our publications (210; 208; 218).

Part III

Mean Field Limits

Chapter 6

Mean field asymptotics of population games

We consider evolutionary games in *finite* population, in which each individual in a interacts with other randomly selected players. The types and actions of each player in an interaction together determine the instantaneous payoff for all involved players. They also determine the transition probabilities to use actions. We provide a rigorous derivation of the asymptotic behavior of this system as the size of the population grows. We show that the large population asymptotic of the microscopic model is equivalent to a macroscopic evolutionary game in which a local interaction is described by a single player against a population profile. We derive various classes of *evolutionary game dynamics*.

6.1   Introduction

We consider a population with finite number of players in which a small frequent interactions occurs among random number of players that are selected at random. Each player is thus involved in infinitely many interactions with other randomly selected players. Each interaction in which a player is involved can be described as one stage of an evolving game. The actions of the players at each stage determine an immediate payoff (also called *fitness* in behavioral ecology) for each player as well as the transition rates of a Markov chain associated with each player. The transition rate is determined of the change of actions and the system state.

This model extends the basic pairwise interaction model in evolutionary games by introducing a random number of interacting player and the rate of transition for several players which has opportunities to changes its action. At each time slot, a one-shot game with unknown number of players replaces the matrix games, Instead of a choice of a (possibly mixed) action, a player is now faced with the choice of decision rules (called strategies) and revision of theses strategies that determine what actions should be chosen at a given interaction for given present and past observations.

This model with a finite number of players, called *mean field interaction* (43), become more complex to analyze if the number of players grows (because a huge action profile space is required to describe all the of players). Then taking the asymptotic as the number of players grows to infinity, the whole behavior of the population is replaced by a deterministic limit that

represents the system's state or population profile, which is fraction of the population at each type that use a given action.

In this paper we study the asymptotic behavior of the system in which the population profile evolves in time. For large number of players, under mild assumptions (see Section 6.2), the mean field converges to a deterministic measure that satisfies a non-linear ordinary differential equation for a fixed stationary strategy. We show that the mean field interaction is asymptotically equivalent to an evolutionary game with random number of interacting players. At a given time, any given player sees an equivalent game against a collective of players whose state evolves according to an ordinary differential equation (ODE) which we explicitly compute. In addition to providing the exact limiting asymptotic, the ODE approach provides tight approximations for fixed large N of the random process. We derive several evolutionary game dynamics and learning-based dynamics such as Smith dynamics, replicator dynamics, logit dynamics, Brown-von Neumann-Nash dynamics etc. We give sufficient conditions for convergence to equilibria and asymptotic results for non-convergent dynamics. We then apply to Aloha-based protocols with power control, evolution of technologies and resource selection in heterogenous networks.

6.2 The setting

Consider the following interaction model consisting of :

- One population of players with size $N \in \mathbb{N}$.

- Each player has its own state. A state of a player is has two components: the type and the action. The state of player j at time t is denoted by $X_j^N(t) = (\theta_j, A_j^N(t))$. Each player j has to make type-dependent sequential decisions in interaction where he will be involved. For each type $\theta \in \Theta$ (finite), there is finite set of possible actions $a \in \mathcal{A}_\theta$.

$$\mathcal{S} = \{(\theta, a) \mid \theta \in \Theta, a \in \mathcal{A}_\theta\}$$

Let $d = \sharp \mathcal{S}$ be the total number of states.

- Time is discrete and takes its value in the set

$$\frac{\mathbb{N}}{N} := \{0, \frac{1}{N}, \frac{2}{N}, \ldots\}.$$

The *global detailed description* of the system at time t is $X^N(t) = (X_1^N(t), \ldots, X_N^N(t))$. We assume that $X^N(t)$ is Markov.

- Introduce the transition kernel of $X^N(t)$, $\mathcal{K}^N(x_1, \ldots, x_N; x_1', \ldots, x_N')$

$$= \mathbb{P}\left(X_1^N(t + \frac{1}{N}) = x_1', \ldots, X_N^N(t + \frac{1}{N}) = x_N' | X_1^N(t) = x_1, \ldots, X_N^N(t) = x_N\right) \qquad (6.1)$$

- (Indistinguishability per type) We assume that all players with the same type are indistinguishable i.e play the same role when facing an interaction. This implies that the transition kernel of $X^N(t)$ is invariant with respect to changes in the labeling of players. In other words, for any permutation σ of the index set $\{1, 2, \ldots, N\}$:

$$\mathcal{K}^N(\theta_1, a_1, \ldots, \theta_N, a_N, a_N; \theta_1, a_1', \ldots, \theta_N, a_N')$$
$$= \mathcal{K}^N(\theta_1, a_{\sigma(1)}, \ldots, \theta_N, a_{\sigma(N)}; \theta_1, a_{\sigma(1)}', \ldots, \theta_N, a_{\sigma(N)}') \qquad (6.2)$$

Thanks to indistinguishability per type assumption we will be able to reduce the study of

$$X^N(t) = (X_1^N(t), X_2^N(t), \ldots, X_N^N(t))$$

to a lower dimension one. This will be used to describe the dynamics of the system.

- Define $M^N(t)$ to be the current population profile or occupancy measure i.e

$$M_{\theta,x}^N(t) = \frac{1}{N_\theta} \sum_{j=1}^{N_\theta} \delta_{\{X_j^N(t)=\theta,a\}}.$$

  where N_θ is the number of players with the type θ. For simplicity we denote by N indeed of N_θ. At each time t, $M^N(t)$ is in the finite set $\{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}^{\sharp S}$. and $M_{\theta,a}^N(t)$ is fraction of population of type θ (also called subpopulation θ) using action a at time t.

- Local interaction: At time t, a random set of players $\mathcal{B}^N(t) \subset 2^{\{1,2,\ldots,N\}}$ are randomly selected from N players for an interaction. Each player j in the set $\mathcal{B}^N(t)$ takes part in a one-shot game at time t. Each player selected has a revision opportunity of its strategies. The state of player j can change only if j is selected for an interaction.

  The instantaneous payoff of a player j when he/she moves from the state $X_j^N(t)$ at time t to $X_j^N(t + \frac{1}{N})$ at time $t + \frac{1}{N}$ after an interaction is given by

$$g_j^N(X_j^N(t), X_{\mathcal{B}^N(t)\setminus j}^N(t), X_j^N(t + \frac{1}{N}), X_{\mathcal{B}^N(t)\setminus j}^N(t + \frac{1}{N}))$$

  We assume that for any permutation σ of the index set $\{1, 2, \ldots, N\}$:

$$g_j^N(\theta_1, a_1, \ldots, \theta_N, x_N, ; x_1', \ldots, x_N')$$
$$= g_{\sigma(j)}^N(\theta_1, a_{\sigma(1)}, \ldots, \theta_N, a_{\sigma(N)}; x_{\sigma(1)}', \ldots, x_{\sigma(N)}') \tag{6.3}$$

  Note that the game is not necessarily symmetric since the set of strategies and the payoff function can be different between the types.

  The instantaneous expected payoff of a player j in state x at time t is given by

$$r^N(x_j(t), m(t)) = \mathbb{E}\left[g_j^N\left(X_j^N(t), X_{\mathcal{B}^N(t)\setminus j}^N(t), X_j^N(t + \frac{1}{N}), X_{\mathcal{B}^N(t)\setminus j}^N(t + \frac{1}{N}) \mid \Omega^N\right)\right]$$

  where $\Omega^N := \{j \in \mathcal{B}^N(t), X_j^N(t) = x_j(t), M^N(t) = m(t)\}$. We assume that the payoff :

$$r^N : \quad \mathbb{R}^{\sharp S} \longrightarrow \mathbb{R}^d$$
$$m \longmapsto (r^N(x, m))_x$$

  is continuous and bounded. This payoff will represent in averaging over the random number $b(t) = \sharp \mathcal{B}^N(t)$ of interacting players. We will explicit the dependency in next section.

It follows from the indistinguishability per type assumptions that

1. $M^N(t)$ is Markov.

2. $(X_j^N(t), M^N(t))$ is Markov. This means that the evolution of one specific player $X_j^N(t)$ depends on the other players only through the occupancy measure $M^N(t)$.

6.2.1 Revision of strategies

At every time slot, a random number of players $\mathcal{B}^N(t)$ is picked at random; then, given that $\sharp\mathcal{B}^N(t) = k$, an ordered sequence of k players $(j_1, j_2, \ldots, j_k)$ is picked at random uniformly among the

$$N(N-1)(N-2)\ldots(N-k+1)$$

possible ones. Then, the collection of k players does a transition independent of the past according to

$$L^N_{(x_1,\ldots,x_k;x'_1,\ldots,x'_k)}(m,k) := P\Big(X^N_{j_l}(t+\frac{1}{N}) = x'_l, l = 1, \ldots, k \,|\zeta^N, \sharp\mathcal{B}^N(t+\frac{1}{N}) = k\Big)$$

$$\zeta^N := \{X^N_{j_l}(t) = x_l, \, l = 1, \ldots, k, \, M^N(t) = m\}$$

$$J^N_k(m) := \mathbb{P}(\sharp\mathcal{B}^N(t) = k|M^N(t) = m) \tag{6.4}$$

6.2.2 Instantaneous payoffs

The instantaneous payoff of a player is given by:

$$r^N(x,m) = \mathbb{E}_{\mathcal{B}}\Big(r^{N,\mathcal{B}}(x, M^N|M^N(t) = m)\Big) = \sum_{k=1}^{N} r^{N,k}(x,m)J^N_k(m)$$

with

$$\begin{aligned}
r^{N,k}(x,m) &= \sum_{x_2,\ldots,x_k} g^{N,k}(x, x_2, \ldots, x_k)\left(\frac{Nm_{x_2} - 1_{x=x_2}}{N-1}\right) \\
&\times \left(\frac{Nm_{x_3} - 1_{x_3=x} - 1_{x_3=x_2}}{N-2}\right) \\
&\ldots \left(\frac{Nm_{x_k} - \sum_{l=2}^{k-1} 1_{x_k=x_l} - 1_{x_k=x}}{N-k}\right)
\end{aligned}$$

where $g^{N,k}(x, x_2, \ldots, x_k)$ is the payoff of the player state in state x and the $k-1$ other players respectively with type-action $x_2 = (\theta_2, a_2), \ldots, x_k = (\theta_k, a_k)$.

6.3 Convergence to differential equation

The drift is defined as the expected change in M^N in one time slot:

$$f^N(m,t) = \mathbb{E}\left(M^N(t+\frac{1}{N}) - M^N(t)|M^N(t) = m\right) = \sum_{k\geq 1} f^{N,k}(m)J^N_k(m)$$

with

$$
\begin{aligned}
f^{N,k}(m) &= \mathbb{E}\left(M^N(t+\frac{1}{N}) - M^N(t)|M^N(t) = m, \sharp \mathcal{B}^N(t+\frac{1}{N}) = k\right) \\
&= \sum_{x_1',\ldots,x_k'} \sum_{x_1,\ldots,x_k} m_{x_1}\left(\frac{Nm_{x_2} - 1_{x_1=x_2}}{N-1}\right) \\
&\quad \times \left(\frac{Nm_{x_3} - 1_{x_3=x_1} - 1_{x_3=x_2}}{N-2}\right) \\
&\quad \ldots \times \left(\frac{Nm_{x_k} - \sum_{l=1}^{k-1} 1_{x_k=x_l}}{N-k}\right) \\
&\quad \times \frac{1}{N} L^N_{x;x'}(m,k)\left(\sum_{l=1}^{k}(\vec{e}_{x_l'} - \vec{e}_{x_l})\right)
\end{aligned}
$$

where $x_l = (\theta_l, a_l); x_l' = (\theta_l, a_l')$.

We are interested in the case where the number N of players is large compared to the *expected number* of possible number of simultaneous interacting players.

The following diagram illustrates the problem.

Figure 6.1: *Non-Commutative Diagram.*

We study the mean field convergence i.e $\lim_{N \to \infty} M^N$ weakly or better almost surely where for any time t, the sequence of measures $M^N(t) \in \Delta(\mathcal{S})$ satisfying some dynamical systems. Necessarily this requires the weak convergence of $M^N(0)$ to $m(0)$. For $B \subseteq S$, define

$$
M^N(t)(B) := \frac{1}{N}\sum_{j=1}^{N} \delta_{\{X_j^N(t) \in B\}}
$$

We assume that the following limits exist as N goes to infinity. (a) $L^N_{x;x'}(m,k) \to L_{x;x'}(m,k)$, (b) $r^N(x,a,m) \to r(x,a,m)$ and (c) $J_k^N(m) \to J_k(m)$.

Proposition 6.3.0.1. *For every k, the random measure*

$$
M^{N,k} = \frac{1}{N(N-1)(N-2)\ldots(N-k+1)} \sum_{\substack{j_1,j_2,\ldots,j_k \\ distinct}} \delta_{(X_{j_1}^N,\ldots,X_{j_k}^N)}
$$

converge weakly to a deterministic measure on $\mathcal{S}^k$. In particular for $k = 1$ the population profile

$$
M^{N,1} = M^N = \frac{1}{N}\sum_{j=1}^{N} \delta_{X_j^N}
$$

converges in distribution.

Proposition 6.3.0.2. *At each time t, the marginal measure* $M^N(t) = \frac{1}{N} \sum_{j=1}^{N} \delta_{X_j^N(t)}$ *converges to the solution of the ODE*

$$\dot{m}(t) = f(m(t)) \tag{6.5}$$

with initial condition $\lim_{N \longrightarrow \infty} M^N(0) = m_0.$

6.4 From mean field interactions to evolutionary games

Proposition 6.4.0.3 (Asymptotically equivalent game). *When N goes to infinity, the mean field interaction model with random set* $\mathcal{B}^N(t)$ *of players is equivalent to an evolutionary game in which a local interaction at time t is described by*

- *each player is facing a population profile* $m(t)$,
- *the instantaneous payoff of a player with the type* θ *and action a is*

$$r_{\theta,a}(m(t)) := \lim_{N \longrightarrow \infty} r_{\theta,a}^N(m(t)|X_j^N = (\theta, a), M^N(t) = m(t))$$

The Proposition 6.4.0.3 says that for large N, all interactions can be replaced to any one player with an average or effective interaction and an appropriate payoff function. This reduces any large population problem into an effective one-player problem.

At the asymptotic regime we have the following: At each time t, a random number $b(t) = k$ of players has opportunity to revise their strategies. The expected rate to move from $x_1, \ldots, x_k$ to $x_1', \ldots, x_k'$ is $\sum_{k \geq 1} L_{x,x'}(m(t), k) J_k(m(t))$

6.4.1 Equilibrium state analysis

We say that a population profile (the vector of frequencies) $\vec{m}^*$ is a stationary *equilibrium state* if for all vector m of frequencies, the following variational inequality holds:

$$\langle \vec{m}^* - \vec{m}, r(\vec{m}^*) \rangle \geq 0,$$

where $r(\vec{m}) = (r_x(m))_{x \in \mathcal{S}}$, and $\langle ., . \rangle$ is the inner product in the d dimensional Euclidean space $\mathbb{R}^d$. We recall that given a closed convex and non-empty set Δ_d and a continuous function r, solving the variational inequality problem defined by Δ_d and r means finding a vector $\vec{m}^*$ such that

$$\langle \vec{m}^* - \vec{m}, r(\vec{m}^*) \rangle \geq 0$$

Geometrically, the task of variational inequality problem is to find a vector $\vec{m}^*$ such that the image of $\vec{m}^*$ under the reward function r will form an angle more than or equal to 90ř with any vector with tail $\vec{m}^*$ and head $\vec{m} \in \Delta_d$.

Proposition 6.4.1.1 (Existence). *For any distribution* $\mathcal{B}$ *and any continuous function r on the non-empty, convex and compact subset* Δ_d *of the Euclidean space* $\mathbb{R}^d$, *the evolving game has a least one stationary equilibrium state.*

Sketch of the Proof. We first show that $\vec{m}^* \in \Delta_d$ is a solution of the variational inequality problem if and only if for any $\varsigma > 0$, $\vec{m}^*$ is a fixed point of the map $\Pi_{\Delta_d}(I + \varsigma \vec{r})$ from Δ_d to Δ_d.

Multiplying the inequality $\langle \vec{m}^* - \vec{m}, r(\vec{m}^*) \rangle \geq 0$ by $\varsigma > 0$, and adding $\langle \vec{m}^*, \vec{m} - \vec{m}^* \rangle$ to both sides of the resulting inequality, one obtains

$$\langle \vec{m} - \vec{m}^*, \vec{m}^* - [\vec{m}^* + \varsigma r(\vec{m}^*)] \rangle \geq 0.$$

Recall that the projection map on Δ_d which is convex and closed set is characterized by:

$$z \in \mathbb{R}^d, \ z' = \Pi_{\Delta_d} z \iff \langle z' - z, m - z' \rangle \geq 0, \ \forall m \in \Delta_d$$

Thus,

$$\vec{m}^* = \Pi_{\Delta_d}(\vec{m}^* + \varsigma r(\vec{m}^*))$$

According to Schauder's fixed point theorem, given a map $\alpha : \Delta_d \longrightarrow \Delta_d$, with α continuous, there is at least one $m \in \Delta_d$, such that $\vec{m} = \alpha(\vec{m})$. Observe that since the projection Π_{Δ_d} and $(I + \varsigma \vec{r})$, are both continuous, $\Pi_{\Delta_d}(I + \varsigma \vec{r})$ is also continuous by composition.

It follows from convexity and compactness of Δ_d and the continuity of $\Pi_{\Delta_d}(I + \varsigma \vec{r})$ that the Brouwer-Schauder fixed point theorem can be applied to the map $\Pi_{\Delta_d}(I + \varsigma \vec{r})$. We conclude that at least one stationary equilibrium state exists.

$\square$

6.4.2 Dynamics of evolving games

The general dynamics obtained from the mean field interaction is given

$$\frac{d}{dt} m_{\theta,a}(t) = f_{\theta,a}(m(t))$$

where the drift limit can be expanded as

$$f_y(m(t)) = \sum_{k \geq 1} J_k(m(t)) f_y^k(m(t)),$$

$$f_y^k(m(t)) = \sum_{x_1,\ldots,x_k} \left(\prod_{l=1}^k m_{x_l}(t) \right) \left(\sum_{j=1}^k \eta_{x;y}^j(m(t),k) \right) - m_y(t) \left(\sum_{j=1}^k \sum_{x_{-j}} \left(\prod_{l=1,l \neq j}^k m_{x_l}(t) \right) \eta_{y,x_{-j}}^t \right),$$

where

$$y = (\theta,a), \ \eta_{x;y}^j(m(t),k) = \sum_{x'_{-j}} L_{x;(y,x'_{-j})}(m(t),k),$$

$$(y, x'_{-j}) = (x'_1, x'_2, \ldots x'_{j-1}, y, x'_{j+1}, \ldots, x'_k)$$

and

$$\eta_{(y,x_{-j})}^t(m(t),k) = \sum_{x'} L_{(y,x_{-j});x'}(m(t),k).$$

6.4.3 Single player selected per time slot

Suppose now that at each time slot, only one player between the N is randomly selected and has a chance to change its strategy, i.e. $\sharp \mathcal{B}^N = 1$ w.p. 1.

Thus H2 and H3 are automatically satisfied. The resulting ODE becomes

$$\frac{d}{dt} m_x(t) = \sum_{x'} m_{x'} L_{x,x'}(m) - m_w \sum_{x'} L_{x,x'}(m)$$

The term $\sum_{x'} m_{x'} L_{x',x}(m)$ is the *incoming flow* in to x' and the *outgoing flow* from x is $m_x \sum_{x'} L_{x,x'}(m)$.

We then obtain a large class of standard evolutionary game dynamics.

$\gamma-$**Smith dynamics** $\gamma-$Smith dynamics are obtained for $L_{\theta,b;\theta,a}(\vec{m}) = \lambda \max(0, r_{\theta,a}(\vec{m}) - r_{\theta,b}(\vec{m}))^\gamma$

$$\frac{d}{dt} m_{\theta,a} = \lambda \sum_b m_{\theta,b} \max(0, r_{\theta,a}(m) - r_{\theta,b}(m))^\gamma - \lambda m_{s,a} \sum_b \max(0, r_{s,b}(m) - r_{s,a}(m))^\gamma$$

Replicator dynamics Replicator dynamics is obtained for $L_{\theta,b;\theta,a}(\vec{m}) = \lambda m_{\theta,a} \max(0, r_{\theta,a}(m) - r_{\theta,b}(m))$

$$\frac{d}{dt} m_{\theta,a} = \lambda \sum_b m_{\theta,b} m_{s,a} \max(0, r_{\theta,a}(m) - r_{\theta,b}(\vec{m})) - \lambda m_{\theta,a} \sum_b m_{\theta,b} \max(0, r_{\theta,b}(m) - r_{\theta,a}(m))$$

$$= \lambda m_{\theta,a} \left[r_{\theta,a}(m) - \sum_b m_{\theta,b} r_{\theta,b}(m) \right]$$

The following result gives possible rest points for linear payoffs (in the population profile) of the replicator dynamics even when the ODE does not necessarily converge.

Convergence of time averages under the replicator dynamics Suppose that the replicator dynamics has a solution $(m_{\theta,a}(t))$ in the relative interior of the $(d-1)-$simplex i.e there is an $\epsilon > 0$ such that $\epsilon \le m_{\theta,a}(t)$ for all $t \ge 0$. Then

$$\lim_{T \longrightarrow +\infty} d_{RP} \left(\frac{1}{T} \int_0^T m_{\theta,a}(t) \right) = 0$$

where RP is the set of rest point of replicator dynamics and

$$d_{RP}(m) = \min_{m^* \in RP} \| m - m^* \| .$$

This says that the time average value of the population profile over the time interval $[0, T]$ converges to the set of interior rest point of the replicator dynamics as T goes to infinity.

Logit dynamics Logit dynamics (called also *regularized version of best response dynamics* or *Boltzman dynamics*) when

$$\epsilon > 0, L_{\theta,b;\theta,a}(m) = \lambda \frac{e^{\frac{r_{\theta,a}(m)}{\epsilon}}}{\sum_b e^{\frac{r_{\theta,b}(m)}{\epsilon}}}$$

$$\frac{d}{dt} m_{\theta,a} = \lambda \frac{e^{\frac{r_{\theta,a}(m)}{\epsilon}}}{\sum_b e^{\frac{r_{s,b}(m)}{\epsilon}}} - \lambda m_{\theta,a}$$

BNN dynamics Brown-von Neumann-Nash dynamics for

$$L_{\theta,b;\theta,a}(m) = \lambda \sum_b m_{\theta,b} \max(0, r_{\theta,a}(m) - \sum_b m_{\theta,b} r_{\theta,b}(m))$$

$$\frac{d}{dt} m_{\theta,a} = \lambda \sum_b m_{\theta,b} \max(0, r_{\theta,a}(m) - \sum_b m_{\theta,b} r_{\theta,b}(m)) - \lambda m_{\theta,a} \sum_b \max(0, r_{\theta,b}(m) - \sum_b m_{s,b} r_{\theta,a}(m))$$

Note that in general the trajectories of the mean field dynamics need not to converge.

6.4.4 Dynamics for pairwise local interaction

Suppose now that at each time slot, exactly two players among N are randomly selected i.e. $\sharp \mathcal{B} = 2$ w.p. 1. Here too H2 and H3 are automatically satisfied. Then the mean dynamics is given by

$$\sum_{x_1',x_2'} \sum_{x_1,x_2} m_{x_1} m_{x_2} L_{x_1,x_2;x_1',x_2'}(\vec{m}) \left(\sum_{l=1}^{2} (\vec{e}_{x_l'} - \vec{e}_{x_l}) \right) \tag{6.6}$$

where $L_{x_1,x_2;x_1',x_2'}(\vec{m})$ is the probability that the first object in interaction movers to state x_1' and the second one to state x_2' given that the first was in state x_1, the second in state x_2, and the system state was $\vec{m}$ just before the interaction.

We can expand Equation (6.6) as

$$
\begin{aligned}
\frac{d}{dt} m_x &= \sum_{x_1,x_2,x_2'} m_{x_1} m_{x_2} L_{x_1,x_2;x,x_2'}(\vec{m}) + \sum_{x_1,x_2,x_1'} m_{x_1} m_{x_2} L_{x_1,x_2;x_1',x}(\vec{m}) \\
&\quad - m_x \sum_{x_2,x_1',x_2'} m_{x_2} L_{x,x_2;x_1',x_2'}(\vec{m}) - m_x \sum_{x_1,x_1',x_2'} m_{x_1} L_{x_1,x;x_1',x_2'}(\vec{m}) \\
&= \sum_{x_1,x_2} m_{x_1} m_{x_2} \eta^1_{x_1,x_2;x}(\vec{m}) + \sum_{x_1,x_2} m_{x_1} m_{x_2} \eta^2_{x_1,x_2;x}(\vec{m}) - m_x \sum_{x_2} m_{x_2} \eta^t_{x,x_2}(\vec{m}) - m_x \sum_{x_1} m_{x_1} \eta^t_{x_1,x}(\vec{m})
\end{aligned}
$$

with

$$\eta^1_{x_1,x_2;x} = \sum_{x_2'} L_{x_1,x_2;x,x_2'}, \quad \eta^2_{x_1,x_2;x} = \sum_{x_1'} L_{x_1,x_2;x_1',x}, \quad \eta^t_{x_1,x_2} = \sum_{x_1',x_2'} L_{x_1,x_2;x_1',x_2'} \tag{6.7}$$

The term $\eta^j_{x_1,x_2;x}$(resp. $\eta^t_{x_1,x}, \eta^t_{x,x_1}$) corresponds to the *rate of transfer* of a player with $x_1 = (\theta_1, a_1)$ to switch into $x = (\theta, a)$ (resp. from x) after an encounter with another player with $x_2 = (\theta_2, a_2)$.

6.4.5 Equilibrium and rest point

Proposition 6.4.5.1 (sufficient condition)**.** *Suppose that the drift limit $\vec{f}$ satisfies*

$$\vec{f}(m) \neq 0 \implies \langle \vec{f}(m), r(m) \rangle = \sum_{s,a} r_{s,a}(m) f_{s,a}(m) > 0 \tag{6.8}$$

where

$$f_y(m) = \sum_{k \geq 1} J_k(m) f_y^k(m),$$

$$f_y^k(m) = \sum_{x_1,\dots,x_k} \left(\prod_{l=1}^{k} m_{x_l} \right) \left(\sum_{j=1}^{k} \eta_{x;y}^j(m,k) \right) - m_y \left(\sum_{j=1}^{k} \sum_{x_{-j}} (\prod_{l=1,l \neq j}^{k} m_{x_l}) \eta_{y,x_{-j}}^t \right),$$

where

$$y = (s,a), \quad \eta_{x;y}^j(m,k) = \sum_{x'_{-j}} L_{x;(y,x'_{-j})}(m,k),$$

$$(y, x'_{-j}) = (x'_1, x'_2, \dots x'_{j-1}, y, x'_{j+1}, \dots, x'_k)$$

and

$$\eta_{(y,x_{-j})}^t(m,k) = \sum_{x'} L_{(y,x_{-j});x'}(m,k).$$

Then any stationary equilibrium state is a rest point of ODE.

Proof. Suppose that the condition

$$\vec{f}(m) \neq 0 \implies \langle \vec{f}(m), r(m) \rangle = \sum_y r_y(m) f_y(m) > 0$$

holds. Let m^* be a stationary equilibrium state (it exists by Proposition 6.4.1.1). We want to prove that $f(m^*) = 0$. Since, m^* is a stationary equilibrium state, it solves the variational inequality problem

$$\forall h \in \{z = (z_{s,a})_{s,a}, \sum_{s,a} z_{s,a} = 0\}, \langle h, r(m^*) \rangle \leq 0.$$

Define the vector $x = (x_{s,a})_{s,a}$ such that $x_{s,a} = f_{s,a}(m^*)$. By construction, the drift limit satisfies

$$\sum_{s,a} f_{s,a}(m^*) = 0.$$

Thus,

$$0 \geq \langle x, r(m^*) \rangle = \langle f(m^*), r(m^*) \rangle.$$

This condition is equivalent to $f(m^*) = 0$. $\qquad \square$

The condition given by Equation (6.8) is called positive correlation (168) condition. In the one object per time slot model with single internal state case, it has been shown that (see (168)) that the replicator dynamics, the Brown-von Neumann-Nash dynamics, $\theta-$Smith dynamics or general pairwise comparison dynamics are positively correlated. As a corollary, the set of rest points of all these dynamics contains the set of equilibrium states.

Proposition 6.4.5.2. *Suppose that the polymatrix of transition L satisfies*

$$L_{s,a,x_{-j};s,b}(m) > 0 \iff a,b \in \mathcal{A}_s, \; r_{s,a}(m) < r_{s,b}(m)$$

for each j, x_{-j} and m. Then

1. *The mean dynamics satisfies the condition of Proposition 6.4.5.1, Equation (6.8).*

2. *m^* is an equilibrium state if and only if for all $x_j = (\theta, a)$, one of the following conditions is satisfied:*

    (a) $\sum_b L_{s,a,x_{-j};s,b}(m) = 0$

(b) $m^*_{s,a} = 0$.

3. *Any rest point of the ODE is a stationary equilibrium state.*

Proof. The point (3) follows from the points $2 - (a)$ and $2 - (b)$. The point The point $2 - (a)$ of Proposition is the optimality of action a used in state s and the point $2 - (b)$ says that if a is not used if a is not a good strategy. We now proof the first part of the Proposition. Suppose that m such that $f(m) \neq 0$ with

$$f_y(m) = \sum_{k \geq 1} J_k(m) f^k_y(m),$$

$$f^k_y(m) = \sum_{x_1,\dots,x_k} \left(\prod_{l=1}^k m_{x_l} \right) \left(\sum_{j=1}^k \eta^j_{x;y}(m,k) \right) - m_y \left(\sum_{j=1}^k \sum_{x_{-j}} (\prod_{l=1,l\neq j}^k m_{x_l}) \eta^t_{y,x_{-j}} \right),$$

One has,

$$\sum_y r_y(m) f_y(m) = \sum_y \sum_{k \geq 1} J_k(m) f^k_y(m) r_y(m) = \sum_{k \geq 1} J_k(m) \left(\sum_y f^k_y(m) r_y(m) \right) \tag{6.9}$$

We can rewrite $\sum_y r_y(m) f_y(m)$ as $\sum_{k \geq 1} J_k(m) \sum_y [A^k_y - B^k_y]$ where

$$A^k_y = \sum_{j=1}^k \sum_{x_1,\dots,x_k} \sum_{x'_{-j}} (\prod_{l=1}^k m_{x_l}) L_{x_j,x_{-j};y,x'_{-j}}(m,k) r_y(m),$$

$$B^k_y = m_y \left[\sum_{j=1}^k \sum_{x_{-j}} (\prod_{l=1,l\neq j}^k m_{x_l}) \left(\sum_{(x'_j,x'_{-j})} L_{y,x_{-j};x'_j,x'_{-j}}(m,k) r_y(m) \right) \right] \tag{6.10}$$

$$= \sum_{j=1}^k \sum_{x_{-j}} \sum_{(x'_j,x'_{-j})} (m_y \prod_{l=1,l\neq j}^k m_{x_l}) L_{y,x_{-j};x'_j,x'_{-j}}(m,k) r_y(m) \tag{6.11}$$

By interchanging the roles of y and x'_j we obtain that

$$B^k_y = \sum_{j=1}^k \sum_{(x_j,x_{-j})} \sum_{x'_{-j}} (\prod_{l=1}^k m_{x_l}) L_{x_j,x_{-j};y,x'_{-j}}(m,k) r_{x_j}(m) \tag{6.12}$$

Then, $\sum_{k \geq 1} J_k(m) \left(\sum_y f^k_y(m) r_y(m) \right)$ is exactly

$$\sum_{k \geq 1} \sum_{j=1}^k \sum_{(x_j,x_{-j})} \sum_{(y,x'_{-j})} (\prod_{l=1}^k m_{x_l}) J_k(m) L_{x_j,x_{-j};y,x'_{-j}}(m,k) \max[0, r_y(m) - r_{x_j}(m)].$$

Since there exists at least one component y such $f_y(m) \neq 0$, and

$$L_{x_j,x_{-j};z,x'_{-j}}(m,k)[r_z(m) - r_{x_j}(m)] = L_{x_j,x_{-j};z,x'_{-j}}(m,k) \max[0, r_z(m) - r_{x_j}(m)] \geq 0$$

for all z, we conclude there exists $k \geq 1$ such that $L_{x_j,x_{-j};y,w'_{-j}}(m,k) \max[0, r_y(m) - r_{x_j}(m)] > 0$. This implies that $\sum_y r_y(m) f_y(m) > 0$. $\qquad \square$

Cooperative vs Competitive mean field limit We say that the mean dynamics is cooperative (resp. competitive) if all $w \neq w'$ $\frac{d}{dm_{w'}} f_w(m) \geq 0$ (resp. $\frac{d}{dm_{w'}} f_w(m) \leq 0$). In the bidimensional (three strategies) case it has been shown in (101) that the trajectories of cooperative game dynamics converge starting from almost any initial condition in the simplex. Example of cooperative mean dynamics includes dynamics of supermodular population games (168). Generic convergence of cooperative (resp. competitive) dynamics with strict inequality can be found in (190).

Dynamic equilibrium

- In many cases in evolving games, it is known that the trajectories of evolutionary game dynamics can be chaotic, may have cycle limits and may not converge. Under several evolutionary game dynamics, the expected payoff obtained in the cycle is not so far from candidate to be a "static" equilibrium payoff. This suggests a *time dependent equilibrium* approach in evolving games.

- a trajectory is a dynamic equilibrium if the *time average payoff under this trajectory leads to an equilibrium payoff*

Example: the trajectory of RSP leads to a "dynamic equilibrium" but it cycles around the Shapley triangle.

$m : [t_0, \infty[\longrightarrow \Delta_n$ is a dynamic equilibrium if for any other trajectory, there exists t_1 sufficiently large such that

$$\liminf_T \frac{1}{T - t_1} \int_{t_1}^{T} \langle m(t) - m'(t), r(m(t)) \rangle \, dt \geq 0$$

Corollary: a "static" equilibrium state m^* is a special case of dynamic equilibrium where the trajectory is $t \longmapsto m^*$ (constant function).

6.5 A Spatial Non-Reciprocal Random Access Model

Random Access Control algorithms have played an increasingly important role in the development of wired and wireless networks and the performance and stability of these algorithms, such as slotted-Aloha, Carrier Sense Multiple Access (CSMA) is still an open problem. Distributed medium access control, starting from the first version of Abramson's Aloha to the most recent algorithms used in IEEE802.11, have enabled a rapid growth of wireless networks. They aim at efficiently and fairly sharing a resource among users even though each user must decide independently (eventually after receiving some messages or listening, sensing channels) when and how to attempt to use the resource. Random access algorithms have generated a lot of research interest, especially recently in attempts to use multi-hop wireless networks to provide high-speed access to the internet with low-cost and low-energy consumption.

In this section, we restrict our attention to wireless networks, where the resources are receivers, base station or access points and where users interact because of interference, i.e., interfering users cannot transmit simultaneously. There is a collision if another user (mobile) transmits with a greater power level at the same range of the receiver. Motivated by the interest of evolving dense networks, game theory evolving was found to be an appropriate framework to apply in networks. We provide asymptotic analysis of spatial random access based on mean field interaction where interfering users share resources placed on a undirected graph using an

Aloha-type access control. We consider several classes of users. Users are spatially distributed and the arrival processes around each vertex are assumed to be independent. The interactions between users are *non-reciprocal* in the sense that *the set of users causing interference differ from the set of those suffering from these interferences.* The scenario has been introduced and analyzed in (213) in the context of evolutionary games.

Model Description

We consider a spatial non-reciprocal random access model described as follows:

- Let define a system composed of an undirected graph $G = (V, E)$ with finite vertices $V = \{1, 2, \ldots, K\}$, and E is set of edges, and many classes of users distributed in 3D. For each vertex $i \in V$, the one-hop neighborhood of i is denoted by

$$\mathcal{N}_i = \{i\} \cup \{j \in V, \ (i, j) \in E\}.$$

The set $\mathcal{N}_i$ corresponds to the set of 1-adjacent nodes to i.

- Time is slotted. During each time slot t a random number A_t of users arrive in the system. Denote by A_t^i the number of users arriving in vertex i at time t. We assume that arrival process $\{A_t^i\}_t$ are i.i.d in time with expectation $\mathbb{E}(A_t^i) = \alpha^i$. Denote by W_t^i the random variable representing the number of users at vertex i at time t. A user at vertex i require a service with some probability. The user transmits with some probability

$$\frac{u^i}{\sum_{j \in \mathcal{N}_i} W_t^j}.$$

If B_t^i is the total number of users requiring a service at time t. Then, the random variable B_t^i follows a binomial law with parameters W_t^i and $\dfrac{u^i}{\sum_{j \in \mathcal{N}_i} W_t^j}$.

$$B_t^i \sim \text{Bin}\left(W_t^i, \frac{u^i}{\sum_{j \in \mathcal{N}_i} W_t^j}\right)$$

- There is a successful transmission at vertex i if only one of users in $\mathcal{N}_i$ transmit at time t.

$$W_{t+1}^i = W_t^i - 1_{\{B_t^i = 1\}} \prod_{j \in \mathcal{N}_i \setminus \{i\}} 1_{\{B_t^j = 0\}} + A_t^i$$

Given $(W_t^1, \ldots, W_t^K)$, the random variable $B_t^1, \ldots, B_t^K$ satisfy the decoupling conditions. $(W_t^1, \ldots, W_t^K) \in \mathcal{N}^K$ is an irreducible Markov chain.

Drift

Let us look at the expected change in W in one slot when all users use the strategy $u_1, \ldots, u^K$. Define

$$f(u^1, \ldots, u^K, x) = \mathbb{E}[W_{t+1} - W_t | W_t = x] \tag{6.13}$$

The drift is given by

$$
f^i(u^1, \ldots, u^K, x) = \begin{cases} \alpha^i - \dfrac{u^i x^i}{\sum_{j \in \mathcal{N}_i} x^j}\left(1 - \dfrac{u^i}{\sum_{j \in \mathcal{N}_i} x^j}\right)^{x^i - 1} \displaystyle\prod_{j \in \mathcal{N}_i \setminus \{i\}} \left(1 - \dfrac{u^j}{\sum_{j \in \mathcal{N}_i} x^j}\right)^{x^j} & \text{if } x^i > 0 \\[2em] \alpha^i & \text{if } x^i = 0 \end{cases}
$$

For any sequence (x_t^i) such that $x_t^i \longrightarrow +\infty$, define the process

$$
M_k^{x_t} = \frac{W(\lfloor k \sum_i x_t^i \rfloor)}{\sum_i x_t^i}.
$$

$M_k^x \in \mathbb{D}([0, T], \mathbb{R}^K)$ where $\mathbb{D}([0, T], \mathbb{R}^K)$ is the space of cadlag or RCLL (right continuous with left limits) functions from the compact interval $[0, T]$ to $\mathbb{R}^K$ endowed with the Skorokhod topology.

Let $(x_t^i)_t$ be a sequence going to infinity. Then the collection $\{(M_k^x)_k, \ k \in \mathbb{N}\}$ has a compact close in the Skorokhod topology. Any limit m of subsequence of M_k^x is Lipschitz continuous.

If a mean field limit m satisfies $\forall i, \ \displaystyle\sum_{j \in \mathcal{N}_i} m^j(0) > 0$ then m is solution of the ordinary differential equation

$$
i = 1, 2, \ldots, K
$$

$$
\dot{m}^i(t) = \alpha^i - \frac{u^i m^i(t)}{\sum_{j \in \mathcal{N}_i} m^j(t)} e^{-\sum_{k \in \mathcal{N}_i \setminus \{i\}} \frac{u^k m^k(t)}{\sum_{j \in \mathcal{N}_k} m^j(t)}} e^{-\frac{u^i m^i(t)}{\sum_{j \in \mathcal{N}_i} m^j(t)}}
$$

The mean field limit dynamics is topologically equivalent to the Ray-projection evolutionary game dynamics with expected rate transition $\alpha^i - \dfrac{u^i m^i(t)}{\sum_{j \in \mathcal{N}_i} m^j(t)} e^{-\sum_{k \in \mathcal{N}_i} \frac{u^k m^k(t)}{\sum_{j \in \mathcal{N}_k} m^j(t)}}$

Chapter 7

Stochastic mean field games

We introduce a particular class of stochastic mean field games also called Markov Decision Evolutionary Games with N players, in which each individual in a large population interacts with other randomly selected players. The states and actions of each player in an interaction together determine the instantaneous payoff for all involved players. They also determine the transition probabilities to move to the next state. Each individual wishes to maximize the total expected discounted payoff over an infinite horizon. We provide a rigorous derivation of the asymptotic behavior of this system as the size of the population grows to infinity. We show that under any Markov strategy, the random process consisting of one specific player and the remaining population converges weakly to a jump process driven by the solution of a system of differential equations. We characterize the solutions to the team and to the game problems at the limit of infinite population and use these to construct almost optimal strategies for the case of a finite, but large, number of players. We show that the large population asymptotic of the microscopic model is equivalent to a (macroscopic) Markov decision evolutionary game in which a local interaction is described by a single player against a population profile. We illustrate our model to derive the equations for a dynamic evolutionary Hawk and Dove game with energy level.

7.1 Introduction

We consider a large population of players in which frequent interactions occur between small numbers of chosen individuals. Each interaction in which a player is involved can be described as one stage of a dynamic game. The state and actions of the players at each stage determine an immediate payoff (also called *fitness* in behavioral ecology) for each player as well as the transition probabilities of a controlled Markov chain associated with each player. Each player wishes to maximize its expected fitness averaged over time.

This model extends the basic evolutionary games by introducing a controlled state that characterizes each player. The stochastic dynamic games at each interaction replace the matrix games, and the objective of maximizing the expected long-term payoff over an infinite time horizon replaces the objective of maximizing the outcome of a matrix game. Instead of a choice of a (possibly mixed) action, a player is now faced with the choice of decision rules (called strategies) that determine what actions should be chosen at a given interaction for given present and past observations.

This model with a finite number of players, called a mean field interaction model, is in general difficult to analyze because of the huge state space required to describe the sate of all players. Then taking the asymptotics as the number of players grows to infinity, the whole behavior of the population is replaced by a deterministic limit that represents the system's state, which is fraction of the population at each individual state that use a given action.

In this chapter we study the asymptotic *dynamic* behavior of the system in which the population profile evolves in time. For large N, under mild assumptions (see Section 7.3), the mean field converges to a deterministic measure that satisfies a non-linear ordinary differential equation for under any stationary strategy. We show that the mean field interaction is asymptotically equivalent to a Markov decision evolutionary game. When the rest of the population uses a fixed strategy u, any given player sees an equivalent game against a collective of players whose state evolves according to the ordinary differential equation (ODE) which we explicitly compute. In addition to providing the exact limiting asymptotic, the ODE approach provides tight approximations for fixed large N. The mean field asymptotic calculations for large N for given choices of strategies allows us to compute the equilibrium of the game in the asymptotic regime.

Related Work. Mean field interaction models have already been used in standard evolutionary games in a completely different context: that of evolutionary game dynamics (such as replicator dynamics) see e.g. (198) and references therein. The paradigm there has been to associate relative growth rate to actions according to the fitness they achieved, then study the asymptotic trajectories of the state of the system, i.e. the fraction of users that adopt the different actions. *Non-atomic* Markov Decision Evolutionary Games have been applied to firm idiosyncratic random shocks and to cellular communications. Nonatomic mean field games have been studied in (129; 48; 173; 114; 108). The authors in (129) develop a mean field approach for optimal control and differential games - continuous state and time- mean field games with continuum of players and their partial differential equation characterization.

Structure. The remainder of this chapter is organized as follows. In next section we present the model assumptions and notations. In Section 7.3 we present some convergence results of the ODE in the random number of interacting players. In Section 7.5 a resource competition between animals with two types of behaviors and several states is presented. All the sketch of proofs are given in Appendix.

7.2 Model description

7.2.1 Markov Decision Evolutionary Process With N Players

We consider the following model, which we call Markov Decision Evolutionary Game with N players.

- There are $N \in \mathbb{N}$ players.

- Each player has its own state. A state has two components: the *type* of the player and the *internal state*. The type is a constant during the game. The state of player j at time t is denoted by $X_j^N(t) = (\theta_j, S_j^N(t))$ where θ_j is the type. The set of possible states $\mathcal{X} = \{1, \ldots, \Theta\} \times \mathcal{S}$ is finite.

- Time is discrete, taking values in $\dfrac{\mathbb{N}}{N} := \{0, \dfrac{1}{N}, \dfrac{2}{N}, \ldots\}$.

- The *global detailed description* of the system at time t is $X^N(t) = (X_1^N(t), \ldots, X_N^N(t))$.

Define $M^N(t)$ to be the current population profile i.e $M_x^N(t) = \frac{1}{N}\sum_{j=1}^{N} 1_{\{X_j^N(t)=x\}}$. At each

time t, $M^N(t)$ is in the finite set $\{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}^{\sharp\mathcal{X}}$. and $M_{\theta,s}^N(t)$ is the fraction of players who belong to population of type θ (also called subpopulation θ) and have internal state s. Also let $\bar{M}_\theta^N = N\sum_{s\in\mathcal{S}} M_{\theta,s}^N(t)$ be the size of subpopulation θ (independent of t by hypothesis). We do not make any specific hypothesis on the ratios $\frac{\bar{M}_\theta^N}{N}$ as N gets large (it may be constant or not, it may tend to 0 or not).

● *Strategies and local interaction:* At time slot t, an ordered list $\mathcal{B}^N(t)$, of players in $\{1,2,\ldots,N\}$, without repetition, is selected randomly as follows. First we draw a random number of players $K(t)$ such that

$$\mathbb{P}(K(t) = k|M^N(t) = \vec{m}) = J_k^N(\vec{m})$$

where the distribution $J_k^N(\vec{m})$ is given for any N, $\vec{m} \in \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}^{\sharp\mathcal{X}}$. Second, we set $\mathcal{B}^N$ to an ordered list of $K(t)$ players drawn uniformly at random among the $N(N-1)...(N-K(t)+1)$ possible ones. By abuse of notation we write $j \in \mathcal{B}^N(t)$ with the meaning that j appears in the list $\mathcal{B}^N(t)$.

Each player such that $j \in \mathcal{B}^N(t)$ takes part in a one-shot event at time t, as follows. First, the player chooses an action a in the finite set $\mathcal{A}$ with probability $u_\theta(a|s)$ where (θ, s) is the current player state. The stochastic array u is the strategy profile of the population, and u_θ is the strategy of subpopulation θ. A vector of probability distributions u which depend only on the type of the player and its internal state is called *stationary strategy*.

Second, say that $\mathcal{B}^N(t) = (j_1, \ldots, j_k)$. Given the actions $a_{j_1}, ..., a_{j_k}$ drawn by the k players, we draw a new set of internal states $(s'_{j_1}, ..., s'_{j_k})$ with probability $L_{\underline{\theta};\underline{s};\underline{a};\underline{s}'}^N(k, \vec{m})$,

$$\text{where } \underline{\theta} = (\theta_{j_1}, ..., \theta_{j_k}), \ \underline{s} = (s_{j_1}, ..., s_{j_k}), \ \underline{a} = (a_{j_1}, ..., a_{j_k}), \ \underline{s}' = (s'_{j_1}, ..., s'_{j_k})$$

Then the collection of k players makes one synchronized transition, such that

$$S_{j_i}^N(t + \frac{1}{N}) = s'_{j_i} \quad i = 1, \ldots, k$$

Note that $S_j^N(t + \frac{1}{N}) = S_j^N(t)$ if j is not in $\mathcal{B}^N(t)$.

It can easily be shown that this form of interaction has following properties: (1) X^N is Markov and (2) players can be observed only through their state.

The model is entirely specified by the probability distributions J^N, the Markov transition kernels L^N and the strategy profile u. In this paper, we assume that J^N and L^N are fixed for all N, but u can be changed and does not depend on N (though it would be trivial to extend our results to strategies that depend on N, but this appears to be unnecessary complication). We are interested in large N.

It follows from our assumptions that

1. $M^N(t)$ is Markov.

2. for any fixed $j \in \{1, \ldots, N\}$, $(X_j^N(t), M^N(t))$ is Markov. This means that the evolution of one specific player $X_j^N(t)$ depends on the other players only through the occupancy measure $M^N(t)$.

7.2.2 Payoffs

We consider two types of instantaneous payoff and one discounted payoff:

• *Instant Gain:* This is the random gain $G_j^N(t)$ obtained by one player whenever it is involved in an event at time t. We assume that it depends on this player's state just before the event and just after the event, the chosen action, and on the states and actions of all players involved in this event. Formally, if player $j \in \mathcal{B}^N(t)$

$$G_j^N(t) = g^N(x_j, a_j, x'_j, x_{\mathcal{B}^N(t)\setminus j}, a_{\mathcal{B}^N(t)\setminus j}, x'_{\mathcal{B}^N(t)\setminus j})$$

where $x_j = X_j^N(t)$, a_j is the action chosen by player j, $x'_j = X_j^N(t + \frac{1}{N})$, $x_{\mathcal{B}^N(t)\setminus j}$ [resp. $x'_{\mathcal{B}^N(t)\setminus j}$] is the list of states at time t [resp. at time $t + \frac{1}{N}$] of players other than j involved in the event, $a_{\mathcal{B}^N(t)\setminus j}$ is the list of their actions and $g()$ is some non random function defined on the set of appropriate lists. Whenever j is not in $\mathcal{B}^N(t)$, $G_j^N(t) = 0$. We assume that $G_j^N(t)$ is bounded, i.e. there is a non random number C_0 such that, with probability 1: for all j, t: $|G_j^N(t)| \leq C_0$

• *Expected Instant Payoff:* It is defined as the expected instant gain of player j, given the state x of j and the population profile $\vec{m}$. By our indistinguishability assumption, it does not depend on the identity of a player, so we can write it as

$$r^N(u, x, \vec{m}) := \mathbb{E}\left(G_j^N(t) \,\Big|\, X_j^N(t) = x, M^N(t) = \vec{m} \right)$$

Note that this conditional expectation contains the case when j is not in $\mathcal{B}^N(t)$, i.e. when $G_j^N(t) = 0$.

• *Discounted Long-Term Payoff:* It is defined as the expected discounted long term payoff of one player, given the initial state of this player and the population:

$$\bar{r}^N(u; x, \vec{m}) := \mathbb{E}\left(\sum_{\substack{t=0 \text{ step } 1/N}}^{\infty} e^{-\beta t} G_j^N(t) | X_j(0) = x, M^N(0) = \vec{m} \right)$$

where β is a positive parameter (existence follows from the boundedness of G_j^N). The fact that it does not depend on the identity j of the player, but only on its initial state x and the initial population profile $\vec{m}$, follows from the indistinguishability assumption.

We defined the Discounted Long-Term Payoff in terms of the instant gain, as this is the most natural definition. The following proposition shows that the alternative definition, by means of the expected instant payoff, is equivalent.

Proposition 7.2.2.1. *For all player state x and population profile $\vec{m}$*

$$\bar{r}^N(u; x, \vec{m}) \quad = \quad \mathbb{E}\left(\sum_{\substack{t=0 \text{ step } 1/N}}^{\infty} e^{-\beta t} r^N(u, X_j^N(t), \vec{M}^N(t)) | X_j(0) = x, M^N(0) = \vec{m} \right)$$

Sketch of proof of Proposition 7.2.2.1

Let τ^N be the first time after $t = 0$ that $X_j^N(t)$ hits in some given state. We show that

$$\bar{r}^N = \frac{1}{N} \mathbb{E} \sum_{\substack{s=0 \\ \text{step } 1/N}}^{\tau^N} e^{-\beta t} r^N \left(X_j^N(s), M^N(s) \right) \tag{7.1}$$

Define for $t \in \mathbb{N}/N$:

$$Z_t^N = \sum_{\substack{s=0 \\ \text{step } 1/N}}^{t} e^{-\beta s} \left(G^N(s) - r^N \left(X_j^N(s), M^N(s) \right) \right)$$

we have, for $0 \le s \le t$:

$$
\begin{aligned}
Q \quad &:= \quad \mathbb{E} \left(Z_t^N - Z_s^N | \mathcal{F}_s^N \right) \\
&= \sum_{\substack{u'=0 \\ \text{step } 1/N}}^{t} e^{-\beta u'} \mathbb{E} \left(G^N(u') - r^N \left(X_j^N(u'), M^N(u') \right) | \mathcal{F}_s^N \right)
\end{aligned}
$$

which can be written as

$$\sum_{\substack{u'=0 \\ \text{step } 1/N}}^{t} e^{-\beta u'} \mathbb{E} \left(\mathbb{E} \left(G^N(u') - r^N \left(X_j^N(u'), M^N(u') \right) | \mathcal{F}_{u'}^N \right) | \mathcal{F}_s^N \right)$$

$$= 0$$

thus Z_t^N is an $\mathcal{F}_t^N-$ martingale. Now τ^N is a stopping time with respect to the filtration $\mathcal{F}_t^N$ thus, by Doob's stopping time theorem: $\mathbb{E} Z_{t \wedge \tau^N}^N = \mathbb{E} Z_{0 \wedge \tau^N}^N = 0$ Further, $Z_{t \wedge \tau^N}^N \le K |\tau^N|$ for some constant K. Since τ^N is almost surely finite and has a finite expectation, we can apply dominated convergence (with $t \to \infty$) and obtain $\mathbb{E} Z_{\tau^N}^N = 0$.

7.2.3 Focus on One Single Player

We are interested in the following special case (here we make the dependency on the strategy explicit). There are two types of players, i.e. $\Theta = 2$. There is exactly one player (the player of interest) with type 1. All other players have type 2. In this case we use the notation $R^N(u_1, u_2; s, \vec{m})$ for the discounted long-term payoff obtained by the player in type 0, when her strategy is u_1 and all other players's strategy is u_2, given that this player's initial internal state is s and the initial type 2 subpopulation profile is $\vec{m}$. Note that

$$R^N(u_1, u_2; s, \vec{m}) = \bar{r}^N(u_1, u_2; (1, s), \vec{m}')$$

with $m'_{1,s'} = \frac{1}{N} 1_{s=s'}$ and $m'_{2,s'} = m_{2,s'}$ for all $s' \in \mathcal{S}$.

Markov Decision Evolutionary Game

Player j may choose a strategy u^j. We look for a (Nash) equilibrium u such that if all players use u then no player has an incentive to deviate from u. For any finite N one can map this into

a standard Markov game. This is true for both the case where the number of players is known and in the case it is unknown when taking a decision. Therefore we know that a stationary equilibrium exists in the discounted case. A stationary equilibrium is solution of the fixed point equation:

$$\forall j, \ u_{j,\theta} \in \arg\max_{v_{j,\theta}} R^N(v_{j,\theta}, u_{-j}; s, m)$$

By assuming symmetry per type we can show that a stationary equilibrium exists which is a solution of the fixed point equation

$$\forall \theta, u_\theta \in \arg\max_{v_\theta} R^N(v_\theta, u; s, m)$$

Markov Decision Evolutionary Team

We wish to find a stationary u that maximizes R^N averaged over all players.

$$u = (u_1, \ldots, u_\Theta) \in \arg\max_v R^N(v; s, m)$$

7.3 Main Results

7.3.1 Scaling Assumptions

We are interested in the large N regime and obtain that, for any fixed j, (X_j^N, M^N) converges weakly to a simple process. This requires the weak convergence of $M^N(0)$ to some $\vec{m}_0$.

We assume that the parameters of the model and the payoff per time unit converge as $N \to \infty$, i.e.

$$\begin{cases} J_k^N(\vec{m}) \to J_k(\vec{m}) \\ L_{\theta;s;a;s'}^N(k, \vec{m}) \to \quad L_{\theta;s;a;s'}(k, \vec{m}) \\ r^N(u, x, \vec{m}) \to r(u, x, \vec{m}) \end{cases} \tag{7.2}$$

Our main scaling assumption is

H1 $\sum_k k^2 J_k(\vec{m}) < \infty$ for all $\vec{m} \in \Delta$. This ensures that the second moment of the number of players involved in an event per time slot is bounded.

Note that H1 excludes the case where the number of players involved in an event per time slot scales like N (i.e. synchronous transitions of all players at the same time). There may be large N asymptotic results for such cases (130) but the limit is not given by an ODE. In contrast, H1 is automatically true if the number of players involved in an event per time slot is upper bounded by a non random constant. We also need some technical assumptions, which are usually true and can be verified by inspection.

H2 $\sum_k J_k(\vec{m}) > 0$ for all $\vec{m} \in \Delta$ (Δ is the simplex $\{\vec{m} : m_{\theta,s} \geq 0, \sum_{\theta,s} m_{\theta,s} = 1\}$). This ensures that the mean number of players involved in an event per time slot, $\sum_{k \geq 0} k J_k(\vec{m})$ is non zero.

Define the *drift* of $M^N(t)$ as

$$\vec{f}^N(u, \vec{m}) = \mathbb{E}\left(M^N(t + \frac{1}{N}) - M^N(t) | M^N(t) = \vec{m}\right)$$

Note that we make explicit the dependency on the strategy u but not on J and L, assumed to be fixed.

It follows from our hypotheses that

$$\lim_{N\to\infty} N f^N(u, \vec{m}) := f(u, \vec{m}) \tag{7.3}$$

exists.

H3 We assume that the convergence in Equation (7.3) is uniform in $\vec{m}$ and the limit is Lipschitz-continuous in $\vec{m}$. This is in particular true if one can write, for every strategy u, $f^N(u, \vec{m}) = \frac{1}{N} \phi_u(\frac{1}{N}, \vec{m})$, with ϕ_u defined on $[0, \epsilon] \times \Delta$ where $\epsilon > 0$ and Φ_u is continuously differentiable.

H4 $\mathbb{P}(X_j^N(t+1/N) = y | X_j^N(t) = x, M^N(t) = m, M^N(t+1/N) = m')$ converges uniformly in $\vec{m}, \vec{m}'$ and the limit is Lipschitz-continuous in $\vec{m}, \vec{m}'$. This is in particular true if one can write, for every strategy u, as $\zeta_{u,x;y}(1/N, m, m')$. with ζ defined on $[0, 1] \times \Delta \times \Delta$ and $\zeta_{u,x;y}$ is continuously differentiable.

Our model satisfies the assumptions in (43), therefore we have the following result:

Theorem 7.3.2 ((43)). *Assume that* $\lim_{N\to\infty} M^N(0) = \vec{m}_0$ *in probability. For any stationary strategy u, and any time t, the random process $M^N(t) = \frac{1}{N} \sum_{j=1}^{N} \delta_{X_j^N(t)}$ converges in distribution to the (non-random) solution of the ODE*

$$\dot{\vec{m}}(t) = f(u, \vec{m}(t)) \tag{7.4}$$

with initial condition $\vec{m}_0$.

7.3.3 Convergence results

We focus on one player, without loss of generality we can call her player 1, and consider the process (X_1^N, M^N). For any finite N, X_1^N and M^N are not independent, however in the limit we have the following:

Theorem 7.3.4. *Assume that* $\lim_{N\to\infty} M^N(0) = \vec{m}_0$ *and* $\lim_{N\to\infty} X_1^N(0) = x_0 = (\theta_1, s_0)$ *in probability. The discrete time process $(X_1^N(t), M^N(t))$ defined for $t \in \frac{\mathbb{N}}{N}$, converges weakly to the continuous time jump and drift process $(X_1(t), \vec{m}(t))$, where $\vec{m}(t)$ is solution of the ODE Equation (7.4) with initial condition $\vec{m}_0$ and $X_1(t)$ is a continuous time, non homogeneous jump process, with initial state x_0. The rate of transition of $X_1(t)$ from state $x_1 = (\theta_1, s_1)$ to state $x_1' = (\theta_1, s_1')$ is*

$$A(x_1, x_1'; \vec{m}(t), u) = \sum_{k \geq 1} J_k(\vec{m}) A_k(s_1, s_1'; \vec{m}(t), u)$$

with $A_k(s_1, s_1'; \vec{m}(t), u) =$

$$\sum_{\underline{\theta}; \underline{s}; \underline{a}; \underline{s}'} L_{\theta_1, \underline{\theta}; s, \underline{s}; \underline{a}; s', \underline{s}'}(k, \vec{m}(t)) \prod_{j=1}^{k} u_{\theta_j}(a_j | s_j) \prod_{j=2}^{k} m_{\theta_j, s_j}(t)$$

$$\text{where } \underline{\theta} = (\theta_2, ..., \theta_k), \underline{s} = (s_2, ..., s_k)$$
$$\underline{a} = (a_1, ..., a_k), \underline{s}' = (s_2', ..., s_k')$$

Note that, contrary to results based on propagation of chaos, we do not assume that the distribution of player states at time 0 is exchangeable. In contrast, we will use Theorem 7.3.4 precisely in the case where player 1 is different from other players. Theorem 7.3.4 motivates the following definition.

Sketch of Proof of Theorem 7.3.4

To prove the weak convergence of Z^N, we check the following steps: Without loss of generality, we took the set of states as $\mathcal{S} = \{0, 1, 2, \ldots, \sharp\mathcal{S}\}$ X_j^N has a jump r with probability

$$q_{i,i+r}^N(M^N(k)) = \frac{1}{N}L_{i,i+r}^N(M^N(k), u))$$

and M^N is the continuous process with drift f^N.

- We introduce of $\tilde{X}_j^N$ by scaling with step size $\frac{1}{N}$. Then, $Z^N = (X^N, M^N)$ is approximate in some sense by a discrete time process $\tilde{Z}^N = (\tilde{X}^N, \tilde{m}^N)$ where $\tilde{m}^N(k) = m(\lfloor Nt \rfloor)$ m solution of the ODE with $\tilde{X}_j^N$ is the discrete time jump process with transition matrix

$$q_{i,i+r}^N(\tilde{m}^N(k)) = \frac{1}{N}L_{i,i+r}^N(m(\frac{k}{N}), u)).$$

We show that $d(X_j^N, \tilde{X}_j^N) \longrightarrow 0$ for any compact of time intervals.

- $$\tilde{Z}^N = (\tilde{X}^N, \tilde{m}^N) \Longrightarrow (\tilde{X}, \tilde{m})$$

$M^N([Nt]) \longrightarrow m(t)$. We derive the weak convergence of Z^N to (X, m) where m is deterministic and X is random.

Approximation by a discrete time process

The following lemma follows from the lemma 1 and 3 in Benaim and Weibull (2003,2008), in which we incorporate behaviorial strategies.

Lemma 7.3.4.1. *For every $t > 0$ there exists a constant c such that for every $\epsilon > 0$ and N large enough one has*

$$P(\sup_{0 \leq \tau \leq T} ||M^N(\tau) - m(\tau)|| > \epsilon | M^N(0) = m_0, u) \leq 2(\sharp S)e^{-\epsilon^2 CN}$$

for all $m_0 \in \Delta_d$, all every stationary strategy u.

Since C is independent of N, and $(e^{-\epsilon^2 C})^N$ is summable, we can use the dominated convergence theorem: for all $\epsilon > 0$,

$$\sum_N \mathbb{P}\left(sup_{0 \leq \tau \leq T} \| M^N(\tau) - m(\tau) \|_\infty > \epsilon | M^N(0) = m_0, u\right) < \infty,$$

By Borel-Cantelli's lemma, for every fixed $t < \infty$, the random variable $\nu^{N,t} := \sup_{0 \leq \tau \leq t} \| M^N(\tau) - m(\tau) \|_\infty$ converges almost completely towards 0. This $\nu^{N,t}$ implies that converges almost surely to 0.

We introduce of $\tilde{X}_j^N$ by scaling with step size $\frac{1}{N}$. Then, $Z^N = (X^N, M^N)$ is approximate in some sense by a discrete time process $\tilde{Z}^N = (\tilde{X}^N, \tilde{m}^N)$ where $\tilde{m}^N(k) = m(\lfloor Nt \rfloor)$ m solution of the ODE where $\tilde{X}_j^N$ is the discrete time jump process with transition matrix $q_{i,i+r}^N(\tilde{m}^N(k)) =$

$\frac{1}{N}L_{i,i+r}(m(\frac{k}{N}),u))$. Using the lemma 7.3.4.1 and uniform Lipschitz continuity of of L^N, we obtain that

$$\sup_{i,j} \sup_{0 \le \tau \le t} \| q_{i,j}^N(M^N(\tau)) - q_{i,j}(m(\tau)) \| \le K(\epsilon_N + \sup_{0 \le \tau \le t} \|M^N(\tau) - m(\tau)\|).$$

Hence, we can write $\|M^N(\tau) - m(\tau)\| \le K(\epsilon_N + \frac{1}{N^2})$ over set of event $\Omega_\epsilon = \{\|M^N(\tau) - m(\tau)\| \le \epsilon\}$ and $P(\Omega_\epsilon) \ge 1 - 2(\sharp S)e^{-\epsilon^2 CN} \to 1$. Thus,

$$P(X^N_{j,|[0,t]} = \tilde{X}^N_{j,|[0,t]}|k \text{ transitions}) \ge \mathbb{E}(\epsilon^{Bin(\frac{1}{N},Nt)})$$

$$\mathbb{E}(\epsilon^{Bin(\frac{1}{N},Nt)}) = (1 - \frac{1}{N} + \frac{1}{N}\epsilon)^{Nt}$$

$$P(X^N_{j,|[0,t]} = \tilde{X}^N_{j,|[0,t]}|k \text{ transitions}) \ge e^\epsilon$$

and this holds for any ϵ arbitrary small. We define the metric $d(X,Y) = \sum_{k \ge 0} \frac{1}{2^k}d(X_k,Y_k)$ where $d(X_k,Y_k) = 1_{X_k \ne Y_k}$. Then, $d(X^N_{j,|[0,t]}, \tilde{X}^N_{j,|[0,t]}) \longrightarrow 0$ when N goes to infinity.

Convergence of the discrete time process To prove the weak convergence of $(\tilde{X}^N_j, \tilde{M}^N)$, we check the following steps:

- the discrete time empirical measures $\tilde{M}^N$ are tight (follows from Sznitman for finite states) and converges to a martingale problem. The limit $\tilde{m}$ is deterministic measure and is solution of ODE which has the unique solution m (given m_0, u). Thus, $\tilde{m} = m$.

- Conditionally to $\tilde{M}^N$, $\tilde{X}^N_j$ converges to a martingale problem. The jump and drift process $\tilde{X}$ with time dependent transition is given by the limit of the marginal of $A^N(.|\tilde{M}^N, m_0, x_0, u)$. We derive the weak convergence of $(\tilde{X}^N_j, \tilde{M}^N)$ to $(\tilde{X}, \tilde{m})$ where $\tilde{m}$ is deterministic and $\tilde{X}$ is random. For this we use the Theorem 17.25 and its discrete time approximation in Theorem 17.28 pages 344-347 in Kallenberg.

Definition 7.3.4.2. *To a game as defined in Section (7.2.1) we associate a "Macroscopic Markov Decision Evolutionary Game", defined as follows. There is one player, (player 1), with state $X_1(t)$ and a population profile $\vec{m}(t)$. The initial condition of the game is $X_1(0) = x$, $\vec{m}(0) = \vec{m}_0$. The population profile is solution to the ODE (7.4) and $X_1(t)$ evolves as a jump process as in Theorem 7.3.4.*

Further, let $\bar{r}(u; x, \vec{m})$ be the discounted long-term payoff of player 1 in this game, given that $X_1(0) = x$ and $\vec{m}(0) = \vec{m}_0$, i.e.

$$\bar{r}(u; x, \vec{m}) = \mathbb{E}\left(\int_0^\infty e^{-\beta t}r(u, X_1(t), m(t))|X_1(0) = x, \vec{m}(0) = \vec{m}_0\right)$$

We also consider, as in Section (7.2.3), the case with $\Theta = 2$ types and define by analogy $R(u_1, u_2; s, \vec{m})$ as the discounted long-term payoff when player 1 starts in state s and the population profile starts in state $\vec{m}$, with player 1 using strategy u_1 and other players strategy u_2.

In order to exploit the convergence in distribution of the process focused on one player, we need that the payoff be continuous in the topology of this convergence. This is stated in the next theorem.

Theorem 7.3.5. *Let $E = \mathcal{S} \times \Delta$ and $D_E[0, \infty)$ the set of cadlag functions from $[0, \infty)$ to $\mathbb{R}$, equipped with Skorokhod's topology. The mapping*

$$\begin{aligned} D_E[0, \infty) &\to \mathbb{R} \\ (s, m) &\mapsto \int_0^\infty e^{-\beta t}r(u, s(t), m(t))\, dt \end{aligned}$$

is continuous.

Using Theorem 7.3.4 and Theorem 7.3.5 we obtain the following, which is the main result of this paper. It says that when N goes to infinity, the Markov Decision Evolutionary Game with $N(t)$ of players becomes equivalent to the associated Macroscopic Markov decision evolutionary game. This reduces any multi-player problem into an effective one-player problem.

Sketch of Proof of Theorem 7.3.5

Since Skorohod's topology is induced by a metric, it is sufficient to show that whenever $(X_j^N, m^N) \to (x, m)$ in Skorohod's topology, we have:

$$\lim_{N \to \infty} \int_0^\infty e^{-\beta t} r^N(v, X_j^N(t), m^N(t)) dt = \int_0^\infty e^{-\beta t} r(v, x(t), m(t)) dt \tag{7.5}$$

By (73), page 117, there is some sequence of increasing bijections $\lambda_n : [0, \infty) \to [0, \infty)$ s.t.

$$\frac{\lambda_n(t) - \lambda_n(s)}{t - s} \to 1 \text{ uniformly in } t \text{ and } s$$

$$\text{and } \| y_n(t) - y(\lambda_n(t)) \| \to 0 \text{ uniformly in } t$$

over compact subsets of $[0, \infty)$. Fix $\epsilon > 0$, arbitrary and consider

$$
\begin{aligned}
h^N \quad &:= \quad |\int_0^\infty e^{-\beta t} r^N(X^N(t), v, m^N(t)) dt - \int_0^\infty e^{-\beta t} r(x(t), v, m(t)) dt| \\
&\leq \quad \int_0^\infty e^{-\beta t} |r^N(x^N(t), v, m^N(t)) - r(x(t), v, m(t))| dt
\end{aligned}
$$

First let $K = \sup_{x \in \mathcal{S}, v, m \in \Delta} |r(x, v, m)| < \infty$ by hypothesis, and pick some time T large enough such that $e^{-\beta T} K / \beta \leq \epsilon/3$. Thus

$$h^N \leq \epsilon/3 + \int_0^T e^{-\beta t} |r(x^N(t), v, m^N(t)) - r(x(t), v, m(t))| dt \tag{7.6}$$

Second, we use the distance on E defined by

$$d((x, m), (x', m')) = \| m - m' \| + 1_{x \neq x'} \tag{7.7}$$

$$\text{Let } K' = \sup_{x \in \mathcal{S}, v, m \in \Delta_d} \frac{|r(x, v, m) - r(x', v, m')|}{\| m - m' \|} < \infty$$

by hypothesis. It is easy to see that for all $x, x' \in \mathcal{S}$ and $m, m' \in \Delta_d$:

$$\| r(x, v, , m) - r(x', v, m') \| \leq K' d((x, m), (x', m')) \tag{7.8}$$

Thus, by Equation (7.6):

$$h^N \leq \epsilon/3 + K' \int_0^T e^{-\beta t} d\left((x^N(t), m^N(t)), (x(t), m(t)) \right) dt \tag{7.9}$$

By (73), page 117, there is some sequence of increasing bijections $\lambda^N \colon [0, \infty) \to [0, \infty)$ s.t.

$$\frac{\lambda^N(t) - \lambda^N(s)}{t - s} \to 1 \text{ uniformly in } t \text{ and } s$$

$$\text{and } d\left((x^N(t), m^N(t), \ (x^N(\lambda^N(t)), m^N(\lambda^N(t)))) \right) \to 0$$

uniformly in t over compact subsets of $[0, \infty)$. Thus there is some $N_0 \in \mathbb{N}$ such that for $N \geq N_0$ and $t \in [0, T]$:

$$d\left((x^N(t), m^N(t), \ (x^N(\lambda^N(t)), m^N(\lambda^N(t)))) \right) \leq \frac{\epsilon \beta e^{\beta T}}{3K'} \tag{7.10}$$

Thus, by the triangular inequality for d: $h^N \leq$

$$\leq \frac{\epsilon}{3} + K' \int_0^T e^{-\beta t} d\left((x^N(t) m^N(t)), (x(\lambda^N(t)), m(\lambda^N(t))) \right) dt$$

$$+ K' \int_0^T e^{-\beta t} d\left((x(\lambda^N(t)), m(\lambda^N(t))), (x(t), m(t)) \right) dt$$

$$\leq \frac{2\epsilon}{3} + K' \int_0^T e^{-\beta t} d\left((x(\lambda^N(t)), m(\lambda^N(t))), (x(t), m(t)) \right) dt \tag{7.11}$$

Third, let D be the set of discontinuity points of (x, m). Since (x, m) is cadlag, D is enumerable, thus it is negligible for the Lebesgue measure and

$$\int_0^T e^{-\beta t} d\left((x(\lambda^N(t)), a, m(\lambda^N(t))), (x(t), a, m(t)) \right) dt$$

$$= \int_0^T e^{-\beta t} d\left((x(\lambda^N(t)), m(\lambda^N(t))), (x(t), m(t)) \right) 1_{t \notin D} dt$$

Now $\lim\limits_{N \longrightarrow \infty} \lambda^N(t) = t$ and thus for $t \notin D$

$$\lim\limits_{N \longrightarrow \infty} d\left((x(\lambda^N(t)), m(\lambda^N(t))), (x(t), m(t)) \right) = 0$$

and thus by dominated convergence

$$\lim\limits_{N \longrightarrow \infty} \int_0^T e^{-\beta t} d\left((x(\lambda^N(t)), m(\lambda^N(t))), (x(t), m(t)) \right) dt = 0 \tag{7.12}$$

and for N large enough the second term in the right-hand side of Equation (7.11) can be made smaller than $\epsilon/3$. Finally, for N large enough, $h^N \leq \epsilon$. This completes the proof.

Theorem 7.3.6 (Asymptotically equivalent game). *When N goes to infinity we have (a) the discrete time process X_1^N converges in distribution to the continuous time process X_1 (b) $\bar{r}^N(u; x, \vec{m}) \to \bar{r}(u; x, \vec{m})$ and (c) $R^N(u_1, u_2; s, \vec{m}) \to R(u_1, u_2; s, \vec{m})$*

Sketch of Proof of Theorem 7.3.6

Define the discounted stochastic evolutionary game with random number of interacting players in each local interaction in which each player in x with the mixed action $u(.|x)$ receives

$r(u, x, m(t))$ where $m(t)$ is the population profile at t, which evolves under the dynamical system (7.4) and the between states follows the transition kernel L. Then, a strategy of a player is the same as in the microscopic case and the discounted payoffs

$$R(u_1, u_2, s_0, m_0) = \int_0^\infty e^{-\beta t} r(s(t), u_1, m[u_2](t)) dt$$

is the limit of $R^N(u_1, u_2, s_0, m_0)$ when N goes to infinity, where $m[u_2]$ is the solution of the ODE $\dot{m} = f(u_2, m), m(0) = m_0$. It follows that the asymptotic regime of the microscopic game and the Markov decision evolutionary game (macroscopic game) are equivalent.

7.3.7 Case with Global Attractor

Assume that, for some strategy u, the ODE (7.4) has a global attractor $\vec{m}^*$ (this may or may not hold, depending on the ODE). If in addition the model with N players is irreducible, with stationary probability distribution $\wp^N$ for M^N, then $\lim_{N \longrightarrow \infty} \wp^N = \delta_{\vec{m}^*}$ where $\delta_{\vec{m}^*}$ is the Dirac mass at $\vec{m}^*$ (follows from (43)). i.e. the large time distribution of $M^N(t)$ converges, as $N \to \infty$, to the attractor $\vec{m}^*$.

Also, $(X_j^N(t), M^N(t))$ converges to a continuous time, homogeneous Markov jump process with time-independent transition matrix:

$$A(x_1, x_1'; u) = \sum_{k \geq 1} J_k(\vec{m}) A_k(s_1, s_1'; \vec{m}^*, u)$$

Assume that the transition matrix $A(x_1, x_1'; u)$ is also irreducible and let $\pi()$ be its unique stationary probability. Also let π^N be the first marginal of the stationary probability of (X_1^N, M^N). It is natural in this case to replace the definition of the long term payoffs $R^N(u_1, u_2; s, \vec{m})$ and $R^N(u_1, u_2; s, \vec{m})$ by their stationary counterparts

$$
\begin{aligned}
R_{st}^N(u_1, u_2) &:= \sum_s \pi^N(s) R^N(u_1, u_2; s, \vec{m}^*) \\
R_{st}(u_1, u_2) &:= \sum_s \pi(s) R(u_1, u_2; s, \vec{m}^*)
\end{aligned}
$$

7.3.8 Single player per type selected per time slot

Consider the special case where at each time slot, only one player per type between the N is randomly selected and has a chance to change its action, i.e. $\sharp \mathcal{B}^N = 1$ w.p 1.

Thus H1 and H2 are automatically satisfied. The resulting ODE (see (41)) becomes

$$\frac{d}{dt} m_x(t) = \sum_{x'} m_{x'} L_{x', x}(\vec{m}, u, \Theta) - m_x \sum_{x'} L_{x, x'}(\vec{m}, u, \Theta)$$

The term $\sum_{x'} m_{x'} L_{x', x}(\vec{m}, u, \Theta)$ is the *incoming flow* in to x and the *outgoing flow* from x is $m_x \sum_{x'} L_{x, x'}(\vec{m}, u, \Theta)$.

We then obtain a large class of *state-dependent* evolutionary game dynamics. Note that in general the trajectories of the mean dynamics need not to converge. In the case of single player

selected in each time slot of $1/N$ and linear transition in m, the time averages under the replicator dynamics converge its interior rest points or the boundaries of the simplex.

7.3.9   Equilibrium and optimality

Let $\mathcal{U}_s$ be the set of strategies. Consider the optimal control problems

$$(OPT_N) \begin{cases} \text{Maximize } R^N(u, u; s, \vec{m}_0) \\ \text{s.t } u \in \mathcal{U}_s \end{cases} \quad (OPT_\infty) \begin{cases} \text{Maximize } R(u, u; s, \vec{m}_0) \\ \text{s.t } u \in \mathcal{U}_s \end{cases}$$

The strategy u is an $\epsilon-$optimal strategy for the N-optimal control problem if $R^N(u, u; s, \vec{m}_0) \geq -\epsilon + \sup_v R^N(v, v; s, \vec{m}_0)$.

Also consider the fixed-point problems

$$(FIX_N) \begin{cases} \text{find } u \in \mathcal{U}_s \text{ such that} \\ u \in \arg\max_{v \in \mathcal{U}_s}\{R^N(v, u; s, \vec{m}_0)\} \end{cases} \quad (FIX_\infty) \begin{cases} \text{find } u \in \mathcal{U}_s \text{ such that} \\ u \in \arg\max_{v \in \mathcal{U}_s}\{R^N(v, u; s, \vec{m}_0)\} \end{cases}$$

A solution to (FIX_N) or (FIX_∞) is a (Nash) equilibrium. We say that u is an $\epsilon-$equilibrium for the game with N [resp. $N \rightarrow \infty$] players if $R^N(u, u; s, \vec{m}_0) \geq \sup_v R^N(v, u; s, \vec{m}_0) - \epsilon$ [resp. $R(u, u; s, \vec{m}_0) \geq \sup_v R(v, u; s, \vec{m}_0) - \epsilon$].

Note that the definition of equilibrium and optimal strategy may depend on the initial conditions. If, for any $u \in \mathcal{U}_s$, the hypotheses in Section (7.3.7) hold, then we may relax this dependency.

Theorem 7.3.10 (Finite N). *For every discount factor $\beta > 0$ the optimal control problem (OPT_N) (resp. the fixed-point problem (FIX_N)) has at least one $0-$optimal strategy (resp. $0-$equilibrium). In particular, there a ϵ_N-optimal strategy (resp. ϵ_N-equilibrium) with $\epsilon_N \longrightarrow 0$.*

Sketch of Proof of Theorem 7.3.10

We show that for every discount factor $\beta > 0$ the optimal control problem (OPT_N) (resp. the fixed-point problem (FIX_N)) has at least one $0-$optimal strategy. It follows from the existence of equilibria in stationary strategies for finite stochastic games with discounted payoff: The set of pure strategies is a compact space in the product topology (Tykhonov theorem). Thus, the set of behavioral strategies Σ_j is a compact space and also convex as the set of probabilities on the pure strategies. For every player j and every strategy profile σ the marginal of the payoffs and constraints functions are continuous for any $\beta > 0 : \alpha_j \longmapsto R_j^N(\alpha_j, \sigma_{-j}, s, m_0)$. Moreover, the stationary strategies is convex, compact and upper and lower hemi-continuous (as a correspondence). Define

$$\gamma_j(s, m_0, \sigma) = \arg\max_{\alpha_j \in \mathcal{U}_s} R_j^N(\alpha_j, \sigma_{-j}, s, m_0).$$

Then, $\gamma_j(m_0, \sigma) \subseteq \Sigma_j$ is a non-empty, convex and compact set and the product correspondence

$$\gamma : \sigma \longmapsto (\gamma_1(s, m_0, \sigma), \ldots, \gamma_N(s, m_0, \sigma))$$

is upper hemi-continuous (its graph is closed). We now use the Glicksberg generalization of Kakutani fixed point theorem, and there is a stationary strategy profile σ^* such that

$$\sigma^* \in \gamma(s, m_0, \sigma^*).$$

Moreover, if the game has symmetric payoffs and strategies for each type, there is a symmetric per type stationary equilibrium. This completes the proof.

Theorem 7.3.11 (Infinite N)**.** *Optimal strategies (resp. equilibrium strategies) exist in the limiting regime when $N \to \infty$ under uniform convergence and continuity of $R^N \to R$. Moreover, if $\{U^N\}$ is a sequence of ϵ_N-optimal strategies (resp. ϵ_N-equilibrium strategies) in the finite regime with $\epsilon_N \longrightarrow \epsilon$, then, any limit of subsequence $U^{\phi(N)} \longrightarrow U$ is an $\epsilon-$ optimal strategy (resp. $\epsilon-$equilibrium) for game with infinite N.*

Sketch of Proof of Theorem 7.3.11

Let $(U^N)_N$ be a sequence of solution of (FIX_N) i.e equilibrium in the system with N players. Choose a subsequence N_k such that U^{N_k} converges to some point u when k goes to infinity. We can write $R^{N_k}(U^{N_k}, U^{N_k}) - R(U, U) = R^{N_k}(U^{N_k}, U^{N_k}) - R^{N_k}(U, U) + R^{N_k}(U, U) - R(U, U)$. Since $R^N(., .)$ is continuous and converges uniformly to $R(., .)$, R^{N_k} converges uniformly to R, the second term $R^{N_k}(U, U) - R(U, U) \longrightarrow 0$ when $N_k \longrightarrow \infty$ and the first term $R^{N_k}(U^{N_k}, U^{N_k}) - R^{N_k}(U, U)$ can be rewritten as $R^{N_k}(U^{N_k}, U^{N_k}) - R^{N_k}(U, U) = R^{N_k}(U^{N_k}, U^{N_k}) - R(U^{N_k}, U^{N_k}) + R(U^{N_k}, U^{N_k}) - R(U, U) + R(U, U) - R^{N_k}(U, U)$. Each term goes to zero by continuity of R, convergence of U^{N_k} to U, and by uniform convergence of R^N to R. Let U^N be a ϵ_N-equilibrium. Then, $R^N(U^N, U^N) \geq R^N(v, U^N) - \epsilon_N$, $\forall v$. Then any limit U of a subsequence of U^N satisfies $R(U, U) \geq R(v, U) - \epsilon$, $\forall v$. Similarly, if

$$R^N(U^N, U^N) \geq R^N(v, v) - \epsilon_N, \ \forall v$$

then any omega-limit U of the sequence of U^N satisfies $R(U, U) \geq R(v, v) - \epsilon$, $\forall v$ i.e U is an $\epsilon-$optimal strategy. In particular if $(U^N)_N$ is a sequence of ϵ_N-equilibria (resp. optimal strategies) with $\epsilon_N \longrightarrow 0$ when N goes to infinity then any accumulation point U of $(U^N)_N$ is a $0-$equilibrium (resp. $0-$optimal strategy).

7.3.12 Robust equilibrium

We now consider two types of players $\theta = 1, 2$ and define the expected payoff of a player with type 2, the strategy v whenever the other players of type 1 use u when the proportion of the players of type 2 is α as $\bar{R}^N(v, u, \alpha, m_0)$

Definition 7.3.12.1. *The strategy u is $k-$resilient to invasion if for all strategy $v \neq u$*

$$\bar{R}^N(v, u, \alpha, m_0) \leq \bar{R}^N(u, u, \alpha, m_0)$$

$$\forall \alpha \in \{0, \frac{1}{N}, \ldots, \frac{k}{N}\}$$

Since a Nash equilibrium is resilient against unilateral deviation (1-resilient). A $k-$resilient strategy is in particular 1-resilient.

Definition 7.3.12.2. *The strategy u is resilient to invasion by small fraction of deviants (RID) if for all strategy $v \neq u$ there exists an invasion barrier $\bar{\alpha}_v \neq 0$ such that*

$$\bar{R}(v, u, \alpha, m_0) < \bar{R}(u, u, \alpha, m_0) = \bar{R}(u, u, 0, m_0) \ \forall \alpha \in (0, \bar{\alpha}_v).$$

The strategy which is resilient to invasion by small fraction of deviant is more robust than a Nash equilibrium. A Nash equilibrium u is not necessarily a RID.

Proposition 7.3.12.3. *If u is a RID then u is an equilibrium.*

Proof.

$$\forall v, \; R(v, u, m_0) \;=\; \bar{R}(v, u, 0, m_0) \leq \bar{R}(u, u, 0, m_0) = R(u, u, m_0)$$

$\square$

Proposition 7.3.12.4 (RID for Infinite N). *If $\{U^N\}$ is a sequence of k_N-resilient against deviants strategies in the finite regime with $\dfrac{k_N}{N} \longrightarrow \epsilon > 0$, then, any limit of subsequence $U^{\phi(N)} \longrightarrow U$ is a neutrally RID[1] for game with infinite N.*

Proposition 7.3.12.5. *If u is an $\dfrac{\epsilon}{2}-$equilibrium of the Markov decision evolutionary game with infinite players then there exists N_0 such that for all $N \geq N_0$, u is an $\epsilon-$equilibrium of the mean field interaction with N players.*

Proof.

$$R(u, u, m_0) \geq R(v, u, m_0) - \epsilon/2, \; \forall v \in \mathcal{U}_b$$

Since R^N goes to R uniformly, choose N_0 satisfying

$$\forall N \geq N_0, \; \|R^N - R\|_\infty \leq \gamma_N \leq \epsilon/4,$$

$$R^N(u, u, m_0) - R^N(v, u, m_0) = R^N(u, u, m_0) - R(u, u, m_0) + R(u, u, m_0) - R(v, u, m_0)$$
$$+R(v, u, m_0) - R^N(v, u, m_0) \geq -\epsilon/2 - 2\gamma_N \geq -\epsilon$$

$\square$

Proposition 7.3.12.6. *Let u be a strict equilibrium. Then u is a RID for the payoff function $\bar{R}$ defined as*

$$\bar{R} = \alpha R(v, u, m_0) + (1 - \alpha) R(u, u, m_0).$$

Moreover u is non-invadable strategy (NIS) i.e the inequality for RID holds for any uniform threshold ϵ.

Proof. Since u is a strict equilibrium, one has:

$$R(v, u, m_0) < R(u, u, m_0), \; \forall v \neq u, \forall v \in \mathcal{U}_b$$

This implies that $\bar{R} < R(u, u, m_0), \; \forall \alpha \in (0, 1)$. This completes the proof for RID and for NIS. $\square$

7.4 Link with differential population

Each individual maximizes its long-term payoff

$$\bar{r}(u; x, \vec{m}) = \mathbb{E}\left(\int_0^\infty e^{-\beta t} r(u, X_1(t), m(t)) | X_1(0) = x, \vec{m}(0) = \vec{m}_0 \right)$$

subject to the population profile evolution given by

$$\frac{d}{dt} m(t) = f(u, m(t))$$

---

[1]A neutrally RID is a weaker notion of RID with non-strict inequality

and the rate of transition of the process $X_1(t)$ from state $x_1 = (\theta_1, s_1)$ to state $x_1' = (\theta_1, s_1')$ is $A(x_1, x_1'; \vec{m}(t), u)$.

We derive a new class of dynamic games called *differential population games*. Different from standard differential game models, this class of games is described by large population of players in which each player from each class is facing a random vector that evolves according to the population dynamics and the individual state follows a jump and drift process. See (207) for more details on differential population games.

7.5 Illustrating example

We present in this section an example of a dynamic version of the Hawk and Dove problem where each individual has three energy levels. We derive the mean field limit for the case where all users follow a given policy and where possibly one player deviates. We then further simplify the model to only two energy states per player. In that case we are able to fully identify and compute the equilibrium in the limiting Markov decision evolutionary game. Interestingly, we show that the ODE converges to a fixed point which depends on the initial condition.

Consider an homogenous population of N animals. An animal plays the role of a *player*. Occasionally two animals find themselves in competition on the same piece of food. Each animal has three states $x = 0, 1, 2$ which represents its energy level. An animal can adopt an aggressive behavior (Hawk) or a peaceful one (Dove, passive attitude). At the state $x = 0$ there is no action. We describe the fitness of an animal (some arbitrary player) associated with the possible outcomes of the meeting as a function of the decisions taken by each one of the two animals. The fitnesses represent the following:

- An encounter Hawk-Dove or Dove-Hawk results in zero fitness to the Dove and in $\bar{v}$ of value for the Hawk that gets all the food without fight. The state of the Hawk (the winner) is incremented $a = 1_{\{x_H' = \min(x_H + 1, 2)\}}$ and the state of the Dove is $b = 1_{\{x_D' = \max(x_D - 1, 0)\}}$.

- An encounter Dove-Dove results in a peaceful, equal-sharing of the food which translates to a fitness of $\dfrac{\bar{v}}{2}$ to each animal and the state of each animal change with the sum of the two distributions $\dfrac{1}{2}a + \dfrac{1}{2}b$

- An encounter Hawk-Hawk results in a fight in which with $p = 1/2$ chances, one (resp. the other) animal obtains the food but also in which there is a positive probability for each one of the animals to be wounded $1/2$. Then the fitness of the animal 1 is $\dfrac{1}{2}(\bar{v} - c) + \dfrac{1}{2}(-c) = \dfrac{1}{2}\bar{v} - c$, where the $-c$ term represents the expected loss of fitness due to being injured.

$i \backslash j$	(g_i^N, g_j^N)	$X_i^N(t + \frac{1}{N}), X_j^N(t + \frac{1}{N})$
$D - D$	$(\frac{\bar{v}}{2}, \frac{\bar{v}}{2})$	$\frac{1}{2}\delta_{\min(x_1 - 1, 0), \max(x_2 + 1, 2)}$ $+ \frac{1}{2}\delta_{\max(x_1 + 1, 2), \min(x_2 - 1, 0)}$
$D - H$	$(0, v)$	$(\min(x_1 - 1, 0), \max(x_2 + 1, 2))$
$H - H$	$\frac{1}{2}v - c$	$\frac{1}{2}\delta_{\min(x_1 - 1, 0), \max(x_2 + 1, 2)}$ $+ \frac{1}{2}\delta_{\max(x_1 + 1, 2), \min(x_2 - 1, 0)}$

The vector of frequencies of states at time t is given by $M_x^N(t) = \dfrac{1}{N} \sum_{j=1}^{N} 1_{\{X_j^N(t) = x\}}$ for $x = 0, 1, 2$

and the action set is $A_x = \{H, D\}$ in each state $x \neq 0$, $A_0 = \{\}$.

The assumptions in Section 6.2 are satisfied (pairwise interaction, $\sharp \mathcal{B}^N(t) = 2$) and the occupancy measure $M^N(t)$ converges to $m(t)$.

7.5.1 ODE and Stationary strategies

Consider the following fixed parameters $\mu_1 = L_{0,1}$, $\mu_2 = L_{0,2}$. The population profile is denoted by $\vec{m} = (m_0, m_1, m_2)$ and the stationary strategy is described by the parameters v_1, v_2 where $v_1 := u(H|1)$, $v_2 = u(H|2)$

$$\begin{cases} \dot{m}_2 = m_0 L_{0,2} + m_1 L_{1,2}(u, m) - m_2 L_{2,1}(u, m) \\ \dot{m}_1 = m_0 L_{0,1} + m_2 L_{2,1}(u, m) - m_1 L_{1,2}(u, m) - m_1 L_{1,0}(u, m) \\ \dot{m}_0 = m_1 L_{10}(u, m) - (\mu_1 + \mu_2) m_0 \end{cases}$$

where

$$L_{12}(u, m) = m_0 + v_1 \left(\frac{v_1 m_1}{2} + (1 - v_1) m_1 + \frac{v_2 m_2}{2} + (1 - v_2) m_2 \right) + (1 - v_1) \left(\frac{(1 - v_1) m_1}{2} + \frac{(1 - v_2) m_2}{2} \right)$$

$$L_{2,1}(u, m) = v_2 \left(\frac{v_1 m_1}{2} + \frac{v_2 m_2}{2} \right) + (1 - v_2) \left(\frac{(1 - v_1) m_1}{2} + v_2 m_2 + \frac{(1 - v_2) m_2}{2} \right)$$

$$L_{10}(u, m) := v_1 \left(\frac{v_1 m_1}{2} + \frac{v_2 m_2}{2} \right) + (1 - v_1) \left(v_1 m_1 + \frac{(1 - v_1) m_1}{2} + v_2 m_2 + \frac{(1 - v_2) m_2}{2} \right),$$

For $\mathcal{B}^N = \{j_1, j_2\}$, $x_j', x_i \in \{0, 1, 2\}$,

$$\begin{aligned} \frac{d}{dt} m_x &= \sum_{x_1, x_2, x_2'} m_{x_1} m_{x_2} L_{x_1, x_2; x, x_2'}(u, \vec{m}) + \sum_{x_1, x_2, x_1'} m_{x_1} m_{x_2} L_{x_1, x_2; x_1', x}(u, \vec{m}) \\ &\quad - m_x \sum_{x_2, x_1', x_2'} m_{x_2} L_{x, x_2; x_1', x_2'}(u, \vec{m}) - m_x \sum_{x_1, x_1', x_2'} m_{x_1} L_{x_1, x; x_1', x_2'}(u, \vec{m}) \end{aligned}$$

7.5.2 Computation of $R(u_1, u_2; s, \vec{m})$.

We want to compute the value

$$V(u_1, u_2, x) := \mathbb{E}_x \int_0^\infty e^{-\beta t} r(u_1, u_2, x(t), m(t)) \, dt$$

$$s.t. \ \dot{m}(t) = f(u_2, m(t)), m(0) = m_0, \ x(0) = x.$$

$$\begin{aligned} V(u_1, u_2, x) &= \mathbb{E}_x \int_0^\Delta e^{-\beta t} r(u_1, u_2, x(t), m(t)) \, dt + \mathbb{E}_x \int_\Delta^\infty e^{-\beta t} r(u_1, u_2, x(t), m(t)) \, dt \\ &= \mathbb{E}_x \int_0^\Delta e^{-\beta t} r(u_1, u_2, x(t), m(t)) \, dt + \mathbb{E}_x e^{-\beta \Delta} V(u_1, u_2, x(\Delta)) \end{aligned}$$

This implies that

$$0 = \mathbb{E}_x \frac{1}{\Delta} \int_0^\Delta e^{-\beta t} r(u_1, u_2, x(t), m(t)) \, dt + \frac{e^{-\beta \Delta} - 1}{\Delta} \mathbb{E}_x V(u_1, u_2, x(\Delta)) + \frac{\mathbb{E}_x V(u_1, u_2, x(\Delta)) - V(u_1, u_2, x)}{\Delta}$$

195

Using Ito's formula and Lebesgue integration properties, we obtain that: $\dfrac{\mathbb{E}_x V(u_1, u_2, x(\Delta)) - V(u_1, u_2, x)}{\Delta}$ goes to $\sum_{x'} D_{m_{x'}} V(u_1, u_2, x') \dfrac{d}{dt} m_{x'}$, where $D_{m_{x'}} V$ is the derivative of V in a weak sense, $\dfrac{e^{-\beta\Delta} - 1}{\Delta} \longrightarrow -\beta$, and the term

$$\mathbb{E}_x \frac{1}{\Delta} \int_0^{\Delta} e^{-\beta t} r(u_1, u_2, x(t), m(t)) \, dt \longrightarrow r(u_1, u_2, x, m_0)$$

when Δ goes to zero. Thus, we obtain

$$\beta V(u_1, u_2, x) = r(u_{1,x}, u_{2,x}, x, m) + \sum_{x'} (D_{m_{x'}} V(u_1, u_2, x')) f_{x'}(u_2, m) \tag{7.13}$$

where $u_{i,x} = u_i(H|x)$.

The global optimality is then given by the Hamilton-Jacobi-Bellman equation obtained by maximizing the right-hand side of the equation (7.13) over the action set.

$$\beta \Psi(x) = \max_{u_{1,x}, u_{2,x}} \left\{ r(u_{1,x}, u_{2,x}, x, m) + \sum_{x'} (D_{m_{x'}} \Psi(u_1, u_2, x')) f_{x'}(u_2, m) \right\}$$

and optimality conditions of the best response to u_2 is given by

$$\beta \Phi(u_2, x) = \max_{a \in \{H, D\}} \left\{ r(a, u_{2,x}, x, m) + \sum_{x'} (D_{m_{x'}} \Phi(u_2, x')) f_{x'}(u_2, m) \right\}$$

Theses equations are in general difficult to solve and the solutions are not necessarily regular (e.g. viscosity solutions). Numerical approaches based on multi-grid techniques of Hamilton-Jacobi-Bellman-Issacs equations can be found in (128).

7.5.3 The case of two energy levels

In order to derive closed form expressions for solutions of our ODE, we consider two states, i.e., each animal has two states $x = 1, 2$ which represents its energy levels. Thus, the ODE can be expressed as follows:

$$\dot{m}_2(t) = (1 - m_2(t)) L_{1,2}(u, m) - m_2(t) L_{2,1}(u, m) \tag{7.14}$$

which can be rewritten as

$$\dot{m}_2(t) = a_1 + a_2 m_2(t) + a_3 (m_2(t))^2 \tag{7.15}$$

with $a_1 = 1$, $a_2 = \dfrac{u_2}{2} - 2 < 0$, $a_3 = \dfrac{1 - u_2}{2} > 0$.

Let $m[u, m_0](t)$ be the solution of the ODE given u and a initial distribution $m(0) = m_0$. We distinguish two cases:

Case 1: $u_2 = 1$ (fully aggressive when it is possible): the ODE becomes $\dot{m}_2(t) = 1 - \dfrac{3}{2} m_2(t)$ and the solution has the form

$$m_2[1, m_0](t) = \frac{2}{3} [1 - c_1 e^{-\frac{3}{2}t}] \tag{7.16}$$

with $c_1 = 1 - \dfrac{3}{2} m_0$ and $m_1[u, m_0](t) = 1 - m_2[u, m_0](t)$

Case 2: $u_2 \neq 1$, (less aggressive in state 2)

$$m_2[u, m_0](t) = \gamma_-(u) + \frac{\gamma_+(u) - \gamma_-(u)}{1 - c_2 e^{(\gamma_+(u) - \gamma_-(u))a_2 t}}, \tag{7.17}$$

$$c_2 = 1 + \frac{\gamma_+(u) - \gamma_-(u)}{m_2(0) - \gamma_-(u)}, \quad \gamma_-(u) = \frac{2 - u_2/2 - (2 + u_2^2/4)^{\frac{1}{2}}}{1 - u_2} < 1,$$

$$\gamma_+(u) = \frac{2 - u_2/2 + (2 + u_2^2/4)^{\frac{1}{2}}}{1 - u_2} > 1$$

Note that in both cases there is a unique strategy-dependent global attractor.

$$\lim_{t \longrightarrow \infty} m_2[u, m_0](t) = \left\{ \begin{array}{l} \gamma_-(u) \text{ if } u_2 \neq 1 \\ 2/3 \text{ if } u_2 = 1 \end{array} \right.$$

The expected instant payoff of a player using the stationary strategy v when the population profile is $m[u, m_0](t)$, is given by

$$r(v, u, 2, m[u, m_0](t)) = v[\bar{v} - cm_2 u_2] + (1 - v)r(v, u, 1, m[u, m_0](t))$$

$$r(v, u, 1, m[u, m_0](t)) = \frac{1}{2}(1 - m_2[u, m_0](t)u_2)\bar{v}$$

where $m_2[u, m_0](t)$ is given by (7.16) (resp. (7.17)) for $u_2 = 1$ (resp. $u_2 \neq 1$). Now, we can compute explicitly the best response against u for a given initial m_0. Let

$$\beta_2(u, 2, m_0, t) = r(H, u, 2, m[u, m_0](t)) - r(D, u, 2, m[u, m_0](t)).$$

The best response, $\text{BR}(x, u, m[u, m_0](t))$, against u at t is

$$\text{BR}(x, u, m[u, m_0](t)) = \left\{ \begin{array}{l} \text{play Hawk if } \beta_2(u, x, m_0, t) > 0 \\ \text{play Dove if } \beta_2(u, x, m_0, t) < 0 \end{array} \right.$$

This implies that it is better to play Hawk for $\frac{\bar{v}}{2c} > \frac{\gamma}{1 + \gamma}$ where $\gamma = \max(2/3, m_0)$. Since the solution of the ODE is strictly monotone in time for each stationary strategy, there is at most one time for which β_2 is zero. It is easy to see that if $\frac{\bar{v}}{2c} > \frac{2}{3}$ then the strategy which to play Hawk in state 2 and Dove in state 1 is an equilibrium.
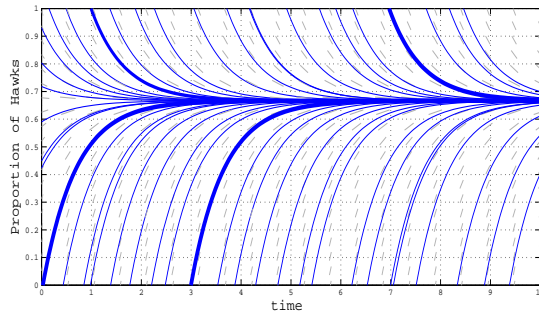
Figure 7.1: Global attractor for $u_2 = 1$
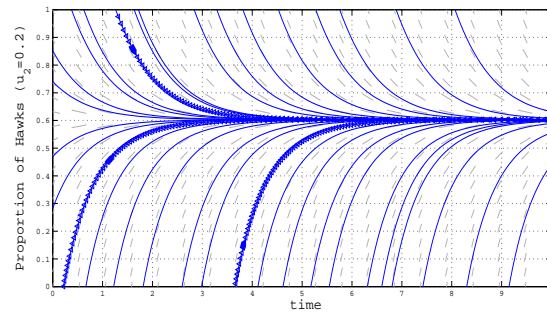
Figure 7.2: *Global attractor for* $u_2 = 0.2$

7.6   Notes

The goal of this chapter has been to develop mean field asymptotic of interactions with large number of players using stochastic games. Due to the curse of the size of the population, the applicability of atomic stochastic games has been severely limited. As an alternative, we proposed a method for Markov decision evolutionary games where players make decisions only based on their own state and the global system state. We have showed under mild assumptions convergence results, where asymptotics were taken in the number of players. The population state profile satisfies a system of non-linear ordinary differential equations. We have considered very simple class of strategies that are functions only of player's own state and the population profile. We applied to Hawk-Dove interaction with several energy level and formulated the ODEs. We show that the best response depends on the initial conditions.

List of Figures

Index

List of publications

Journal papers

- H. Tembine, E. Altman, R. El-Azouzi, Y. Hayel , Evolutionary Games in Wireless Networks , accepted and to appear in IEEE Transactions on Systems, Man and Cybernetics, 2009.

- Y. Hayel, H. Tembine, E. Altman, R. El-Azouzi , A Markov Decision Evolutionary Game for Individual Energy Managaement , accepted and to appear in Annals of Dynamics Games, 2009.

- H. Tembine, E. Altman, R. El-Azouzi, Y. Hayel, *Delayed evolutionary game dynamics with networking applications*, accepted and to appear in Telecommunication Systems Journal, 2009.

- E. Altman , R. El-Azouzi , Y. Hayel , H. Tembine , The evolution of transport protocols: an evolutionary game perspective, Computer Networks Journal, 2009, 53, 10, 1751-1759, July

- E. Altman, Y. Hayel, H. Tembine, R. El-Azouzi , Markov Decision Evolutionary Games with Expected Average Fitness , Evolutionary Ecology Research EER Journal, Special issue in honor of Thomas L. Vincent, May 2009.

Book Chapter

- H. Tembine , E. Altman, Y. Hayel, R. El-Azouzi , Evolutionary Games, Part I, Chapter 3, 31-42, in Altman E., Dini P., Miorandi D., Schreckling, Foundations of BIONETS research, 220 pages, 2008.

International Conferences and Workshops

- H. Tembine, Evolutionary Network Formation Games and Fuzzy Coalitions in Heterogeneous Networks, IFIP WIRELESS DAYS 2009.

- H. Tembine Dynamic Bargaining Solutions for Opportunistic Spectrum Access, IFIP WIRELESS DAYS 2009.

- Q. Zhu, H. Tembine and T. Basar, Evolutionary Games for Additive White Gaussian Noise Multiple Access Control, accepted in IEEE GLOBECOM 2009.

- H. Tembine, J.-Y. Le Boudec, R. ElAzouzi, E. Altman , From Mean Field Interaction to Evolutionary Game Dynamics, In Proc. of IEEE/ACM Proc. 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WIOPT), presented at Physics-Inspired Paradigms in Wireless Communications and Networks (PHYSCOM-NET), June 2009.

- G. HE , H. Tembine and M. Debbah , Deployment Analysis of cooperative OFDM base stations , In Proc. of International IEEE Conference on Game Theory for Networks (GameNets), May, 2009.

- H. Tembine , J. Y. Le Boudec , R. El-Azouzi, E. Altman , Mean Field Asymptotic of Markov decision evolutionary games, In Proc. of International IEEE Conference on Game Theory for Networks (GameNets), May, 2009.

- E. Altman, T. Basar, I. Menache , H. Tembine , A dynamic random access game with energy constraints , In IEEE/ACM Proc. 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WIOPT),2009.

- H. Tembine , E. Altman , R. El-Azouzi , Y. Hayel , Correlated evolutionarily stable strategies in Random Medium Access Control, In Proc. of International IEEE Conference on Game Theory for Networks (GameNets), 2009.

- Q. Zhu , H. Tembine , T. Basar , A Constrained Evolutionary Gaussian Multiple Access Channel Game , In Proc. of International IEEE Conference on Game Theory for Networks (GameNets), May, 2009.

- H. Tembine , E. Altman , R. El-Azouzi, W. H. Sandholm , Evolutionary game dynamics with migration for hybrid power control in wireless communications , IEEE Proc. of 47th IEEE Conference on Decision and Control, December 2008.

- H. Tembine , E. Altman, R. El-Azouzi, Y. Hayel , Stable Networking Games, In Proc. of Forty-Sixth Annual Allerton Conference on Communication, Control, and Computing, Allerton Retreat Center, Monticello, Illinois, October 2008.

- H. Tembine , E. Altman , R. El-Azouzi , Y. Hayel , Battery State-dependent Access Control in Solar-powered Broadband Wireless Networks, in Proc. of Second International EuroFGI Conference on Network Control and Optimization (NETCOOP 2008), Proceedings of Springer LNCS, September 2008.

- E. Altman , Y. Hayel , H. Tembine, R. El-Azouzi , Markov decision evolutionary games with time average expected fitness criterion, ACM Proc. of the Third International Conference on Performance Evaluation Methodologies and Tools (ValueTools), October 2008.

- H. Tembine , E. Altman , R. El-Azouzi , Y. Hayel , Stochastic population games with individual independent states and coupled constraints, ACM Proc. of the International Workshop on Game Theory in Communication Networks, October 2008.

- E. Altman, R. El-Azouzi, Y. Hayel, H. Tembine , An Evolutionary Game approach for the design of congestion control protocols in wireless networks , IEEE/ACM Proc. of 1st Workshop on Physics-Inspired Paradigms in Wireless Communications and Network, October 2008.

- E. Altman, R. El-Azouzi, Y. Hayel , H. Tembine , Evolutionary Power Control Games in Wireless Networks , Book chapter in A. Das et al. (Eds.): IFIP NETWORKING, LNCS 4982, pp. 930-942, 2008.

- H. Tembine , E. Altman, R. El-Azouzi, Y. Hayel , Evolutionary games with random number of players applied to access control , IEEE/ACM Proc. 6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WIOPT), April 2008.

- H. Tembine, E. Altman, R. El-Azouzi , Delayed Evolutionary Game Dynamics applied to Medium Access Control , In Proc. of 4th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS), October 2007.

- Y. Hayel and H. Tembine , Information Dissemination using Epidemic Routing with Delayed Feedback , In IEEE Proc. of Socially and Biologically Inspired Wired and Wireless Networks (BIONETWORKS), 2007.

- H. Tembine , E. Altman, R. El-Azouzi, Y. Hayel , Multiple Access Game in Ad-hoc Network , In ACM Proc. International Workshop on Game Theory in Communication Networks (GameComm), 2007.

- H. Tembine, E. Altman , R. El-Azouzi , Asymmetric Delay in Evolutionary Games , in Proc. of the Second International Conference on Performance Evaluation Methodologies and Tools (ValueTools), 2007.

- H. Tembine , A. Kobbane, K. Ibrahimi , Minority game applied to Selfish Routing , JOS-TIC, 2007.

National conferences

- H. Tembine , E. Altman , R. El-Azouzi, Y. Hayel, EvolNet: Evolutionary games with networking applications, 9eme congres de la Société Fraçaise de Recherche Opérationnelle et d'Aide à la Décision, ROADEF'08/MODE'08.

- H. Tembine, E. Altman, R. El-Azouzi , Conflits entre voisins dans les réseaux mobiles , DOCTISS, 2008.

Intership Report and thesis dissertations

- H. Tembine , Folk theorem for stochastic games with public signals, Intership report, University of Joseph Fourier, France, 2006.

- H. Tembine, Repeated games with imperfect monitoring, Master Thesis Dissertation, Ecole Polytechnique, France, 2006.

- H. Tembine , Computation of Betti numbers in four dimensional spaces, Master Thesis Dissertation, University of Konstanz, Germany, 2005.

Under Submission, Technical report

- H. Tembine , Evolutionary Networking Games , Technical report, University of Avignon, 40 pages, June 2009.

- H. Tembine , Delayed Mean Field Game Dynamics with Migration, Technical report, University of Avignon, France, January, 2009.

- H. Tembine , Dynamic Games and Networks Evolution, Technical report, University of Avignon, France, 700 pages, June 25, 2009.

- H. Tembine , Differential population games, Technical report, University of Avignon, France, August 2009.

- A. Schaffhauser, H. Petitjean, H. Tembine, Biomathematics, Technical report, October 2008.

- H. Tembine, Evolutionary games with continuous action spaces. Part I : rectangular domains. Part II : applications, Technical report, March 2008.

- H. Tembine, Evolutionary games with incomplete information. Part I : dynamics and evolution of beliefs. Part II : application to network security, Technical report, March 2008.

Short abstract

This manuscript presents dynamic foundations of population games with variable number of players and their solutions and stability concepts. We first introduce delayed evolutionary game dynamics and study their stability. Applications to both wired and wireless networks are presented. We then introduce mobility and spatial aspects of players distribution into the network dynamics. This leads to a new class of game dynamics with multicomponent strategies and migration constraints called *evolutionary game dynamics with migration*. We derived such dynamics for hybrid systems such as power control in heterogenous networks, switching between technologies and migration between different classes of users. After that we focus on stochastic population games with multiple classes of players in which each player has its own state and facing to an evolving vector which represents the population profile. We use this model to analyze resource and energy constrained interactions in wireless networks. Finally, we present a class of mean field games. When taking the asymptotics of finite systems, we derive a new class of game dynamics called *mean field game dynamics*. This class contains the standard evolutionary game dynamics based on revision of pure actions. We apply this model to analyze spatial random access game and dynamic resource competition game with individual states. We establish a link between mean field games and *differential population games* in which each player optimizes its long-term objective during its sojourn time in the system subject to the constraint that the population profile evolves according to some mean field game dynamics.

References

[1] O Ait-Hellal and E. Altman. *Analysis of TCP Vegas and TCP Reno,* Telecommunication Systems, 15:381–404, 2000.

[2] O. Ait-Hellal , E. Altman, D. Elouadghiri, M. Erramdani, N. Mikou, *Performance of TCP/IP: the case of two Controlled Sources*, ICCC'97, Cannes, France, November 19-21, 1997.

[3] S. Alouf, I. Carreras, D. Miorandi, G. Neglia, *Evolutionary Epidemic Routing*, Technical Report, INRIA, 2007.

[4] T. Alpcan and T. Başar, *A Hybrid Noncooperative Game Model for Wireless Communications*, in Proc. of 11th International Symposium on Dynamic Games and Applications, Tuscon, AZ, December 2004.

[5] E. Altman, K. Avratchenkov, C. Barakat, A. A. Kherani and B. J. Prabhu, *Analysis of scalable TCP*, 7th IEEE International Conference on High Speed Networks and Multimedia Communications (HSNMC'04) June 30 - July 2, 2004 Toulouse, France.

[6] E. Altman, K. Avrachenkov, and B. Prabhu. *Fairness in MIMD congestion control algorithms, Telecommunication Systems*, 30(4): 387–415.

[7] E. Altman, D. Barman, B. Tuffin and M. Vojnovic, *Parallel TCP Sockets: Simple Model, Throughput and Validation, IEEE Infocom*, 2006.

[8] E. Altman, S. Sarkar, E. Solan, Constrained Markov games with transition probabilities controlled by a single player, Proceedings of the 2nd international conference on Performance evaluation methodologies and tools, 2007.

[9] E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-Azouzi and D. Sadoc Menasche, *Constrained Cost-Coupled Stochastic Games with Independent State Processes*, Operations Research Letters, Vol 36, pp 160-164, 2008.

[10] E. Altman, Y. Hayel , H. Tembine, R. El-Azouzi , *Markov decision Evolutionay Games with Time Average Expected Fitness Criterion*, 3rd International Conference on Performance Evaluation Methodologies and Tools, (Valuetools), Athens, Greece, October, 2008

[11] E. Altman, R. ElAzouzi, Y. Hayel and H. Tembine, *Evolutionary Power Control Games in Wireless Networks*, in Proc. 7th IFIP Networking 2008.

[12] E. Altman, R. ElAzouzi, Y. Hayel and H. Tembine , An Evolutionary Game approach for the design of congestion control protocols i n wireless networks, Physicomnet, Berlin, April 4, 2008.

[13] E. Altman, R. ElAzouzi, Y. Hayel, H. Tembine , *The evolution of transport protocols: an evolutionary game perspective*, Elsevier Computer Networks Journal, Volume 53, Issue 10, Pages 1751-1759, July 2009

[14] E. Altman and Z. Altman . S-modular games and power control in wireless networks. IEEE Trans. on Automatic Control, 48(5):839-842, 2003.

[15] E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-Azouzi, and D. Sadoc Menasche. Constrained stochastic games in wireless networks. In *IEEE Globecom General Symposium*, Washington D.C., 2007.

[16] E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-Azouzi, and D. Sadoc Menasche. Constrained cost-coupled stochastic games with independent state processes. *Operations Research Letters*, 36:160–164, 2008.

[17] E. Altman, J. Galtier, C. Touati , *Fair power and transmission rate control in wireless networks*, IEEE Globecom, November 2002, Taipei, Taiwan.

[18] Altman et al., A survey on networking games in telecommunications, Computers and Operations Research, Volume 33, Issue 2 (February 2006), 286 - 311.

[19] E. Altman :*Constrained Markov Decision Processes*, Chapman and Hall/CRC, 1999.

[20] E. Altman and Y. Hayel : *A Stochastic Evolutionary Game of Energy Management in a Distributed Aloha Network*, In Proc. of INFOCOM 2008.

[21] E. Altman, A. Hordijk, F. M. Spiekma :*Contraction conditions for average and $\alpha-$discounted optimality in countable stat Markov games with unbounded rewards* , 1996.

[22] E. Altman and A. Shwartz , *Constrained Markov Games: Nash Equilibria* , Annals of the International Society of Dynamic Games , vol. 5, Birkhauser, V. Gaitsgory, J. Filar and K. Mizukami, editors, pp. 303-323, 2000.

[23] E. Altman, K. Avrachenkov, R. Marquez *Zero-sum constrained stochastic games with independent state processes*, Mathematical Methods of Operations Research, 62,375-386,2005.

[24] Altman E. and Hayel Y., *Stochastic Evolutionary Games*, Proceedings of the 13th Symposium on Dynamic Games and Applications, 30th June-3rd July, 2008.

[25] E. Altman, Y. Hayel, H. Tembine, R. El-Azouzi , *Markov decision Evolutionay Games with Time Average Expected Fitness Criterion*, Evolutionary Ecology Research EER Journal, Special issue in honor of Thomas L. Vincent 2009.

[26] E. Altman, *Constrained Markov Decision Process*, Chapman and Hall/CRC, 1999

[27] J. Arino, L. Wang, and G. Wolkowicz, *An Alternate Formulation for a Delayed Logistic Equation*, Journal of Theoretical Biology, vol. 241, no. 1, 2006.

[28] N. Andelman, M. Feldman, Y. Mansour , *Strong Price of Anarchy*. In SODA'07, 2007.

[29] E. Anshelevich, A. Dasgupta, J. Kleinberg , E. Tardos ., T. Wexler and T. Roughgarden , *The Price of Stability for Network Design with Fair Cost Allocation*. In 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS), pages 59-73, 2004.

[30] N. Andelman, M. Feldman, Y. Mansour , *Strong Price of Anarchy*. In SODA'07, 2007.

[31] G. Anastassiou , *Quantitative Approximations*, Chapman & Hall/CRC, 2000.

References

[32] R. Auman, *Subjectivity and Correlation in Randomized Strategies*, Journal of Mathematical Economics 1, 67-96, 1974.

[33] R. Aumann, *Correlated Equilibrium as an Expression of Bayesian Rationality*, Econometrica 55, 1-18, 1987.

[34] R. Aumann, *Acceptable Points in General Cooperative n-Person Games*. In Contributions to the Theory of Games, volume 4, 1959.

[35] R. Axelrod *The Evolution of Cooperation*, New York: Basic Books, 1984.

[36] C. Bardenave, D. McDonald, A. Proutiere,*Performance of Random Medium Access Control An asymptotic approach*, ACM Sigmetrics 2008.

[37] T. Başar and G. J. Olsder. Dynamic Noncooperative Game Theory. SIAM, 1999.

[38] T. Basar , R. Srikant : *A stackelberg network game with a large number of followers*, Journal of Optimization Theory and Applications, 115(3):479–90, 2002.

[39] M. J. Beckmann, C. B. McGuire, and C. B. Winsten,*Studies in the Economics of Transportation*. Yale University Press, New Haven, CT, 1956.

[40] R. Bellman and K.L Cooke, *Differential difference equations*, Academic Press, New York, 1963.

[41] M. Benaim and J. W. Weibull . *Deterministic Approximation of Stochastic Evolution in Games*, Econometrica 71, 873-903, 2003.

[42] M. Benaïm, J. Hofbauer and W. H. Sandholm, *Robust Permanence and Impermanence for the Stochastic Replicator Dynamic*.

[43] M. Benaim and J.-Y. Le Boudec, A Class Of Mean Field Interaction Models for Computer and Communication Systems Performance Evaluation, Vol. 65, Nr. 11-12, pp. 823-838, 2008.

[44] M. Benaïm , J. Hofbauer , E. Hopkins, *Learning in Games with Unstable Equilibria*.mimeo, 2005.

[45] L. Berezansky and E. Braverman, *On stability of some linear and nonlinear delay differential equations*, J. Math. Anal. Appl., Vol. 314, No. 2, pages 391-411, 2006.

[46] U. Berger, *Fictitious Play in 2xN Games*, Journal of Economic Theory 120, 139-154, 2005.

[47] Joseph L.F. Bertrand, *Théorie des Richesses: revue de Théories mathématiques de la richesse sociale par Léon Walras et Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot*, Journal des Savants, 1883.

[48] M. Blonski, *Characterization of equilibria in large anonymous games*, University of Mannheim, 2000.

[49] M. Bloem, T. Alpcan, and T. Basar,*A stackelberg game for power control and channel allocation in cognitive radio networks*, in Proc. Workshop on Game Theory in Communication Networks (GameComm), Nantes, France, October 2007.

[50] L.E.J. Brouwer, ́*Uber Abbildung von Mannigfaltigkeiten*, Math. Ann. 71, 97-115, 1912.

[51] T. Bonald, *Comparison of TCP Reno and TCP Vegas: Efficiency and Fairness*, Perform. Eval. 36-37(1-4): 307-332, 1999.

[52] N. Bonneau, E. Altman, M. Debbah, and G. Caire. *An Evolutionary Game Perspective to ALOHA with Power Control*. 19th International Teletraffic Congress, Beijing, China, Aug. 29-Sep., 2005.

[53] D. Braess, *Uber ein Paradoxon aus der Verkehrsplanung. Unternehmensforschung* 12, 258 - 268, 1969.

[54] Brakmo L.S., O'Malley S., and L.L. Peterson. *Tcp vegas: New techniques for congestion detection and avoidance*, Computer Communication Review, Vol. 24, No. 4, pp. 24-35, Oct., 1994.

[55] G. Brown and J. Von Neumann. Solutions of games by differential equations. In H. Kuhn and A. Tucker, editors, Contributions to the Theory of Games I, Annals of Mathematics Studies 24, pages 73-79. Princeton University Press, 1950.

[56] Kim C. Border, *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, 1989.

[57] G.W. Brown, *Iterative Solutions of Games by Fictitious Play*, In Activity Analysis of Production and Allocation, T.C. Koopmans (Ed.), New York: Wiley, 1951.

[58] A. Cahn, *General procedure leading to correlated equilibria*, Int. Journal of Game theory 33: 21-40, 2004.

[59] Lijun Chen, Steven H. Low and John C. Doyle, *Random Access Game and Medium Access Control Design*, 2006.

[60] Gilbert G. Chen and Joel W. Branch and Boleslaw K. Szymanski, *A Self-selection Technique for Flooding and Routing in Wireless Ad-hoc Networks*, J. Netw. Syst. Manage., 14,3, pp.359–380, 2006.

[61] Gilbert Chen and Joel W. Branch and Boleslaw K. Szymanski, *Local Leader Election, Signal Strength Aware Flooding, and Routeless Routing*, Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05) 2005.

[62] S. Chong, R. Nagarajan, and Y. T. Wang, *Designing stable ABR flow control with rate feedback and open loop control: First-order control case*, Perform. Eval., vol. 34, pp. 189Ű206, 1998.

[63] V. Corradi and R. Sarin, *Continuous Approximations of Stochastic Evolutionary Game Dynamics*, Journal of Economic Theory Volume 94, Issue 2, Pages 163-191, October 2000.

[64] A. A. Cournot, *Recherches sur les principes mathématiques de la théorie des richesses*, Researches into the Mathematical Principles of the Theory of Wealth, 1838 (1897, Engl. trans. by N.T. Bacon).

[65] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*, MIT Press, 2003.

[66] Currarini S., Marini M. A.,*A Conjectural Cooperative Equilibrium for Strategic Form Games*, Game Practice and the Environment, Ed. Edgar Elgar, London, 2003.

[67] B.S. Cushing , *When a hawk can damage a dove: an extension of game theory*, Journal of Theoretical Biology, Volume 175, Number 2 , pp. 173-176(4), 1995.

[68] A. Datta, S. Quarteroni, K. Aberer, *Autonomous Gossiping: A self-organizing epidemic algorithm for selective information dissemination in wireless mobile ad-hoc networks*, in proceedings of International Conference on Semantics of a Networked World, 2004.

[69] E.V. Denardo, E. A. Feinberg and U. G. Rothblum, *Splitting of randomized stationary policies in absorbing Markov decision processes*, INFORMS 2007.

[70] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart and D. Terry *Epidemic Algorithms for Replicated Database Maintenance*, ACM Symposium on Principles of distributed computing, 1987.

[71] Dutta P.K.:*A Folk theorem for stocastic games*, Journal of Economic Theory, 66,pp.1-32, 1995.

[72] R. El-Khoury and E. Altman, *Analysis of Scalable TCP*, HET-NETs'04 International working conference on Performance Modelling and Evaluation of Heterogeneous Networks July, West Yorkshire, U.K., 2004.

[73] Ethier S. N. and Kurtz Thomas G.. *Markov Processes, Characterization and Convergence*. Wiley, 2005.

[74] Feinberg Eugene, Schwartz Adam, Handbook Of Markov Decision Processes: Methods and Applications, Kluwer 2002.

[75] Fan, Ky (1952). *Fixed-point and Minimax Theorems in Locally Convex Topological Linear Spaces*. Proc. Natl. Acad. Sci. USA. February; 38(2): 121–126, 1952.

[76] Falomari D. , Mandayam N. , and Goodman D. . *A new framework for power control in wireless data networks: games utility and pricing,* In Proc. Allerton Conference, pages 546–555, 1998.

[77] B. Fette, *Cognitive Radio Technology*, Elsevier Newnes, 2006.

[78] Mark Felegyhazi and Jean-Pierre Hubaux. *Game theory in wireless networks: A tutorial.* EPFL Technical report, 2006.

[79] J. Filar, T. Raghavan, *A Matrix Game Solution of the Single-Controller Stochastic Game*, Mathematics of Operations Research, vol. 9, no. 3, 1984.

[80] Fink A. M. *Equilibrium in a Stochastic $n-$Person Game*, J. Sci. Hiroshima University, 28, 89-93, 1964.

[81] Flesch J., Schoenmakers G., Vrieze K., Stochastic Games on a Product State Space, METEOR 016, Maastricht Research School of Economics of Technology, 2008.

[82] Flesch J., Thuijsman F., Vrieze O J.:*Stochastic games with additive transitions*, European Journal of Operational Research, 2007.

[83] Flesch J., Schoenmakers G.,Vrieze O J.:*Stochastic games on a product space*, Research Memoranda 010, METEOR, Maastricht Research School of Economics of Technology and Organization, 2007.

[84] W. H. Fleming. *Optimal Continuous-Parameter Stochastic Control*. SIA

[85] Forges, F., *An approach to communication equilibria*. Econometrica 54, pp. 1375-1385, 1985.

[86] D. Foster and P. Yong, *Stochastic Evolutionary game dynamics*, Theoreticial Population Biology, Vol 38 No 2, October 1990.

[87] Fudenberg, D. and D.K. Levine, The Theory of Learning in Games Cambridge: MIT Press, 1998.

[88] Gajic V., Rimoldi B. *Game theoretic considerations for the Gaussian Multiple Access Channel*, ISIT 2008.

[89] R. Garg, A. Kamra, V. Khurana, A game-theoretic approach towards congestion control in communication networks SIGCOMM Computer Communication Review archive **32**(3) (July 2002) 47 - 61.

[90] H. Gintis, *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, Princeton University Press, Second Edition, 2009.

[91] Glicksberg, I.L. . *A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium*. Proceedings of the American Mathematical Society 3 (1): 170–174, 1952.

[92] Goodman John C. *A Note on Existence and Uniqueness of Equilibrium Points for Concave N-Person Games*, Econometrica, Vol. 48, No. 1, Jan., 1980.

[93] K. Gopalsamy. Stability and Oscillation in Delay Differential Equations of Population Dynamics. Kluwer Academic Publishers, London, 1992.

[94] J. Haigh , Game theory and evolution. Adv. Appl. Probab 7:8-11, 1975.

[95] W. D. Hamilton, The evolution of altruistic behavior, *American Naturalist* 97, 354-356, 1963.

[96] W. D. Hamilton , The genetical evolution of social behviour, I and II, *Journal of Theoretical Biology* 7, 1-52, 1964.

[97] W.D. Hamilton, *Extraordinary sex ratios*. Science 156: 477-488, 1967.

[98] S. Haykin, *Cognitive Radio: Brain-Empowered Wireless Communications,* IEEE JSAC, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[99] J. Hofbauer and K. Sigmund. *Evolutionary game dynamics*. American Mathematical Society, Vol 40 No. 4, pp. 479-519, 2003.

[100] J. Hofbauer and K. Sigmund, *Theory of Evolution and Dynamical Systems*. Cambridge: Cambridge University Press, 1988.

[101] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press 1998.

[102] J. Hofbauer and W. H. Sandholm , *Stable Games*, in Proc. of IEEE CDC, 2007.

[103] A.J. Hoffman , J.B. Kruskal, *Integral Boundary Points of Convex Polyhedra*, in Kuhn H.W. ,Tucker, A.W., Linear Inequalities and Related Systems, Annals of Mathematics Studies, 38, Princeton: Princeton University Press, pp. 223-246, 1956.

[104] A. I. Houston and J. M. McNamara, *Evolutionarily stable strategies in the repeated hawk-dove game*, Behavioral Ecology, 1991.

[105] A. I. Houston and J. M. McNamara, *Fighting for food: a dynamic version of the Hawk-Dove game*, Evolutionary Ecology, Vol 2, pp. 51-64, 1988.

[106] J. Hromkovic, R. Klasing, A. Pelc, P. Ruzicka and W. Unger, *Dissemination of Information in Communication Networks: Broadcasting, Gossiping, Leader Election, and Fault-Tolerance,* Springer Monograph, 2005.

[107] J. Hu, M. P. Wellman, Nash Q-Learning for General-Sum Stochastic Games, Journal of Machine Learning Research 4 , 1039-1069(2003).

[108] M. Huang, P. E. Caines, R. P. Malhamé, A Locality Generalization of the NCE (Mean Field) Principle: Agent Specific Cost Interactions, CDC 2008.

[109] H. Inaltekin and S. B. Wicker, *A one-shot random access game for wireless networks*, Symposium on Information Theory in Wireless com, 2005.

[110] Inaltekin H. and Wicker S.B., *A One-shot Random Access Game for Wireless Networks- Behavior of Nodes at Nash Equilibria*, IEEE/ACM Trans. on Networking, to appear December 2008.

[111] H. Inaltekin and S. Wicker, *A One-shot Random Access Game*, Inter. Conference on Wireless Networks, Comm. and Mobile Computing, 2005.

[112] L. A Imhof, *Long-run behavior of the stochastic replicator dynamics*, Annals of Appl Probability, Vol 15 No 1B, 1019-1045, 2005.

[113] Van Jacobson *Congestion Avoidance and Contrnol*, SIGCOMM '88, Stanford, CA, Aug. 1988.

[114] B. Jovanovic, R. W. Rosenthal, , *Anonymous Sequential Games*, Working Papers 86-12, C.V. Starr Center for Applied Economics, New York University, 1986.

[115] Ji H. and Huang C.,*Non-cooperative uplink power control in cellular radio systems*. Wireless Networks, 4(3): 233–240, April, 1998.

[116] Ramesh Johari and David Kim Hong Tan, *End-to-end congestion control for the internet: delays and stability*, IEEE/ACM Trans. Netw.,9,6,818–832,2001.

[117] Shizuo Kakutani, *A generalization of Brouwer's fixed point theorem*. Duke Mathematical Journal 8 (3): 457–459, 1941.

[118] Kalai, E., and M. Smorodinsky, *Other solution to Nash's bargaining problem*, Econometrica 43, 513–518, 1975.

[119] L. C. M. Kallenberg, *Linear Programming and Finite Markovian Control Problems*, Mathematical Centre Tracts 148, Amsterdam, 1983.

[120] H. Kameda, *How harmful the paradox can be in braess/cohen-kelly-jeffries networks*, In IEEE INFOCOM, New York, NY, USA, June 2002.

[121] G. Karumanchi, S. Muralidharan, R. Prakash *Information Dissemination in Partitionable Mobile Ad Hoc Networks*, in proceedings of Symposium on Reliable Distributed Systems, 2006.

[122] T. Kelly, *Scalable TCP: Improving Performance in Highspeed Wide Area Networks*, Computer Comm. Review 32(2), April 2003.

[123] Kim H., Han Y., Koo J.,*Optimal subchannel allocation scheme in multicell OFDMA systems*, in Proc. IEEE VTC, May 2004.

[124] Yannis A. Korilis, Aurel A. Lazar and Ariel Orda, *Achieving network optima using Stackelberg routing strategies*, IEEE/ ACM Transactions on Networking, vol 5,1,161–173, 1997.

[125] Lars P. Koch, *Evolution and Correlated Equilibrium*. European Economic Association Econometric Society Parallel Meetings, 2006.

[126] Thomas G. Kurtz,*Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes*, Journal of Applied Probability, Vol. 7, No. 1, pp. 49-58, Apr., 1970.

[127] T. G. Kurtz, *Limit Theorems for Sequences of Jump Markov Processes Approximating Ordinary Differential Processes*, Journal of Applied Probability, Vol. 8, No. 2,pp. 344-356, Jun., 1971.

[128] Kushner, Harold J., and Paul G. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, New York: Springer-Verlag, 1992.

[129] J. M. Lasry and P. L. Lions, Mean field games, Jpn. J. Math., vol. 2 (1), pp 229-260,2007.

[130] Le Boudec J.Y., McDonald D. , and Mundinger J., *A Generic Mean Field Convergence Result for Systems of Interacting Objects*. In QEST 2007.

[131] G. Li and H. Liu, *Downlink dynamic resource allocation for multi-cell OFDMA system*, in Proc. IEEE VTC, Oct. 2003

[132] J. Lipman, P. Boustead, J. Chicharo, J. Judge *Resource Aware Information Dissemination in Ad Hoc Networks*, in proceedings of International Conference on Networks, 2003.

[133] X. Li, S. Ruin and J. Wei, *Stability and bifurcation in delay-differential equations with two delays*, Journal of math. analysis and applications, vol. 236, no2, pp. 254-280, 1999.

[134] Li,G. and Liu, H., *Downlink dynamic resource allocation for multi-cell OFDMA system*, in Proc. IEEE VTC, Oct., 2003.

[135] L. Lopez, A. Fernandez and V. Cholvi. *A Game Theoretic Analysis of Protocols Based on Fountain Codes*. IEEE ISCC, June 2005

[136] L. Lopez, G. Rey, A. Fernandez and S. Paquelet, *A mathematical model for the TCP tragedy of the commons*, Theoretical Computer Science, 343, pp. 4–26, 2005.

[137] R. Ma, S.Lee, J. Lui, D. Yau, *A Game Theoretic Approach to Provide Incentive and Service Differentiation in P2P Networks*, Sigmetrics/performance 2004.

[138] A. W. Marshall and I. Olkin , *Inequalities: Theory of of Majorization and Its Applications*, Academic Press, 1979.

[139] J. M. McNamara, A. I. Houston, *If animals know their own fighting ability, the evolutionary stable levels of fighting is reduced*, Journal of theoretical Biology, 232(2005), 1-6.

[140] J. M. McNamara, S. Merad and E. J. Collins, *The hawk-dove game as an average cost problem*, Advances of Applied Probability, vol. 23, 1991.

[141] J. McNamara, S. Merad, E. Collins, *The Hawk-Dove Game as an Average-Cost Problem*, Advances in Applied Probability, vol. 23, no. 4, 1991.

[142] I. Menache and E. Altman :*Battery-State Dependent Power Control as a Dynamic Game*, In Proc. of WiOpt'08, Berlin, 31 March–04 April 2008.

[143] D.S. Menasche, D.R. Figueiredob, and E. de Souza e Silva, *An evolutionary game-theoretic approach to congestion control*, Performance Evaluation Vol. 62, 1-4, Pages 295-312,October 2005.

[144] Meshkati F., Poor H., Schwartz S., *Energy-Efficient Resource Allocation in Wireless Networks*, IEEE Signal Processing Magazine, vol. 58, 2007.

[145] F. Meshkati, H. Poor, S. Schwartz, N. Mandayam, *An Energy-Efficient Approach to Power Control and Receiver Design in Wireless Data Networks*, IEEE Trans. on Communications, vol. 52, 2005.

[146] F. Meshkati, H. Poor, S. Schwartz, *Energy-Efficient Resource Allocation in Wireless Networks*, IEEE Signal Processing Mag., vol. 58, 2007.

[147] D.S. Menasche, D.R. Figueiredob, and E. de Souza e Silva, *An evolutionary game-theoretic approach to congestion control*, Performance Evaluation Volume 62, Issues 1-4, Pages 295-312, October 2005.

[148] Miękis J. and Patkowski T., *Population dynamics with a stable efficient equilibrium,* Journal of Theoretical Biology, Vol. 237, Issue 4, 21, Pages 363-368, December 2005.

[149] Monderer D, Shapley LS (1996) *Potential games*. Games Economic Behav 14:124–143.

[150] Morgenstern, Oskar and John von Neumann (1944), *The Theory of Games and Economic Behavior*, Princeton University Press.

[151] Roger B. Myerson. *Refinements of the Nash equilibrium concept*. International Journal of Game Theory, 15:133–154, 1978.

[152] Nachbar, J. *Evolutionary Selection Dynamics in Games: Convergence and Limit Properties*, International Journal of Game Theory 19, 59-89, 1990.

[153] J.F. Nash, *Equilibrium Points in N-Person Games*. Proc. Nat. Acad. Sci. U.S.A. 36: 48–49, 1950.

[154] J. Nash, *Non-Cooperative Games*, The Annals of Mathematics 54(2): 286–295, 1951.

[155] Neel J., Reed J., and MacKenzie A., *Cognitive Radio Network Performance Analysis, Cognitive Radio Technology*, B. Fette, ed., Elsevier, Burlington MA August,2006.

[156] Neyman A, *Correlated equilibrium and potential games*. Int. J. Game Theory 26:223–227, 1997.

[157] Nikaido H. and Isoda K., *Note on non-cooperative convex games*. Pacific Journal of Mathematics, 5:807–815, 1955.

[158] Niyato D., Hossain E., and Fallahi A.,*Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization*. IEEE Transactions on Mobile Computing, 6(2):221-236, 2007.

[159] J. Oechssler and F. Riedel, *On the dynamics fondation of evolutionary stability in continuous models* , J. Econ. Theory 107, 223-252,2002.

[160] Perez-Guirao, M.D. Luebben, R. Kaiser, T. Jobmann, Evolutionary Game Theoretical Approach for IR-UWB Sensor Networks, IEEE ICC 2008.

[161] Ponstein J., *Existence of Equilibrium Points in Non-Product Spaces*, SIAM Journal on Applied Mathematics, Vol. 14, No. 1, pp. 181-190,Jan., 1966.

[162] Robinson, J. *An Iterative Method of Solving a Game*, Annals of Mathematics 54, 296-301, 1951.

[163] J. B. Rosen, *Existence and Uniqueness of Equilibrium Points for Concave N-Person Games*, Econometrica, Vol. 33, Issue 3, pp. 520-534, Jul. 1965.

[164] T. Roughgarden, Stackelberg scheduling strategies, Proc. of the 33th annual ACM symposium on Theory of computing, 2001.

[165] G.-C. Rota , *The Number of Partitions of a Set*, American Mathematical Monthly 71(5): 498–504, 1964.

[166] Y. E. Sagduyu, A. Ephremides, *SINR-Based MAC Games for Selfish and Malicious Users*, in Information Theory and Applications Workshop 2007, San Diego, CA, Jan. 2007.

[167] W.H. Sandholm. *Potential games with continuous player sets*. J. Eco. Theory, 97:81-108, 2001.

[168] W. H. Sandholm, Population Games and Evolutionary Dynamics, to be published by MIT Press, 2009.

[169] L. Samuelson, *Evolutionary Games and Equilibrium Selection*, MIT Press, 1997.

[170] L. Samuelson, *Evolution and Game Theory*, *The journal of Economics Perspectives*, Vol 16, No 2, 2002.

[171] C. Saraydar, N. Mandayam, D. Goodman *Efficient Power Control via Pricing in Wireless Data Networks*, IEEE Trans. on Communication, vol. 50, 2002.

[172] S. Sarkar, E. Altman, R. ElAzouzi, Y. Hayel, *Information Concealing Games*, in Proc. of IEEE INFOCOM 2008.

[173] D. Schmeider , *Equilibrium points of nonatomic games*. Journal of Statistical Physics 7 , pp. 295-300, 1973.

[174] G. Scutari, S. Barbarossa, and D. P. Palomar, *Potential games: A framework for vector power control problems with coupled constraints*, in Proc. IEEE ICASSP, 2006.

[175] R. Selten, *A reexamination of the perfectness concept for equilibrium points in extensive games*. International Journal of Game Theory 4:25–55, 1975.

[176] L. S. Shapley. *Stochastic games*. In Proceedings of the National Academy of Science USA, volume 39, pages 1095-1100, 1953.

[177] L. Shapley ,*Some Topics in Two-Person Games*, In Advances in Game Theory M. Drescher, L.S. Shapley, and A.W. Tucker (Eds.), Princeton: Princeton University Press, 1964.

[178] N. Shimkin , *Stochastic games with average cost constraints*, Annals of the International Society of Dynamic Games, Vol. 1: Advances in Dynamic Games and Applications, Eds. T. Basar and A. Haurie, Birkhauser, 1994.

[179] N. Shimkin, A Survey of Uniqueness Results for Selfish Routing, NetCoop - International Conference on Network Control and Optimization. In: Lect. Notes Comp. Sci. 4465, pp. 33-42, June 2007.

[180] H. Singh, S. Singh, *Energy consumption of TCP Reno, New Reno and SACK in multi-hop wireless networks*, ACM SIGMERTICS, Jun. 2002.

[181] S. Shakkottai, E. Altman and A. Kumar, *The Case for Non-cooperative Multihoming of Users to Access Points in IEEE 802.11 WLANs*, proceedings of IEEE Infocom, 2006.

[182] M. Sion. *On general minimax theorems*, Pac. J. Math. 8, pp. 171–176,1958.

[183] S. Shakkottai, E. Altman, A. Kumar, *The case for non-cooperative multihoming of users to access points in ieee 802.11 wlans*. IEEE Infocom, 2006.

[184] S. Shakkottai, E. Altman and A. Kumar, *Multihoming of Users to Access Points in WLANs: A Population Game Perspective*, IEEE Journal on Selected Areas in Communication, Special Issue on Non-Cooperative Behavior in Networking Vol 25 No 6, 1207-1215, Aug 2007.

[185] A. J. Shaiju and P. Bernhard, *Evolutionarily Robust Stragies: Two Nontrivial Examples and a Theorem*, Proc. of ISDG, 2006.

[186] M. Smith, *Game Theory and the Evolution of Fighting*, In John Maynard Smith, On Evolution (Edinburgh University Press), pp.8-28, 1972.

[187] J. Maynard Smith and GR. Price, *The logic of animal conflict*, Nature 246, 15-18, 1973.

[188] Maynard Smith, J. *Evolution and the Theory of Games*, Cambridge University Press, 1982.

[189] Smith, M. J. *The stability of a dynamic model of traffic assignment: an application of a method of Lyapunov*. Transportation Science, 18: 245-252, 1984.

[190] H. L. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, American Mathematical Society,1995.

[191] E. Souza and D.A. Agarwal. *A High Speed TCP study: Characteristics and deployment issues*, Technical Report LBNL-53215, Lawrence Berkeley National Laboratory, 2002.

[192] S. Sorin :*Asymptotic properties of a non-zero sum game*, International Journal of Game Theory, 15, pp. 101-107, 1986.

[193] E. Solan, *Stochastic Games*, Thesis, Hebrew University of Jerusalem, 1998.

[194] H. V. Stackelberg. *Marktform und Gleichgewicht*. Julius Springer, Vienna, Austria, 1934.

[195] Ekaterina Shevkoplyas, *Shapley Value in cooperative differential games with random duration*, ISDG 2008.

[196] A.S. Sznitman, *Topics in propagation of chaos*, In P.L. Hennequin, editor, Springer Verlag Lecture Notes in Mathematics 1464, Ecole d'Eté de Probabilités de Saint- Flour XI , pages 165-251, 1991.

[197] P. Taylor and L. Jonker, *Evolutionary stable strategies and game dynamics*. Mathematical Biosciences, 16:76-83, 1978.

[198] Tanabe, Y., *The propagation of chaos for interacting individuals in a large population*, Mathematical Social Sciences,51, 2,pp.125-152, 2006.

[199] Kok-Keong Tan, Jian Yu and Xian-Zhi Yuan, *Existence Theorems of Nash Equilibria for Non-Cooperative N-Person Games*, International Journal of Game Theory, 24: 217–222, 1995.

[200] A. Tang, J. Wang, S. Low, and M. Chiang. *Equilibrium of heterogeneous congestion control: Existence and uniqueness*, IEEE/ACM Transactions on Networking, Oct 2007.

[201] Takashi H. and Yuya K.,*Meta-evolutionary game dynamics for mathematical modelling of rules dynamics*, ECAL: 107-117, 2003.

[202] Takashi Ui, *Correlated equilibrium and concave games*, International Journal of Game Theory,37,April 2008.

[203] Taylor Peter D.: *Evolutionarily Stable Strategies with Two Types of Player*, Journal of Applied Probability, Vol. 16, No. 1 , pp. 76-83, Mar., 1979.

[204] Tao Yi and Wang Zuwang. Effect of Time Delay and Evolutionary Stable Strategy. J. theor. Biol, 187, 111-116, 1997.

[205] Tao Y. and Wang Z., . *Effect of time delay and evolutionarily stable strategy*, J. Theor. Biol. 187, 111-116,1997.

[206] Takahashi M.:*Equilibrium points for stochastic non-cooperative $n-$person games*, Journal of Science of Hiroshima University, pp. 95-99, 1964.

[207] H. Tembine , Dynamic Games and Networks Evolution, Technical report, University of Avignon, France, 700 pages, June 25, 2009.

[208] H. Tembine, J. Y. Le Boudec, R. ElAzouzi, E. Altman, *From mean field interaction to evolutionary game dynamics*, in proceedings IEEE/ACM WIOPT, Physcomnet, June 2009.

[209] Tembine H., Altman E., ElAzouzi R., Sandholm W. H., *Evolutionary game dynamics with migration for hybrid power control in wireless communications*, in Proc. of 47th IEEE Conference on Decision and Control (CDC), 2008.

[210] Tembine H., Le Boudec J. Y., ElAzouzi R., Altman E., *Mean field asymptotics of Markov decision evolutionary games*, In Proc. IEEE Gamenets, May 2009.

[211] H. Tembine, E. Altman, R. Elazouzi, Y. Hayel, *Stable Networking Games* , in Proc. 46th Annual Allerton Conference on Communication, Control, and Computing, 2008.

[212] H. Tembine, E. Altman, R. ElAzouzi, Y. Hayel, *Battery State-dependent Access Control in Solar-powered Broadband Wireless Networks*, in Proc. of Second International EuroFGI Conference on Network Control and Optimization (NETCOOP 2008).

[213] H. Tembine, E. Altman, R. El-Azouzi and Y. Hayel: *Evolutionary games with random number of interacting players applied to access control*, in Proc. of WiOpt, 31 March-04 April 2008.

[214] H. Tembine, E. Altman and R. ElAzouzi, *Asymmetric Delay in Evolutionary Games*, International Conference on Performance Evaluation Methodologies and Tools(VALUETOOLS), Oct. 23-25, 2007.

[215] H. Tembine, E. Altman and R. El-Azouzi, *Delayed Evolutionary Game Dynamics applied to the Medium Access Control*, Proc. of Workshop on Socially and Biologically Inspired Wired and Wireless Networks (Bionetworks), in conjunction with the Fourth IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, Oct., 2007.

[216] H. Tembine, E. Altman and R. El-Azouzi and Y. Hayel, *Multiple Access Game in Ad-hoc Network*, Proc. of First International Workshop on Game theory for Communication networks (GameComm), Oct., 2007.

[217] H. Tembine, E. Altman, R. El-Azouzi, Y. Hayel , *Stochastic population games with individual independent states and coupled constraints*, In Proc. of Second International Workshop on Game theory in Communication networks (Gamecomm), Oct. 2008.

[218] H. Tembine, E. Altman , R. El-Azouzi, Y. Hayel, *Correlated evolutionarily stable strategies in random medium access control*, In Proc. of International IEEE Conference on Game Theory for Networks (GameNets), 2009.

[219] H. Tembine, E. Altman, R. El-Azouzi, Y. Hayel, Evolutionary games in wireless networks, to appear in IEEE Transactions on Systems, Man and Cybernetics, 2009.

[220] D. Topkis, *Equilibrium points in nonzero-sum n-person submodular games*. SIAM Journal on Control and Optimization, 17:773-787, 1979.

[221] B. Thomas, *On Evoloutionary Stable Sets*, Journal of Mathematical Biology, vol. 22, pp. 105-115, 1985.

[222] A. Vahdat, D. Becker, *Epidemic Routing for Partially-Connected Ad Hoc Networks*, Technical Report, Duke University, 2000.

[223] E. van Damme , *Stability and perfection of Nash equilibria*, Springer, Berlin, 1991.

[224] E. van Damme *A relationship between perfect equilibria in extensive form games and proper equilibria in normal form games.* International Journal of Game Theory 13:1–13, 1984.

[225] Vieille N. , *Equilibrium in 2-Person Stochastic Games I: A Reduction.* Israel J. Math. 119 55Ű91, 1997.

[226] Vieille N.:*Equilibrium in two-person stochastic games: a reduction* , Israel Journal of Mathematics, 119, pp.93-126.

[227] Vieille N.:*Stochastic games: Recent results.* In Handbook of Game Theory, pages 1833-1850. Elsevier Science, 2002.

[228] Vieille N., Rosenberg D., Solan E.:*Stochastic Games with a Single Controller and Incomplete Information,*SIAM Journal on Control and Optimization, 43, 86-110, 2004.

[229] Vincent T.L, Brown J.S. *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics.* Cambridge, UK: Cambridge Univ. Press. 400 pp.2005.

[230] T. L. Vincent and T. L. S. Vincent, *Evolution and control system design*, IEEE Control Systems Magazine, Vol 20 No. 5, Oct. 2000.

[231] Viossat Y., *The Replicator Dynamics Does not Lead to Correlated Equilibria*, Games and Economic Behavior, 59, 397-407,2007.

[232] B. Wang, K. J. R. Liu, and T. C. Clancy, *Evolutionary Game Framework for Behavior Dynamics in Cooperative Spectrum Sensing*, to appear in Globecom 2008.

[233] Wardrop, J. G., *Some theoretical aspects of road traffic research*, Proceedings, Institute of Civil Engineers, PART II, Vol.1, pp. 325–378, 1952.

[234] J.W. Weibull, *Evolutionary Game Theory*, Cambridge, MA: MIT Press, 1995.

[235] G. Y. Weintraub, L. Benkard, B. Van Roy, *Oblivious Equilibrium: A mean field Approximation for Large-Scale Dynamic Games*, Advances in Neural Information Processing Systems, Vol 18, 2006.

[236] D. J. White, *Markov Decision Processes*, J. Wiley, 1993.

[237] E. Wright, *A Non-linear Difference-Differential Equation*, J. Reine Angew. Math., vol. 494, 1955.

[238] B. Xu, A. Ouksel and O. Wolfson, *Opportunistic Resource Exchange in Inter-vehicle Ad-hoc Networks*, in Mobile Data Management, 2004.

[239] B. Yang, G. Feng, X. Guan, *Noncooperative Random Access Game via Pricing in Ad Hoc Networks*, Proceedings of the 46th IEEE Conference on Decision and Control New Orleans, LA, USA, Dec. 12-14, 2007.

[240] Y. Zheng and Z. Feng, *Evolutionary Game Theoretic Model and End-to-End Flow Control in the Internet*, proceedings of ICII, Beijing, 2001.

[241] X. Zhang, G. Neglia, J. Kurose, D. Towsley *Performance Modeling of Epidemic Routing*, Computer Networks, 2007.

[242] Q. Zhu , H. Tembine , and T. Başar , *A constrained evolutionary Gaussian multiple access channel game*, Proc. International Conference on Game Theory for Networks (GameNets), Istanbul, Turkey, May 13-15, 2009 .

[243] M. Zorzi and R. Rao, *Energy efficiency of TCP in a local wireless environment*, Mobile Networks and Applications, Vol. 6, No. 3,2001.