



New approaches to evolutionary games and decision dynamics

Ilaria Brunetti

► To cite this version:

Ilaria Brunetti. New approaches to evolutionary games and decision dynamics. Other [q-bio.OT]. Université d'Avignon, 2015. English. NNT : 2015AVIG0204 . tel-01305006

HAL Id: tel-01305006

<https://theses.hal.science/tel-01305006>

Submitted on 6 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

Présentée pour obtenir le grade de Docteur en Sciences de l'Université d'Avignon et
des Pays de Vaucluse France

SPÉCIALITÉ : Informatique

École Doctorale 536: *Agrosciences et Sciences*

Laboratoire d'Informatique d'Avignon

Nouvelles Approches aux Jeux Évolutionnaires et Processus de Décision

par

Ilaria Brunetti

Soutenue publiquement le 8 Decembre devant un jury composé de :

Samson Lasaulce	Directeur de recherche CNRS au LSS Paris	Rapporteur
Frank Thuijsman	Professeur à l'Universié de Maastricht	Rapporteur
Tijani Chahed	Professeur à Telecom Sud-Paris	Examineur
Yezekael Hayel	Maitre de conférence à l'Université d'Avignon	Examineur
Eitan Altman	Directeur de recherche à l'INRIA Sophia-Antipolis	Directeur de thèse



Laboratoire LIA, Avignon



INRIA, Sophia Antipolis



THESIS

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

In Computer Science

Doctoral School 536: *Agrosciences et Sciences*

Laboratoire d'Informatique d'Avignon

New Approaches to Evolutionary Games and Decision Dynamics

by

Ilaria Brunetti

Committee

Samson Lasaulce	CNRS Director of Research at LSS Paris	Reviewer
Frank Thuijsman	Associate Professor at Maastricht University	Reviewer
Tijani Chahed	Professor at Telecom Sud-Paris	Examiner
Yezekael Hayel	Associate professor at Avignon University	Examiner
Eitan Altman	Director of research at l'INRIA Sophia-Antipolis	Supervisor



Laboratoire LIA, Avignon



INRIA, Sophia Antipolis

Dedicated to my beloved *nonna Gnagna*,
who used to repeat to me that
“A living donkey is better than a dead scientist”,
but who, if she was still here,
would be deeply proud of me.

Résumé

Mots clés: Théorie des Jeux Évolutionnaires, Processus de Décision Markovien, Systèmes Dynamiques, Contrôle et Optimisation

La théorie des jeux évolutionnaires (EGT) constitue un cadre simple pour étudier le comportement de populations larges dont les membres sont engagés en interactions stratégiques. L'origine de l'EGT est due à Maynard Smith ([Maynard Smith, 1972](#)), qui utilise les concepts de base de la Théorie des Jeux (GT) classique pour étudier l'évolution des espèces en biologie. Tandis que la GT considère des agents rationnels qui choisissent leurs actions pour maximiser leur propre utilité, dans l'EGT, au départ, les joueurs sont censés être les membres d'une espèce, pour lesquels l'hypothèse de rationalité ne peut pas être considérée. Les actions sont interprétées comme des traits héréditaires et l'utilité correspond au *fitness* darwinien (ou succès reproductif). Après la diffusion de l'EGT en biologie, les économistes comprennent que l'approche évolutionnaire peut être très utile pour l'étude des équilibres de la GT et que ses modèles peuvent représenter des problèmes dans autres contextes que la biologie. De plus, les modèles évolutionnaires ont l'avantage de ne pas nécessiter d'hypothèse de rationalité, étant un concept très difficile à définir et caractériser ([Weibull, 1998](#)). Déjà Nash, dans sa thèse, avait suggéré que son concept d'équilibre (l'équilibre de Nash) peut avoir deux interprétations, une rationaliste et une autre qu'il appelle «action de masse» ([Nash, 1950](#)). L'EGT est désormais considérée comme un enrichissement important de la GT et s'applique à un vaste éventail de problèmes dans des contextes différents, comme l'économie, l'informatique, les télécommunications, les sciences sociales. En étant une théorie assez récente, des nombreuses applications restent encore à identifier, ainsi que plusieurs aspects théoriques à approfondir.

Dans les jeux évolutionnaires standards, les individus d'une grande population sont sélectionnés aléatoirement de façon répétitive pour jouer un jeu symétrique entre deux joueurs. Le *fitness* d'un joueur est défini comme une fonction de son action ainsi que de la distribution des actions au sein de la population. Les actions donnant un *fitness* plus élevé, se diffusent dans la population. Le concept d'équilibre est la stratégie évolutionnairement stable (SES ou ESS), introduite par Maynard Smith et Price ([Maynard Smith and Price, 1973](#)). Il s'agit d'une stratégie telle que, si elle est adoptée par l'ensemble de la population, ne peut être envahie par une stratégie différente, et est donc stable par rapport à la déviation d'une (petite) fraction de la population à une stratégie différente (mutante). D'un point de vue biologique, l'ESS peut être vue comme une généralisation du concept darwinien de «survie du plus apte», alors que d'un point de vue des jeux, il s'agit d'un raffinement conceptuel de l'équilibre de Nash.

Tandis que l'ESS est un concept statique qui vise à expliquer les processus de mutation, Taylor et Jonker ([Taylor and Jonker, 1978](#)) définissent des dynamiques qui permettent de comprendre comment une population atteint une situation stable, la dynamique du réplicateur, en mettant l'accent sur le processus de sélection. Ils formalisent ces dynamiques à travers un système d'équations différentielles et montrent le lien entre les points stationnaires de ce système et les équilibres du jeu. Dans les jeux évolutionnaires standards le joueur est l'individu qui choisit ses actions pour maximiser son propre *fitness*. Dans cette thèse nous proposons une nouvelle approche pour la modélisation de l'évolution, où le joueur est formé par un ensemble d'individus. Nous considérons toujours des interactions entre individus mais nous supposons qu'ils maximisent le *fitness* du group auquel ils appartiennent.

Une direction importante pour le développement de l'EGT est celle des jeux stochastiques. La notion de hasard est implicite dans la notion d'ESS, en étant une stratégie stable face à des mutations aléatoires. Même la dynamique du réplicateur peut être vue comme l'approximation déterministe d'un processus stochastique, où le hasard disparaît quand la population est suffisamment large. Dans la deuxième partie de ce manuscrit, nous introduisons une classe particulière de jeux évolutionnaires stochastiques, les Jeux Evolutionnaires Markoviens (MDEG), défini par Altman et Hayel ([Altman and Hayel, 2010](#)). Dans ce contexte, chaque joueur est associé à un état individuel qui évolue dans le temps selon un Processus de Décision Markovien (MDP). Un joueur interagit à répétition avec d'autres joueurs, et il change d'états. L'action choisie détermine son *fitness* et aussi la probabilité de transition à l'état suivant. En considérant des politiques stationnaires, Altman et Hayel montrent que, si la distribution sur les états est stationnaire par rapport à la politique choisie, il est possible de transformer le jeu MDEG dans un jeu évolutionnaire standard et calculer ses ESSs.

Motivés par l'importance de comprendre les dynamiques qui amènent à une situation stable, nous présentons ici une nouvelle approche dynamique des MDEG. À différence de l'approche statique adopté par Altman et Hayel, en ce travail nous considérons les dynamiques des états individuels et couplée avec les politiques et nous les décrivons à travers des équations différentielles interdépendantes.

Dans la troisième partie du manuscrit, nous poursuivons l'étude des jeux stochastiques dynamiques dans un contexte différent, la théorie du contrôle. Nous définissons un système stochastique dynamique contrôlé simultanément par deux joueurs engagés dans un jeu à somme non nulle (et non constante) et nous montrons que le problème stochastique peut être approximé à travers un jeu dynamique déterministe.

Acknowledgements

First of all, I would like to thank my supervisor Prof. *Eitan Altman*, for giving me the possibility to live this experience, for his immense knowledge, for always trusting in my capabilities and for letting me travel around the world to present our results.

Besides my advisor, I warmly thank the rest of my committee members: Prof. *Frank Thuijsman* and Prof. *Samson Lasaulce*, for the enthusiasm and the precious time they dedicated to report my thesis, Prof. *Tijani Chahed*, and Prof. *Yezekael Hayel*.

I'm particularly grateful to Yezekael Hayel also for his constant and careful support during the last two years, without which this thesis would probably not exist.

I would also like to thank Prof. *Rachid El-Azuzi* for his motivation and his stimuli: our "lively" exchanges have been fruitful and important for my work.

I wish to express my most sincere gratitude to Prof. *Vladimir Gaitsgory*, who gave me the opportunity to join an internship at Macquarie University. During this period, he attentively guided and encouraged me, turning the last months of my PhD into a thrilling adventure on the other side of the world...and the other side of Mathematics.

I'm thankful to *Laurie Vermeersch*, whose daily assistance *behind the scenes* of Inria's MAESTRO group is fundamental.

An immense thank, from the bottom of my heart, goes to my *parents* and my *sister*, for their unconditional love and essential support, and for being such an important presence in my life. I thank them and my dear *Spartaco* also for coming here to assist my defense and share this important moment with me.

A big thank goes to my "almost cousin" *Andrea and his family*, for being the first ones welcoming me in Avignon among "*The Italians*". An infinite number of thanks go to *Pascale*, *Pierre* and *Lily*, for taking care of me as if I was a member of their family. Thank to my *aussi* sister *Rachel* for making me feel at home even on the other side of the world.

I also need to thank the Avignon friends, who made my stay here pleasant despite my well known non-affection for the city. I can't be too wordy and I can't name you all, but a special, sincere and deep *grazie* goes to *Alessandra*, *Juan*, *Marta*, *Monica*, *Sofia* and *Ugo*, who shared with me the most of this journey and who, through their love and their support, made the way to this accomplishment easier.

Last but not least, I thank *Angela Barusi*, who believed in me from the very first time we met and gave me the opportunity to live all those experiences that allowed me to express the "foodie" side of me, helping me to go through these three years.

Contents

Résumé

Introduction **xii**

I Deterministic Evolutionary Games and Group of Players **1**

CHAPTER 1—An Introduction to Evolutionary Game Theory

1.1	Origins of Evolutionary Game Theory and its Developments	3
1.2	A Static Approach to EGT	4
1.2.1	Normal Form Games	4
1.2.2	Evolutionarily Stable Strategy (ESS)	5
1.2.3	Population Games and Evolutionarily Stable State	7
1.3	A Dynamic Approach to EGT	7
1.3.1	Replicator Equation	9
1.3.2	Convergence and Stability: the Folk Theorem of EGT	9
1.4	The Hawk-Dove Game	10

CHAPTER 2—Group Players

2.1	Introduction	14
2.2	The General Model	17
2.2.1	The Case of a Finite Population of Individuals	18
2.3	Hawk-Dove Game with Group Players	18
2.3.1	Existence and Uniqueness of the Equilibrium	19

2.3.2	Convergence to the Equilibrium	26
2.4	Conclusion	27
CHAPTER 3—A New Equilibrium Concept: Group Fitness Evolutionarily Stable Strategy		
3.1	Introduction	29
3.2	Group Fitness Evolutionarily Stable Strategy	31
3.2.1	Nash Equilibrium and Group-Players	34
3.2.2	GFESS and Standard ESS	34
3.3	Analysis of N Groups Games with Two Pure Actions	35
3.4	Some Examples	39
3.4.1	Hawk-Dove Game	40
3.4.2	Stag Hunt Game	41
3.4.3	Prisoner's Dilemma	42
3.5	GFESSs in a Multiple Access Control Game	42
3.5.1	A MAC Game	43
3.6	Conclusion	50
II Individual State and Policy Dynamics in Evolutionary Games		51
CHAPTER 4—Markov Decision Evolutionary Games		
4.1	Introduction	53
4.2	Markov Decision Process	56
4.2.1	Occupation measures	57
4.3	Markov Decision Evolutionary Games	58
CHAPTER 5—State Policy Couple Dynamics		
5.1	Introduction	62
5.2	General Model	64
5.2.1	The Individual State Dynamical Model	64
5.2.2	Population of Players	65
5.2.3	Policies Dynamics	66
5.2.4	Two Time-Scales Behavior	67
5.3	Complete Characterization of the Game with Two States and Two Actions	71
5.3.1	Individual State and Its Dynamics	71

5.3.2	Individual Fitness	72
5.3.3	Equilibrium Profile	73
5.3.4	Policy Based Replicator Dynamics	74
5.3.5	State-Policy Coupled Dynamics	75
5.3.6	Singular Perturbations Approximation Method	76
5.3.7	Matrix Game Approximation Technique	81
5.3.8	Relation Between the Equilibria	83
5.3.9	Applications in Network Systems	86
5.4	Conclusion	89
CHAPTER 6—A Dynamic Approach for the Study of a Hawk-Dove Problem		
6.1	Introduction	91
6.2	The Model	92
6.2.1	Deriving the Expected Fitness and the Equilibria	94
6.3	The Dynamic Hawk-Dove Game and Group Players	97
6.3.1	Group Fitness and Equilibria	99
6.4	Conclusion	102
III Stochastic Hybrid Dynamics		105
CHAPTER 7—Hybrid Stochastic Systems		
7.1	Introduction	107
7.2	Hybrid Game	108
7.3	Averaged Dynamic Game	111
7.4	Construction of ANE Policies - Main Results	112
7.4.1	Proofs of Propositions 15 and 16	113
7.5	Conclusion	122
Conclusions		125
CHAPTER 8—Conclusions and Future Perspectives		
8.1	Summary of Contributions	125
8.2	Perspectives	127

List of Publications

List of figures 131

Bibliography 135

Introduction

“The word “model” sounds more scientific than “fable” or “fairy tale” although I do not see much difference between them. [...] The author of a fable draws a parallel to a situation in real life. He has some moral he wishes to impart to the reader. The fable is an imaginary situation that is somewhere between fantasy and reality. Any fable can be dismissed as being unrealistic or simplistic, but this is also the fable’s advantage. Being something between fantasy and reality, a fable is free of extraneous details and annoying diversions. In this unencumbered state, we can clearly discern what cannot always be seen in the real world. On our return to reality, we are in possession of some sound advice or a relevant argument that can be used in the real world. We do exactly the same thing in economic theory.”

Ariel Rubinstein, *Dilemmas of an Economic Theorist*, 2006

Evolutionary Game Theory (EGT) constitutes a simple framework to study the behavior of large populations whose individuals are repeatedly engaged in pairwise strategic interactions. The birth of EGT is marked by the pioneering work of Maynard Smith ([Maynard Smith, 1972](#)), who uses classical Game Theory (GT) as a tool to explain and to predict quantitative and qualitative aspects of biological evolution of species. While GT considers rational agents choosing their actions in order to maximize their utility function, originally in EGT the players are supposed to be the members of an animal species, for which rationality can’t be assumed. Actions are interpreted as inheritable traits and the utility corresponds to the Darwinian fitness (or reproductive success).

After the spread of EGT in biology, economists understand that the evolutionary ap-

proach is useful to investigate the foundations of game theoretic solution concept (Friedman, 1991) and it can be adopted in social science contexts, to predict human behavior with no need of rationality assumption, which may be complex to define and to interpret (Weibull, 1998). It's worth to mention that in his thesis, Nash already notes that his solution concept (the Nash equilibrium) could have two interpretations, one rationalistic and one that he called the "mass action interpretation": *"We shall now take the "mass-action" interpretation of equilibrium points. In this interpretation solutions have no great significance. It is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to cumulate empirical information on the relative advantages of the various pure strategies at their disposal."* (Nash, 1950).

EGT is nowadays considered as an important enrichment of GT and it's applied in a wide variety of fields, spanning from economics to computer science. Evolutionary models allow, for example, to explain and predict different aspects of human behavior, as the evolution of language (Nowak et al., 1999), the spread of culture (Enquist and Ghirlanda, 2007) and moral behavior (Harms and Skyrms, 2008), as well as to study telecommunication networks dynamic problems, as the evolution of Internet Transport Protocols (Altman et al., 2009), the formation of wireless networks (Shakkottai et al., 2006), (Altman et al., 2008b), (Altman and Hayel, 2008) (Bonneau et al., 2005) and congestion control problems (Menasche et al., 2005), (Zheng and Feng, 2001b), (Zheng and Feng, 2001a). Being a relatively young mathematical theory, there still remain many possible applications yet to be identified, and many interesting theoretical issues to be addressed or deeper explored.

In standard evolutionary games, individuals in a large population are repeatedly and randomly selected to play a symmetric two-person game. The fitness is defined as a function of both the behavior of the individual as well as of the distribution of behaviors among the whole population. Actions with higher fitness are supposed to spread within the population. The main solution concept, first introduced by Maynard Smith and Price (Maynard Smith and Price, 1973), is the *Evolutionarily Stable Strategy* (ESS), which is such that, if a population adopt it, it is *uninvadable* by any other strategy, which means that it is robust against deviations of a (possibly small) fraction of the population to a different strategy (*mutations*). From a biological point of view it can be seen as a generalization of Darwin's idea of survival of the fittest, while from a game theoretical perspective it constitutes a refinement of the Nash Equilibrium. Maynard Smith equilibrium concept has been enriched with an explicit dynamic foundation by Taylor and Jonker (Taylor and Jonker, 1978). In order to explain how a population reaches a stable situation, they introduce the *replicator dynamics*, which highlights the role of se-

lection. It is formalized by a system of ordinary differential equations and it establishes that an action spreads if its fitness is larger than the averaged fitness in the population. While in standard EGT, the interacting individual is the player, choosing the actions to play in order to maximize its own fitness, in the first part of this dissertation we propose a new approach to model evolution, where the player is supposed to be a whole group. We still consider pairwise interactions among individuals but we assume that they maximize the fitness of the group they belong to, which is thus the actual player of the game.

An interesting direction for future developments of EGT is towards *stochastic games*. The notion of randomness is somehow implicitly carried in the concept of evolutionary stability, as the mutations in the population are random events. The replicator equation can also be derived as the deterministic approximation of a stochastic process, where all randomness is averaged away when the population size is sufficiently large. Sandholm ([Sandholm, 2010](#)) defines the notion of *revision protocol*, which specifies the general rule followed by players updating their actions in time, generating a continuous time Markov process over the finite set of actions' distributions (states of the system). He rigorously proves that, when the population is sufficiently large, the stochastic process converges in a finite-horizon to the deterministic *mean dynamics*, defined as the expected motion. We present, in the second part of the manuscript, a particular class of stochastic evolutionary games, that of *Markov Decision Evolutionary Games* (MDEG), introduced by Altman and Hayel ([Altman and Hayel, 2010](#)). In this framework, each player is associated with an individual state which evolves in time according to a Markov Decision Process (MDP). During his finite life time a player repeatedly meets other users in random pairwise interactions and it may move among different states; the actions played by an individual determines his immediate fitness and the transition probabilities to the next state. Restricting (without loss of generality) to stationary policies, Altman and Hayel prove that, if the distribution over the individual states is assumed to be stationary with respect to the currently used policy, it is possible to transform the MDEG into a standard evolutionary game and to compute its ESSs. Motivated by the importance of understanding the dynamics leading to a stable situation, we present in this dissertation, our new dynamical approach to MDEG. In contrast with the static approach adopted by Altman and Hayel, we study here the local dynamics of individual states and the dynamics intrinsically related to the distribution of policies in the population, describing them by interdependent differential equations.

In the third part of the manuscript we pursue the study of stochastic dynamics in a different context, that of control theory. We define a hybrid stochastic dynamical system jointly controlled by two players involved in a non-zero sum game and we prove that

the problem can be approximated by an averaged deterministic differential game.

The manuscript is organized as follows.

Part I: Deterministic Evolutionary Games and Groups of Players

In Part I, we first briefly introduce Evolutionary Game Theory (EGT), pointing out the main notions that will be used and developed in this work. After a hint of the history of EGT, we provide the main equivalent definitions of the ESS. We then illustrate the dynamic aspect of EGT through the replicator dynamics and the relation between the rest points of the replicator equation and the equilibria of the game. We then describe one of the most studied examples in evolutionary games, that of the Hawk-Dove.

In Chapter 2 we present our new approach to evolutionary games, in which the concept of the player as a single individual is replaced by that of a player as a whole group. Even if we still consider pairwise interactions among individuals, we suppose that individuals maximize the fitness of the group they belong to. In order to provide simple but meaningful results, we analyze the Hawk-Dove game in this framework, considering the case of an infinite and that of a finite population of individuals. We obtain, in both cases, a concave game. This allows us to prove the existence and the uniqueness of a symmetric Nash equilibrium through Rosen's results (Rosen, 1965). We explicitly compute the equilibria as a function of the number of groups. We then define a gradient based dynamics as a counterpart of the replicator equation. We obtain that the fact of teaming together makes individuals less aggressive at equilibrium. We deepen the study of group-players in Chapter 3, where we consider groups with different relative sizes and we introduce a new concept of equilibrium, the **Group Fitness Evolutionarily Stable Strategy (GFESS)**. The stability required by the GFESS is related to a notion of deviation within a group, and thus of a fraction of the player (instead of a fraction of the whole population). We first define the GFESS in a general case and we then characterize it for two-actions games. We illustrate our results through three classical examples, the Hawk-Dove game, the Stag Hunt game and the Prisoner's Dilemma. We then generalize the definition of group fitness in order to study an application in multiple access control in slotted Aloha.

Part II: Individual State and policy Dynamics in Markov Decision Evolutionary Games

We begin the second part of the dissertation presenting MDEG as defined by Altman and Hayel (Altman and Hayel, 2010), combining MDP and EGT. In Chapter 5 we present our new dynamical approach to MDEG, in which players are associated with

an individual state evolving according to a continuous time MDP. We first introduce the concept of interdependent dynamics of states and policies in a general case, and we then define the State Policy coupled Dynamics (SPcD) system in a particularly simple scenario, with two states and two actions. We establish the relation between the equilibria of the defined system and the equilibria of the game, and we then find the closed-form solutions of the system. The system is solved by assuming that the processes of states and policies move with different velocities, which allows us to apply two different approximation techniques: the singular perturbation method, and a matrix approximation technique. In Chapter 6 we define a dynamic model for the Hawk-Dove game in a MDEG type of framework, in order to study the impact of the aggressive behavior of adults on the evolution of young individuals. As in MDEG models, players are associated with a MDP, but transitions probabilities here do not depend on the player's action but on the action of its opponent. By considering the stationary distribution over the states, we transform the game into an equivalent standard evolutionary game, and we compute the equilibria. We then combine the notion of group-players presented in Chapter 2 and MDEG, studying the dynamic Hawk-Dove game for the case of group players.

Part III: Stochastic Hybrid Dynamics

In the last part of the manuscript, we extend the theory of control for an hybrid stochastic dynamical system to the case of two players non-zero sum games. The system evolves in continuous time and it is subjected to abrupt changes of the parameters, determined by two (discrete time) Markov decision processes, each of which is controlled by a player that aims at minimizing its objective function. As we did in Chapter 5, we assume a two time scales behavior of the system: the lengths of the time intervals between the "jumps" of the parameters are assumed to be small, which means that parameters evolve faster than the state of the system. This allows us to approximate the hybrid game with a deterministic averaged dynamic game. We prove that an asymptotic Nash equilibrium of such hybrid game can be constructed on the basis of a Nash equilibrium of a deterministic averaged dynamic game. We conclude the dissertation with a last chapter resuming our main contributions and providing the possible future developments of our work.

"Maybe needless to say, the discussion is limited by idiosyncratic limitations to my memory, knowledge and understanding. I apologize for omissions and misrepresentations.", (Weibull, 1998).

Notations

Lowercase letters are mostly used for real numbers, vectors of real numbers and functions, while capital letters indicate matrices and sets. Bold lowercase letters are usually used for vectors, non-bold one for real numbers. We will use the term "action" and "policy", avoiding the term "strategy" which may be misleading. In fact, in evolutionary game theory literature it is often used to refer to actions (pure or mixed), while in Stochastic Games the same term refers to policy.

$\dot{x}(t)$ indicates the derivative $dx(t)/dt$;

\mathcal{A} : set of pure actions, with $|\mathcal{A}| = M$;

$\Delta(\mathcal{A}) = \{\mathbf{p} \in \mathbb{R}^K | \sum_{i=1}^M p_i = 1\}$: set of mixed actions;

\mathcal{S} : set of individual states, with $|\mathcal{S}| = K$;

\mathcal{U} : set of general policies;

\mathcal{U}_M : set of Markov policies;

\mathcal{U}_S : set of stationary policies;

\mathcal{U}_D : set of deterministic (stationary) policies;

$F(\cdot, \cdot)$: fitness function of an individual;

$\bar{F}(\cdot)$: average fitness in the population;

$\Gamma_i(\cdot, \cdot)$: fitness function of a group player G_i ;

α_i : normalized size of a group i ;

\mathcal{Q} : set of transition probabilities;

$\mathcal{Q}_s(s', a)$ (homogeneous time) transition probability from state $s' \in \mathcal{S}$ to state $s \in \mathcal{S}$ given action $a \in \mathcal{A}$;

\mathcal{R} set of transition rates;

$\mathcal{R}_s(s', a)$ transition rate from state $s' \in \mathcal{S}$ to state $s \in \mathcal{S}$ given action $a \in \mathcal{A}$;

$T_s(u)$: average time that an individual playing deterministic policy $u \in \mathcal{U}_D$ spends in state $s \in \mathcal{S}$;

\mathbf{H} : set of all observable state-action histories $h^t = \{s_l, a_l, l = 0, \dots, t\}$, $t = 0, 1, 2, \dots$;

$W = \bigcup_{u_s \in \mathcal{U}_S} \omega(u_s)$: set of all vectors of steady state probabilities of state action pairs.

ANE: Asymptotic Nash Equilibrium

CT-MDP: Continuous Time Markov Decision Process

EGT: Evolutionary Game Theory

ESS: Evolutionarily Stable Strategy

GFESS: Group Fitness Evolutionarily Stable Strategy

GT: Game Theory

MAC: Multiple Access Control

MDP: Markov Decision Process

NE: Nash Equilibrium

SPcD: State Policy couple Dynamics

Part I

Deterministic Evolutionary Games and Group of Players

Chapter 1

An Introduction to Evolutionary Game Theory

"The relation between species, or among the whole assemblage of an ecology, may be immensely complex; and at Dr. Cavalli's invitation I propose to suggest that one way of making this intricate system intelligible to the human mind is by the analogy of games of skill, or to speak somewhat more pretentiously, of the Theory of Games."

R. Fisher, Polimorphism and Natural Selection, 1958

Summary

This chapter gives a brief introduction to evolutionary game theory, pointing out the main notions that will be used and developed in this work. After a hint of the history of EGT, we provide the main equivalent definitions of the equilibrium concept, the ESS. We then study the dynamic aspect of EGT through the replicator dynamics and we establish the relation between the rest points of the replicator equation and the equilibria of the game. We conclude the chapter presenting one of the most studied examples in evolutionary games, that of the Hawk-Dove.

1.1 Origins of Evolutionary Game Theory and its Developments

Evolutionary Game Theory (EGT) has originally developed to formally describe and predict quantitative and qualitative aspects of biological evolution by using the mathematical theory of games.

Before the birth of Game Theory (GT), Fisher ([Fisher, 1930](#)) analyzes the dynamics of the sex ratio in a species as a competition between individuals. In order to explain the stable sex ratio observed in most of mammal species, he supposes that each individual maximizes its own fitness, defined as the expected number of *grandchildren*, depending on the relative frequency of males and females in the population. Even if the formalism he adopts is not the same, Fisher's approach is conceptually very close to game theory. The first explicit attempt to apply GT formalism in evolutionary biology is by Lewontin ([Lewontin, 1961](#)), who describes the evolution of the genetic mechanism as a game played between nature and a species, where the latter seek actions that minimize the probability of extinction. The equilibrium is represented by a *maxmin* action, the species doing the best against the worst of the nature; similar ideas have been developed a few years later by Slobodkin and Rapoport ([Slobodkin and Rapoport, 1974](#)).

Hamilton ([Hamilton, 1967](#)) uses game theory terminology to study the sex ratio of species, in situations where certain underlying assumptions of Fisher's argument do not hold. He models the choice of the sexes of offspring of an individual as the choice of an action in a game and defines the concept of *unbeatable* strategy, which is very close to the notion of evolutionarily stable strategy, the key equilibrium concept introduced by Maynard Smith and Price, ([Maynard Smith and Price, 1973](#)) which marks the official birth of EGT. The authors study animals' conflicts as games, where actions are behavioral phenotypes and the payoff represents the fitness, i.e. the number of offspring of an individual. With his treatise *Evolution and the Theory of Games* ([Maynard Smith, 1982](#)), Maynard Smith brings EGT into widespread circulation.

While Maynard Smith's notion of ESS captures the *mutation mechanism* through a static definition, Taylor and Jonker ([Taylor and Jonker, 1978](#)) focus on the *selection mechanism* and model the dynamic process by which the distribution of actions in a population evolves to a stable situation as a system of ordinary equations; Schuster and Sigmund ([Schuster and Sigmund, 1983](#)) call this model the *replicator dynamics*.

Since the late 1980s, there is a growing interest in EGT by economists, who see the value of the evolutionary approach to game theory in contexts other than biology, both as a method of providing foundations for the equilibrium concepts of traditional game theory, and as a tool for selecting among equilibria. Furthermore the evolutionary approach do not need the rationality assumption required by standard games. EGT mod-

els start to be applied to study the behavior of populations of active decision makers (Friedman, 1991), (Mailath, 1992), (Nachbar, 1990), (Selten, 1991). Even if the majority of work in EGT has been undertaken by biologists and economists, closely related models have been applied in a wide range of disciplines, including sociology (Zhang, 2004), (Bisin and Verdier, 2001), (Kuran and Sandholm, 2008), (Sandholm, 2005), computer science (Shakkottai et al., 2006), (Zheng and Feng, 2001a), (Sandholm, 2001), (Sandholm, 2005), and transportation science (Sandholm, 2001).

1.2 A Static Approach to EGT

1.2.1 Normal Form Games

We briefly introduce here **two players normal form symmetric games**, which will be considered to model pairwise interactions involved in EGT models. Suppose that each of the two players disposes of a finite set of pure actions, \mathcal{A} , with $|\mathcal{A}| = M$. Let $\Delta(\mathcal{A}) = \{\mathbf{p} \in \mathbb{R}_+^M \mid \sum_{i \in \mathcal{A}} p_i = 1\}$ be the set of mixed actions, that are probability measures over the action space. Note that a pure action $i \in \mathcal{A}$ can be represented through the unit vector $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in \Delta(\mathcal{A})$ with all elements equals to zero except for the element in position i , which equals 1. The payoff can be described by a matrix $A = (a_{ij}) \in \mathbb{R}^M \times \mathbb{R}^M$, where the entry a_{ij} corresponds to the payoff that a player gets using pure action $i \in \mathcal{A}$ in an interaction with a player using $j \in \mathcal{A}$. If mixed actions are considered, the expected payoff of an individual playing \mathbf{p} against an opponent using \mathbf{q} , with $\mathbf{p}, \mathbf{q} \in \Delta(\mathcal{A})$, is given by:

$$\mathbf{p}^T A \mathbf{q} = \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{A}} p_i a_{ij} q_j \quad (1.1)$$

An action $\mathbf{q}^* \in \Delta(\mathcal{A})$ is a (symmetric) **Nash equilibrium** of the two players-symmetric normal form game if:

$$\mathbf{q}^{*T} A \mathbf{q}^* \geq \mathbf{p}^T A \mathbf{q}^* \quad \forall \mathbf{p} \in \Delta(\mathcal{A}). \quad (1.2)$$

If (1.2) holds with strict inequality, \mathbf{q}^* is a **strict Nash equilibrium**.

1.2.2 Evolutionarily Stable Strategy (ESS)

Consider a large population of players, where individuals are repeatedly matched at random in a two-players normal form game, as described in the previous section.

Maynard Smith's **ESS** is defined as an action that, if adopted by the whole population, is robust against invasions by a small group of individuals playing a different (*mutant*) action. Suppose that the whole population is programmed to play an (*incumbent*) action $\mathbf{q} \in \Delta(\mathcal{A})$, and that a fraction ϵ of *mutants* deviate to an action $\mathbf{p} \in \Delta(\mathcal{A})$.

Definition 1. The mixed action \mathbf{q} is an ESS if $\forall \mathbf{p} \neq \mathbf{q}$, there exists some $\epsilon_p > 0$ such that:

$$\forall \epsilon \in (0, \epsilon_p) \quad \mathbf{q}^T A(\epsilon \mathbf{p} + (1 - \epsilon) \mathbf{q}) > \mathbf{p}^T A(\epsilon \mathbf{p} + (1 - \epsilon) \mathbf{q}). \quad (1.3)$$

Equation 1.3 requires that, if the size of the invading group is sufficiently small, the incumbent's expected payoff from a random match in the post-entry population exceeds that of any mutant.

The stable population is said to be *monomorphic* if all individuals are assumed to adopt the same action, *polymorphic* if individuals are allowed to take different actions (Maynard Smith, 1982).

It can be easily proved (see e.g. (Weibull, 1995)) that the following conditions are equivalent to Definition 1.

Proposition 1. $\mathbf{q} \in \Delta(\mathcal{A})$ is an ESS if and only if it satisfy:

- *Nash Condition:*

$$\mathbf{q}^T A \mathbf{q} \geq \mathbf{p}^T A \mathbf{q} \quad \forall \mathbf{p} \in \Delta(\mathcal{A}), \quad (1.4)$$

- *Stability Condition:*

$$\mathbf{q}^T A \mathbf{q} = \mathbf{p}^T A \mathbf{q} \Rightarrow \mathbf{q}^T A \mathbf{p} \geq \mathbf{p}^T A \mathbf{p} \quad \forall \mathbf{p} \in \Delta(\mathcal{A}), \mathbf{p} \neq \mathbf{q}. \quad (1.5)$$

The inequality 1.4 corresponds to the definition of the Nash equilibrium, while the second condition requires that, if the mutant action \mathbf{p} is an alternative best reply to the incumbent action \mathbf{q} , then the payoff of the incumbent action against the mutant one is higher than the one of the mutant action against itself. This definition can be interpreted from a biologic perspective: \mathbf{q} is an ESS if an arbitrary rare mutant \mathbf{p} does no better than \mathbf{q} in its most frequent contests against \mathbf{q} . If it does as well in these, then it does worse than \mathbf{q} in its rare contests against another mutant. It immediately follows from (1.4) that a symmetric strict Nash equilibrium is an ESS (while the converse is not true).

Another equivalent definition of the ESS which is often used, is due to Hofbauer and Sigmund.

Theorem 1. [Theorem 6.4.1 (Hofbauer and Sigmund, 1998)] The action $\mathbf{q} \in \Delta(\mathcal{A})$ is an ESS if and only if

$$\mathbf{q}^T \mathbf{A} \mathbf{p} > \mathbf{p}^T \mathbf{A} \mathbf{p}$$

for all $\mathbf{p} \neq \mathbf{q}$ in some neighbourhood of \mathbf{q} in $\Delta(\mathcal{A})$.

Other equivalent definitions of ESS can be found in the literature. The evolutionary stability can be characterized, for example, by introducing the concept of *invasion barrier*, which is useful in the case of a finite population (Weibull, 1995). An important distinction to be made is that among ESS and evolutionarily stable *state* (see the next subsection for the definition of the latter): an analysis of the relation between these two notions can be found in (Thomas, 1984). Some weakening criteria for evolutionary stability are the concept of *neutral* stability (Weibull, 1995) and that of *local* stability (Pohley and Thomas, 1983), while an important refinement of the ESS is the *regular* ESS (Taylor and Jonker, 1978). Maynard Smith and Parker (Maynard Smith et al., 1976) extend the study of the ESS to the case of asymmetric games and Selten (Selten, 1980) proves that, for these games, no mixed action can be evolutionarily stable. Taylor (Taylor, 1979) and Cressman (Cressman, 1992) study the evolutionary stability in the case of two types of players (see Chapter 2). In Chapter 3 we briefly present the different existing definitions of *N-species* ESS, first introduced by Garay and Varga (Garay and Varga, 2000), and we define our new notion of Group Fitness Evolutionary Stable Strategy. Cressman (Cressman, 1992) devotes a text to the development of *static* ESS conditions in diverse theoretical models of evolutionary biology, providing a wide overview of the different definitions of the evolutionary stability notion. Evolutionary stability can also be defined adopting a *dynamical* point of view: see e.g. (Hofbauer and Sigmund, 2003), where the authors state a necessary and sufficient condition for an action to be evolutionarily stable, which is related to the replicator equation. Also Pholey and Thomas' local stability has been defined in order to find a static concept in better agreement with the dynamical aspects of the process of natural selection. Cressman et al. (Cressman et al., 2001) focused on a dynamical approach to the ESS in the case of *N* species. An *application driven approach* leads to a setwise generalization of the ESS, the *Evolutionarily Stable Set* (ES): see (Thomas, 1985), (Cressman, 1992), (Weibull, 1995).

1.2.3 Population Games and Evolutionarily Stable State

Population games study strategic interactions within a large polymorphic populations of individuals playing **pure** actions (Sandholm, 2009). In this scenario, the notion of evolutionary stability is associated with the *state* of the population instead that to a mixed action (see e.g. (Taylor and Jonker, 1978), (Hofbauer and Sigmund, 1998) and (Sandholm, 2009)). More precisely, a **population state** (also called *profile*), is defined by a vector $\mathbf{x} = (x_1, \dots, x_M)$, with $\sum_{i=1}^M x_i = 1$, where x_i is the proportion of individuals in the population playing pure action $i \in \mathcal{A}$. Note that $\mathbf{x} \in \Delta(\mathcal{A})$, so it is formally equivalent to a mixed action in $\Delta(\mathcal{A})$. A population game is identified by the continuous vector valued **payoff function** $F : \Delta(\mathcal{A}) \rightarrow \mathbb{R}^M$, depending on the state of the population. Let $F_i(\mathbf{x})$ denote the payoff of pure action i in a population in state \mathbf{x} . If mixed actions are considered, the payoff of mixed action $\mathbf{q} \in \Delta(\mathcal{A})$, denoted by $F(\mathbf{q}, \mathbf{x})$, can be thought as the average fitness of a group of individuals such that a proportion q_i of the group uses pure action i , against a population in state \mathbf{x} , i.e:

$$F(\mathbf{q}, \mathbf{x}) = \sum_{i=1}^M q_i F_i(\mathbf{x}).$$

Definition 2. A population state \mathbf{q} is a (symmetric) Nash equilibrium profile if:

$$F(\mathbf{q}, \mathbf{q}) \geq F(\mathbf{x}, \mathbf{q}),$$

for all population states $\mathbf{x} \in \Delta(\mathcal{A})$.

Definition 3. A population state \mathbf{q} is evolutionarily stable if:

$$F(\mathbf{q}, \mathbf{x}) > F(\mathbf{x}, \mathbf{x}),$$

for all $\mathbf{x} \neq \mathbf{q}$ in a neighborhood of \mathbf{q} .

Remark 1. If the payoff function F is linear in the population state, then evolutionarily stable state and ESS coincide.

1.3 A Dynamic Approach to EGT

The notion of dynamics has been introduced into evolutionary games by Taylor and Jonker (Taylor and Jonker, 1978), in order to provide a dynamic foundation of Maynard Smith's static concept of evolutionary stability. Through a system of differential equations, the authors describe the evolution of the distribution of actions within a pop-

ulation, where the share of individuals playing a certain action is supposed to change according to the actions' success (represented by the fitness function). They then show the relation between the stable points of their dynamics and the ESSs of the game. The model has been named the *replicator dynamics* by Schuster and Sigmund (Schuster and Sigmund, 1983).

The need for a dynamical approach to games is already felt at the earliest stages of classical game theory, when Neumann and Morgenstern, in the introduction of their treatise (Neumann and Morgenstern, 1947), suggest the idea of completing their static solution concept with some notion of dynamics. The first dynamic models are due to Brown and Von Neumann, who defined the *Brown-Von Neumann-Nash dynamics*, as a tool for computing the equilibria in zero-sum games (Brown and Neumann, 1950). In the last decades, a large number of dynamics has been introduced in GT framework: see e.g. *gradient-based dynamics* (Rosen, 1965), *fictitious play* (Gilboa and Matsui, 1991), *projection dynamics*, *best response dynamics* (Fudenberg and Tirole, 1991), *Boltzman dynamics* and *logit dynamics* (Fudenberg and Levine, 1998).

When a dynamic approach is adopted, a player is supposed to occasionally reconsider its choice of action, adjusting its action in response to several information, like its own current payoff, the average current payoff, the historical actions of the others, etc. In (Sandholm, 2010), the author introduces, in population games framework, the notion of *revision protocols*, which specify the general rule followed by players updating their actions in time (see e.g. *imitative protocol*, *natural selection protocol*, *evaluative protocol*). A revision protocol generates a continuous time Markov process over the finite set of states of the system. For large populations of players and finite time games, this stochastic process can be approximated by its expected motion, given by the deterministic *mean dynamics*. Then, each revision protocol can be viewed as defining a map, called *deterministic evolutionary dynamic*, from population games to mean dynamics. On the basis of the revision protocols which induce them, it is possible to identify different families of deterministic evolutionary dynamics. Sandholm defines four main classes of deterministic dynamics for population games: *imitation* (which includes the replicator dynamics), *excess payoff*, *pairwise comparison* and *perturbed pairwise comparison* dynamics.

In the following subsection, we give the formal definition of the standard replicator dynamics in EGT. In Section 2.3.2 we use Rosen's gradient-based dynamics in our new framework of group-players evolutionary games. In Chapter 5 we study the replicator dynamics of policies coupled with the dynamics of individual states in a particular Markov decision evolutionary game (see Chapter 4 for an introduction to MDEG). In Chapter 7 we deal with a different kind of dynamical system, a *hybrid stochastic system*

for a two players non-zero sum game.

1.3.1 Replicator Equation

While the ESS notion is related to the *mutation mechanism* in a population, the replicator dynamics focuses on the *selection mechanism*, that favors some behavioral phenotypes over others. Consider a large population of haploid individuals programmed to play the same action during their entire lifetime and suppose that offspring inherit the parent's action. In the standard setup for replicator dynamics, only pure actions are allowed and thus a vector $\mathbf{q} \in \Delta(\mathcal{A})$ is interpreted as a population state and not as a mixed action. This means that q_i represents the share of the population adopting pure action i (instead of representing the probability that an individual plays i); in our mathematical treatment, we do not distinguish between the two interpretations.

The replicator dynamics consists in a system of differential equations describing how the frequencies of pure actions evolve in time depending on their success. The evolution of q_i , with $i \in \mathcal{A}$, is expressed as:

$$\dot{q}_i(t) = q_i(t)(F_i(\mathbf{q}(t)) - \bar{F}(\mathbf{q}(t))), \quad (1.6)$$

where $F_i(\mathbf{q})$ denotes the immediate fitness of an individual playing pure action i in a population whose state is \mathbf{q} and $\bar{F}(\mathbf{q}) = \sum_{i=1}^M p_i F_i(\mathbf{q})$ is the average immediate fitness in the population (in state \mathbf{q}). The **growth rate** $\dot{q}^i(t)/q_i(t)$ of the fraction of the population using action i is thus equal to the difference between the immediate fitness of that action and the current average fitness in the population.

In a two-action game, with $\mathcal{A} = \{1, 2\}$, if q indicates the share of the population playing action 1, we have that $\bar{F}(q) = qF_1(q) + (1 - q)F_2(q)$, and thus, by substituting this into (1.6) we obtain:

$$\dot{q}(t) = q(t)(1 - q(t))(F_1(q(t)) - F_2(q(t))). \quad (1.7)$$

1.3.2 Convergence and Stability: the Folk Theorem of EGT

It is easy to see that any Nash equilibrium is a rest point of the replicator dynamics, that is an equilibrium of the ODE (1.6). In fact, if an action in $\Delta(\mathcal{A})$ is a NE, all pure strategies in its support earn the same maximal payoff against that strategy, and thus they all earn the population average payoff. This implies that the right side of (1.6) is zero for that action.

The **folk theorem of evolutionary game theory** (Cressman, 2003), (Hofbauer and Sigmund, 2003) establishes the relation between the rest points of the replicator dynamics, and the Nash equilibria of the symmetric game. Before stating the folk theorem, we define the following notions:

- an orbit is **interior** if it is such that $x(t) \in \text{int}\Delta := \{x \in \Delta \mid x_i > 0, \forall i = 1, \dots, M\}$, $\forall t \geq 0$;
- a rest point x^* is **Lyapunov stable** if, for every neighborhood U_{x^*} of x^* there exists a neighborhood V_{x^*} of x^* such that $x(0) \in V_{x^*}$ implies $x(t) \in U_{x^*}$, $\forall t \geq 0$;
- a rest point x^* is **attracting** if it has a neighborhood U_{x^*} such that $x(t) \rightarrow x^*$ for $t \rightarrow \infty$ holds for $\forall x \in U_{x^*}$;
- a rest point x^* is **asymptotically stable** (or an attractor) if it is both stable and attracting.

Theorem 2. [Theorem 2.5.3 (Cressman, 2003)]

- i. any strict Nash equilibrium is asymptotically stable;
- ii. if a rest point is the limit of an interior trajectory, then it is a Nash equilibrium;
- iii. if a rest point is Lyapunov stable, then it is a Nash equilibrium,

Any ESS is asymptotically stable (Hofbauer et al., 1979), but the converse does not hold in general; only in the special case of **two-actions games**, we have that dynamic stability is equivalent to evolutionary stability. More precisely, the following result holds.

Theorem 3. [Theorem 2.5.4 (Cressman, 2003)] Every interior trajectory that is not initially at rest, converges to an ESS. Furthermore, for the replicator dynamics, the following three statements are equivalent:

- i. $q \in \Delta(\mathcal{A})$ is stable.
- ii. q is asymptotically stable.
- iii. q is an ESS.

1.4 The Hawk-Dove Game

One of the most studied examples of evolutionary games, is the Hawk-Dove game, first introduced by Maynard Smith to study the level of aggressiveness in a population of animals competing for a natural resource (Maynard Smith, 1972). It has been later used

in a wide variety of fields, spanning from biology to economics, where the notion of aggressive behavior can be translated into a short-term oriented maximization of “selfish” individuals (see e.g. (Tomassini et al., 2010), (Hanauske et al., 2010)). In engineering, for example, it is used in (Altman et al., 2009) to study competition of congestion control algorithms in communication networks, while in (Altman et al., 2008a) it serves to study the interactions between mobile phones that can choose which power to use when transmitting packets.

The standard model has been reformulated in many different ways and a large number of revisited versions of the Hawk-Dove game can be found in the literature. Houston and McNamara (Houston and McNamara, 1988), for example, study a *repeated* version of the Hawk-Dove model, including a state variable representing animal’s level of energy resources and they define an ESS which depends on that variable. Cressman (Cressman, 1992) defines a *density dependent* Hawk-Dove game, where he modifies the payoff matrix including in the individual’s fitness a term that is independent of its strategy, to reflect the biological intuition that population growth rates decrease as density increases (“background fitness”). Crowley (Crowley, 2000) defines a *generalized Hawk-Dove game*, modeling interactions between individuals that may differ in size, where size represents “resource holding power”. He considers three different situations, based on the amount of information that interacting individuals have about their sizes: the symmetric case, in which no information about sizes is used, the asymmetric case, in which the individuals know their relative sizes, and a mixed-symmetry case, in which each individual only knows its own size.

In Chapter 2 we define a Hawk-Dove model in a new framework, where players are supposed to be (symmetric) groups of interacting individuals. In Chapter 3 we define the notion of GFESS and we compute such equilibrium as a function of the size of the groups for the group-players Hawk-Dove game. In Chapter 6 we study a particular example of a dynamic version of the game, where players are characterized by an individual state and transition probabilities between two states depend on the action of player’s opponent.

The Standard Hawk-Dove Game

Suppose that two animals contest for a resource (food, territory,...) of value V . They dispose of two possible pure actions, Hawk (H) and Dove (D), where the first correspond to an aggressive behavior, the second to a passive one. The game is played as follows:

- if two Hawks meet, they fight and each of them has equal probability to win the fight and to be injured. The fight has a cost C ;
- if two Doves meet, there is no fight and they equally share the resource;
- if an Hawk meets a Dove, the Hawk gets the resource and the Dove retreats without being injured.

The payoff matrix associated with the game is:

$$\begin{array}{cc} & \begin{array}{cc} H & D \end{array} \\ \begin{array}{c} H \\ D \end{array} & \begin{pmatrix} (V-C)/2 & V \\ 0 & V/2 \end{pmatrix} \end{array} \quad (1.8)$$

There are three possible outcomes, depending on the values of V and C .

- If $V > C$, the game admits a unique pure strict Nash equilibrium, which is also evolutionarily stable (H, H) .
- If $V = C$, the symmetric pure actions pair (H, H) is still the unique Nash equilibrium; in this case it's not strict (as $F_H(H) = F_D(H)$) but it is an ESS.
- If $C > V$, it's an anti-coordination game, which admits three Nash equilibria: the two non-symmetric pure actions pair (H, D) , (D, H) and a mixed NE (q^*, q^*) , with $q^* = V/C$, where q^* is the probability of playing H . The only ESS is (q^*, q^*) : being a random matching game, a non-symmetric Nash equilibrium can't be an ESS.

The replicator equation of the game is:

$$\begin{aligned} \dot{q} &= q(1-q)(F_H(q) - F_D(q)) \\ &= q(1-q)\left(\frac{1}{2}(V-C)q + V(1-q) - \frac{V}{2}(1-q)\right) \\ &= \frac{q(1-q)}{2}(V-Cq), \quad q \in [0, 1]. \end{aligned}$$

It's easy to verify the folk theorem statements and that, for the Hawk-Dove game the interior orbit of the replicator equation converges to the ESS (q^*, q^*) .

Group Players

"A society of ants, bees, or termites achieves a kind of individuality at a higher level. Food is shared to such an extent that one may speak of a communal stomach. Information is shared so efficiently by chemical signals and by the famous 'dance' of the bees that the community behaves almost as if it were a unit with a nervous system and sense of organs of its own."

Richard Dawkins, *The Selfish Gene*, 1976

Summary

In this chapter we present a new model for evolutionary games, in which the concept of the player as a single individual is substituted by that of a player as a whole group. We still consider pairwise interactions among these individuals, but we suppose that the fitness they maximize is that of their group. We analyze a Hawk-Dove game with group players and, as it results to be concave, we prove the existence of a symmetric Nash equilibrium through Rosen's conditions. We explicitly compute it and we obtain that the fact of teaming together makes individuals less aggressive at equilibrium. We finally define a gradient based dynamics as a counterpart of the replicator equation, such the equilibrium of the game is asymptotically stable for such dynamics.

2.1 Introduction

We introduce in this chapter a new scenario for evolutionary games, where the notion of the player as a single individual is replaced by that of a player as a whole group. We still consider pairwise interactions among individuals but we assume that they maximize the fitness of the group they belong to. We suppose that the number of groups is finite and we consider two different scenarios: the case of an infinite and that of a finite population of individuals. We allow pairwise interactions within members of the same group and between individuals of different groups. Groups are supposed to be monomorphic, that is, all individuals in a group play the same (possibly mixed) action. As individuals are indistinguishable, the behavior of an individual is fixed and does not depend on the individual it encounters.

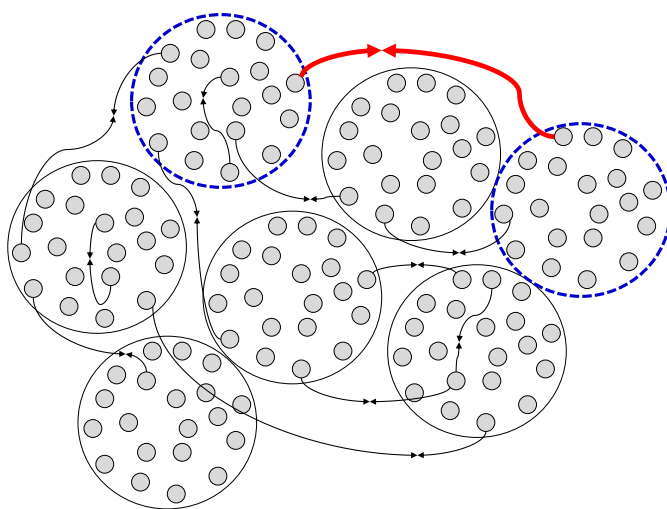


Figure 2.1: Individuals are divided into a finite number of symmetric groups. Each individual can interact with a member of its own group or of a different one.

In order to provide simple but meaningful results, we analyze the Hawk-Dove game in this framework, and we characterize its equilibria. We adopt the Nash equilibrium as solution concept, rather than the ESS. As a matter of facts, in our game, even if the population of individuals may be infinitely large, the number of players is finite, and, as the Nash equilibrium requires stability against the deviation of one player, in our framework this implies stability against the deviation of a whole group of individuals. In the next chapter we present the GFESS, a new concept of equilibrium for group-players games, which considers local deviations within a group. We show here that we obtain a *concave* game and we thus apply Rosen's results ([Rosen, 1965](#)) to prove that the

game admits a unique Nash equilibrium. We finally define a gradient-base dynamics, which converges to the considered equilibrium.

Motivations

In standard EGT, each interacting individual is the player of the game and it chooses its actions in order to maximize its own utility. Many example showing the modeling weaknesses of this assumption can be found. Since in evolutionary games the fitness is defined as related to the reproduction rate, the classical evolutionary game paradigm cannot represent those situations in which only one selected member of a group is responsible for reproduction: in a beehive, for example, the fitness is related to the entire swarm and not to a single bee. Furthermore, in many species, we find altruistic behaviors, which may hurt the individual adopting it, favoring instead the group it belongs to. Altruistic behaviors are typical of parents toward their children: they may incubate them, feed them or protect them from predators at a high cost for themselves. Another example can be found in flock of birds: when a bird sees a predator it gives an "alarm call" to warn the rest of the flock, attracting the predator's attention to itself. Also the stinging behavior of bees is an altruistic one: it serves to protect the hive, but it's lethal for the bees which strives. Other examples can be found also different contexts. As already mentioned, in engineering applications to wireless communication, power control games have frequently been studied in the framework of standard EGT (see e.g. ([Altman and Hayel, 2008](#))). Papers that consider these games usually assume that each mobile can control selfishly its power. In practice however the protocols for power control are not determined by the users of the terminal but by the equipment constructors; this implies that the real competition is among a final number of equipment constructors.

State of Art

The notion of group is not new in evolutionary game theory and different models considering the presence of a finite number of sub-populations have been developed. In those cases, however, the group has a different connotation from the one we introduce here, since it mainly serves to distinguish players with different characteristics and thus to represent asymmetric contexts between individuals. The first evolutionary model considering groups has been introduced by Taylor ([Taylor, 1979](#)), who describes a population composed of two types of individuals, I and J , with inter-group and intra-groups interactions. The state of the population as a whole is described by

the pair of vectors (\mathbf{p}, \mathbf{q}) , which are the distributions over the pure actions of a player of type I and J respectively, with $\mathbf{p} = (p_1, \dots, p_m)$ and $\mathbf{q} = (q_1, \dots, q_n)$. If the population is in state (\mathbf{p}, \mathbf{q}) and a sub-population is in state (\mathbf{r}, \mathbf{s}) , the average fitness of a player of type I is given by: $F(\mathbf{r}|\mathbf{p}, \mathbf{q}) = \sum_{i=1}^m F(i|p, q)$ and the average fitness of type J is: $G(\mathbf{s}|\mathbf{p}, \mathbf{q}) = \sum_{j=1}^n G(j|p, q)$. The ESS for the two populations model is defined as a pair of actions $(\mathbf{p}^*, \mathbf{q}^*)$ which is stable against simultaneous small deviations of \mathbf{p} and \mathbf{q} , which means that, for all (\mathbf{r}, \mathbf{s}) with $\mathbf{r} \neq \mathbf{p}$ or $\mathbf{s} \neq \mathbf{q}$:

$$F(\mathbf{r}|\bar{\mathbf{p}}, \bar{\mathbf{q}}) + G(\mathbf{s}|\bar{\mathbf{p}}, \bar{\mathbf{q}}) < F(\mathbf{p}^*|\bar{\mathbf{p}}, \bar{\mathbf{q}}) + G(\mathbf{q}^*|\bar{\mathbf{p}}, \bar{\mathbf{q}}),$$

where $\bar{\mathbf{p}} = \epsilon \mathbf{r} + (1 - \epsilon) \mathbf{p}^*$, and $\bar{\mathbf{q}} = \epsilon \mathbf{s} + (1 - \epsilon) \mathbf{q}^*$. If the fitness functions are linear in (\mathbf{p}, \mathbf{q}) , with $F(\mathbf{r}|\bar{\mathbf{p}}, \bar{\mathbf{q}}) = \mathbf{r}(A\bar{\mathbf{p}} + B\bar{\mathbf{q}})$, $G(\mathbf{s}|\bar{\mathbf{p}}, \bar{\mathbf{q}}) = \mathbf{s}(C\bar{\mathbf{p}} + D\bar{\mathbf{q}})$, then the ESS condition translates into the conditions: $\mathbf{r}(A\bar{\mathbf{p}} + B\bar{\mathbf{q}}) + \mathbf{s}(C\bar{\mathbf{p}} + D\bar{\mathbf{q}}) \leq \mathbf{p}^*(A\bar{\mathbf{p}} + B\bar{\mathbf{q}}) + \mathbf{q}^*(C\bar{\mathbf{p}} + D\bar{\mathbf{q}})$, where, if the equality holds, then $\mathbf{r}(A\mathbf{r} + B\mathbf{s}) + \mathbf{s}(C\mathbf{p} + D\mathbf{s}) < \mathbf{p}^*(A\mathbf{r} + B\mathbf{s}) + \mathbf{q}^*(C\mathbf{r} + D\mathbf{s})$. Cressman studied (Cressman, 1992) the same scenario but he gives a weaker definition of the ESS: while Taylor's ESS performs better in both groups than any other mutant action pair, he requires just that $(\mathbf{p}^*, \mathbf{q}^*)$ performs better for (at least) one of the two groups, and thus, either:

$$\mathbf{r}(A\bar{\mathbf{p}} + B\bar{\mathbf{q}}) < \mathbf{p}^*(A\bar{\mathbf{p}} + B\bar{\mathbf{q}})$$

or

$$\mathbf{s}(C\bar{\mathbf{p}} + D\bar{\mathbf{q}}) < \mathbf{q}^*(C\bar{\mathbf{p}} + D\bar{\mathbf{q}}).$$

Cressman provides different equivalent definitions of this ESS in the two-species context. Both Taylor's and Cressman's definitions have been later extended to a multipopulation scenario: see the following chapter for an introduction to the N -species ESS.

Our framework is closer to multipopulation games as defined by Sandholm (Sandholm, 2010). He considers a society composed of p populations, each with mass m^i and a set of strategies $S^i, i = 1, \dots, p$. The *population state* \mathbf{x}^i of population i is given by a distribution over the set of strategies S^i , while a *social state* \mathbf{x} describes the behavior in all p populations. The game is defined by the payoff function F , which assigns to each social state a vector of payoffs, one for each strategy in each population. By $\bar{F}^i(\mathbf{x})$, Sandholm denotes the average payoff function in population i . He thus also considers the payoff associated with a whole group of interacting individuals, but, while Sandholm focuses on the different dynamics in this kind of games, introducing the notion of revision protocol, we mainly adopt here a static approach and we consider different solution concepts. In

particular, we study here the Nash equilibrium for a group-player Hawk-Dove game, while in the following chapter we define and we characterize a new equilibrium, the GFESS, in a general framework.

The chapter is structured as follows: in the next section we describe our framework and we define the fitness of a group-player in two different cases: that of an infinite population of individuals and that of a finite one. In Section 2.3 we study the Hawk-Dove game in both contexts: we first prove, in Subsection 2.3.1 the existence and the uniqueness of the equilibrium and we then explicitly compute it. We finally define a gradient based dynamics in Section 2.3.2.

2.2 The General Model

We consider a population of individuals divided into N groups, G_1, \dots, G_N with $N \geq 2$; for simplicity of presentation we will consider symmetric groups of the same size. Let $\mathcal{A} = \{a_1, a_2 \dots a_M\}$ be the set of pure actions and $\Delta(\mathcal{A})$ be the corresponding set of mixed actions. We suppose that the sub-population in each group is *monomorphic*, that is all individuals in the same group adopt the same action. Individuals are indistinguishable and their behavior is fixed. We thus associate to each group G_i the mixed action (or state vector) $\mathbf{p}^i = (p_1^i, \dots, p_M^i) \in \Delta(\mathcal{A})$, with $\sum_{l=1}^M p_l^i = 1$. We define a **multiaction** as the vector of all N group's actions, denoted by $\mathbf{p} = (\mathbf{p}^1, \dots, \mathbf{p}^N) \in \Delta(\mathcal{A})^N$. Let π_{ij} be the probability that an individual in group G_i meets an individual in group G_j . The expected fitness of a (group) player G_i playing $\mathbf{p}^i \in \Delta(\mathcal{A})$ in a population whose multiaction is \mathbf{p} is defined by:

$$\Gamma_i(\mathbf{p}^i, \mathbf{p}^{-i}) = \pi_{ii}(\mathbf{p}^i, \mathbf{p}^i) + \sum_{j=1, j \neq i}^N \pi_{ij} F(\mathbf{p}^i, \mathbf{p}^j), \quad \mathbf{p}^i, \mathbf{p}^j \in \Delta(\mathcal{A}), \quad (2.1)$$

where $\mathbf{p}^{-i} = (\mathbf{p}^1, \dots, \mathbf{p}^{i-1}, \mathbf{p}^{i+1}, \dots, \mathbf{p}^N)$, and $F(\mathbf{p}^i, \mathbf{p}^j)$ (resp. $F(\mathbf{p}^i, \mathbf{p}^i)$) denotes the immediate fitness of an individual playing mixed action \mathbf{p}^i against an opponent playing \mathbf{p}^j (resp. \mathbf{p}^i). The function Γ_i is linear in \mathbf{p}^{-i} . Note that, as the groups are symmetric, the expected fitness only depend on the strategy used by the player and by the population, i.e. $\Gamma_i(\mathbf{p}^i, \mathbf{p}^{-i}) = \Gamma(\mathbf{p}^i, \mathbf{p}^{-i})$, $\forall i = 1, \dots, N$. The probabilities π_{ij} depend on the size of the population. In what follows we see the case of an infinite and that of a finite population of individuals.

The Case of an Infinite Population of Individuals

If the population of individuals is infinitely large, we can assume that the probability of an interaction among two individuals of the same group equals the probability of an interaction among actors of different groups. We thus have that:

$$\pi_{ii} = \pi_{ij} = 1/N \quad \forall i, j = 1, \dots, N.$$

The resulting expected fitness of a (group) player G_i playing $\mathbf{p}^i \in \Delta(\mathcal{A})$ in a population whose multi-action is \mathbf{p} is given by:

$$\Gamma(\mathbf{p}^i, \mathbf{p}^{-i}) = \frac{1}{N} F(\mathbf{p}^i, \mathbf{p}^i) + \frac{1}{N} \sum_{j=1, j \neq i}^N F(\mathbf{p}^i, \mathbf{p}^j), \quad \mathbf{p}^i, \mathbf{p}^j \in \Delta(\mathcal{A}). \quad (2.2)$$

2.2.1 The Case of a Finite Population of Individuals

In order to study the impact of the size of groups on the equilibrium, we now consider a finite population of individuals, with size NK , which is thus divided into N groups of size K . The probability that an individual in group G_i meets an individual in group G_j is given by:

$$\pi_{ij} = \begin{cases} \frac{K-1}{NK-1} & \text{if } i = j \\ \frac{K}{NK-1} & \text{if } i \neq j \end{cases} \quad i, j = 1, \dots, N$$

The expected fitness of a group playing action $\mathbf{p}^i \in \Delta(\mathcal{A})$ in a population whose multi-action is \mathbf{p} can be rewritten as:

$$\Gamma(\mathbf{p}^i, \mathbf{p}^{-i}) = \frac{K-1}{NK-1} F(\mathbf{p}^i, \mathbf{p}^i) + \frac{K(N-1)}{NK-1} \sum_{j=1, j \neq i}^N F(\mathbf{p}^i, \mathbf{p}^j), \quad \mathbf{p}^i, \mathbf{p}^j \in \Delta(\mathcal{A}) \quad (2.3)$$

2.3 Hawk-Dove Game with Group Players

We study here the Hawk-Dove game in the framework of group players. Let $\mathcal{A} = \{H, D\}$; the payoff associated with actors' pairwise interactions is given by the follow-

ing matrix:

$$\begin{array}{cc} & \begin{array}{cc} H & D \end{array} \\ \begin{array}{c} H \\ D \end{array} & \begin{pmatrix} 1/2 - \delta & 1 \\ 0 & 1/2 \end{pmatrix} \end{array}, \quad (2.4)$$

which is obtained from (1.8) with $V = 1$ and $C = 2\delta$. With some abuse of notation, since a mixed action is given by $\mathbf{p} = (p_H, p_D)$ with $p_D = 1 - p_H$, the action of a group i is represented simply by the probability of being aggressive, $p^i \in [0, 1]$, $i = 1, \dots, N$.

2.3.1 Existence and Uniqueness of the Equilibrium

As in our model the number of player is finite and each player corresponds to a group, we adopt the Nash equilibrium as solution concept rather than the ESS. This choice is also justified by the fact that, as the Nash equilibrium is stable against the deviation of on player, in our model this implies the stability against the deviation of a whole group of individuals, as the player is the group.

The existence and the uniqueness of the equilibrium for *concave games* have been proved by Rosen (Rosen, 1965). He considers N players games where each player disposes of a set of mixed actions Δ_i , $i = 1, \dots, N$, and where player's i payoff function $F_i(\mathbf{p})$ depends on the vector of actions of all the the N players, $\mathbf{p} = (p_1, \dots, p_N) \in \Delta := \Delta_1 \times \dots \times \Delta_N$. The game is said to be *concave* if the product space of actions Δ is convex, closed and bounded $\forall \mathbf{p} \in \Delta$, and $F_i(\mathbf{p})$ is continuous in \mathbf{p} and concave in p_i (for fixed \mathbf{p}_{-i}). Then, every concave game admits a Nash equilibrium. Rosen also introduces the weighted sum of payoffs $\sigma(\mathbf{p}, \mathbf{r}) = \sum_{i=1}^N r_i F_i(\mathbf{p})$, with $r_i \in \mathbb{R}_+$ and its **pseudogradient**:

$$g(\mathbf{p}, \mathbf{r}) = \begin{bmatrix} r_1 \nabla_1 F_1(\mathbf{p}) \\ r_2 \nabla_2 F_2(\mathbf{p}) \\ \vdots \\ r_N \nabla_N F_N(\mathbf{p}) \end{bmatrix}. \quad (2.5)$$

The equilibrium point of the game $\mathbf{p} \in \Delta$ is proved to be unique when $\sigma(\mathbf{p}, \mathbf{r})$ is **diagonally strictly concave** for a positive \mathbf{r} , which means that for a fixed non-negative vector $\mathbf{r} \geq 0$ and every vector $\mathbf{q} \neq \mathbf{p}$:

$$(\mathbf{p} - \mathbf{q})' g(\mathbf{q}, \mathbf{r}) + (\mathbf{q} - \mathbf{p})' g(\mathbf{p}, \mathbf{r}) > 0.$$

Hofbauer and Sandholm (Hofbauer and Sandholm, 2008) establish the connection be-

tween diagonally strictly concave games and *stable* population games. They define a p -players game and a p unit masses population game, proving that the diagonal concavity of the first one is equivalent to the stability of the latter.

In what follows we prove that the infinite and the finite-population models are concave games and they both satisfy the diagonal concavity condition, which prove the existence and the uniqueness of the equilibrium. In (Brunetti and Altman, 2013) we present this model also in a third case, where the population of individuals is finite and the number of groups is random, but it does not satisfy the diagonally concavity condition.

Infinite Population

From the definition of $\Delta(\mathcal{A})$, it trivially follows that the product space $\Delta = \Delta(\mathcal{A})^N = [0, 1]^N$ is convex, closed and bounded. When considering an infinite population and the Hawk-Dove payoff matrix (2.4), the fitness of a (group) player adopting action $p^i \in [0, 1]$ in a population whose multiaction is $\mathbf{p} = (p^1, \dots, p^N) \in [0, 1]^N$ (2.2) equals

$$\begin{aligned} \Gamma(p^i, \mathbf{p}^{-i}) &= \frac{1}{N} F(p^i, p^i) + \frac{1}{N} \sum_{j=1, j \neq i}^N F(p^i, p^j) \\ &= \frac{1}{N} \left[\left((-\delta)(p^i)^2 + \frac{1}{2} \right) + \sum_{j=1, j \neq i}^N \left((-\delta p^j + \frac{1}{2}) p^i + \frac{1 - p^j}{2} \right) \right]. \end{aligned} \quad (2.6)$$

Function Γ in (2.6) is continuous in \mathbf{p} and it's easy to verify that it's also concave in p^i , i.e.

$$\Gamma(\epsilon p^i + (1 - \epsilon) q_i, \mathbf{p}^{-i}) > \epsilon \Gamma(p^i, \mathbf{p}^{-i}) + (1 - \epsilon) \Gamma(q_i, \mathbf{p}^{-i}),$$

with $q_i \in [0, 1]$. From Theorem 1 in (Rosen, 1965), it follows that there exists an equilibrium point \mathbf{p}^* for our game. We verify that it is also unique. Let $\sigma(\mathbf{p}, \mathbf{r}) := \sum_{i=1}^N r_i \Gamma(\mathbf{p})$, $r_i \geq 0$ and its **pseudogradient** (Rosen, 1965):

$$g(\mathbf{p}, \mathbf{r}) := \begin{bmatrix} r_1 \nabla_1 \Gamma(p^1, \mathbf{p}^{-1}) \\ r_2 \nabla_2 \Gamma(p^2, \mathbf{p}^{-2}) \\ \vdots \\ r_N \nabla_N \Gamma(p^N, \mathbf{p}^{-N}) \end{bmatrix}. \quad (2.7)$$

To prove that $\sigma(\mathbf{p}, \mathbf{r})$ is diagonally strictly concave for a positive $r \geq 0$, we need to prove that the inequality $(\mathbf{p} - \mathbf{q})' g(\mathbf{q}, \mathbf{r}) + (\mathbf{q} - \mathbf{p})' g(\mathbf{p}, \mathbf{r}) > 0$ holds for all $\mathbf{q} = (q^1, \dots, q^N) \neq \mathbf{p}$. A sufficient condition for this inequality to hold is given in Theorem 6 in (Rosen,

1965). It states that, if the symmetric matrix:

$$[G(\mathbf{p}, \mathbf{r}) + G'(\mathbf{p}, \mathbf{r})],$$

is negative definite, where $G(\mathbf{p}, \mathbf{r})$ is the Jacobian with respect to \mathbf{p} of $g(\mathbf{p}, \mathbf{r})$:

$$G(\mathbf{p}, 1) = \frac{\partial g(\mathbf{p}, 1)}{\partial \mathbf{p}} = \begin{bmatrix} \frac{\partial g(\mathbf{p}, 1)}{\partial p^1} & \cdots & \frac{\partial g(\mathbf{p}, 1)}{\partial p^N} \\ \vdots & \ddots & \vdots \\ \frac{\partial g(\mathbf{p}, 1)}{\partial p^N} & \cdots & \frac{\partial g(\mathbf{p}, 1)}{\partial p^N} \end{bmatrix},$$

then $g(\mathbf{p}, 1)$ is diagonally strictly concave. We fix $\mathbf{r} = (1, \dots, 1)$; the pseudogradient of our game is:

$$g(\mathbf{p}, 1) = \begin{bmatrix} \frac{-2\delta p^1}{N} + \frac{1}{N} \sum_{j \neq 1} (-\delta p^j + \frac{1}{2}) \\ \frac{-2\delta p^2}{N} + \frac{1}{N} \sum_{j \neq 2} (-\delta p^j + \frac{1}{2}) \\ \cdots \\ \frac{-2\delta p^N}{N} + \frac{1}{N} \sum_{j \neq N} (-\delta p^j + \frac{1}{2}) \end{bmatrix}. \quad (2.8)$$

We thus obtain that:

$$G(\mathbf{p}, 1) = -\frac{\delta}{N}(\bar{\mathbf{1}} + I),$$

where $\bar{\mathbf{1}}$ is the $N \times N$ matrix with all the elements equal to 1 and I is the identity matrix. The sum $(\bar{\mathbf{1}} + I)$ has $N - 1$ eigenvectors of the form $(1, 0, -1, 0, \dots)'$ with a corresponding eigenvalue 1 and one eigenvector $(1, 1, \dots, 1)'$ with eigenvalue $N + 1$, which means that $G(\mathbf{p}, 1)$ has strictly negative eigenvalues and thus it is strictly negative definite. From theorem 2 in (Rosen, 1965), this proves the uniqueness of the Nash equilibrium of our game. We explicitly compute it and we resume our results in the following proposition.

Proposition 2. *The group-players Hawk-Dove game with an infinite population of individuals admits a unique symmetric Nash equilibrium which is given by*

$$p^* = \frac{N - 1}{N + 1} \frac{1}{2\delta}. \quad (2.9)$$

Proof. We have already proved the existence and the uniqueness of the equilibrium through Rosen's conditions. We thus compute the symmetric equilibrium value p^* by

maximizing the group fitness function. We have that:

$$\frac{\partial \Gamma(p^i, \mathbf{p}^{-i})}{\partial p^i} = \frac{-2\delta}{N} p^i + \frac{1}{N} \sum_{j=1, j \neq i}^N \left(\frac{1}{2} - \delta p^j \right).$$

By imposing $\frac{\partial \Gamma(p^i, \mathbf{p}^{-i})}{\partial p^i} = 0$, we get:

$$p^i = \frac{1}{2\delta N} \sum_{j=1, j \neq i}^N (-\delta p^j + \frac{1}{2}).$$

To obtain the symmetric Nash equilibrium of the game, we set $p^i = p^j = p^*$, which leads to:

$$p^* = \frac{N-1}{N+1} \frac{1}{2\delta}.$$

□

It is interesting to study the two extreme cases $N \rightarrow \infty$ and $N = 2$.

- When the number of groups is infinite, i.e. $N \rightarrow \infty$, we obtain:

$$p^* = \frac{1}{2\delta}$$

which is the value of the equilibrium of the corresponding standard Hawk-Dove game. This is consistent with a similar result in (Haurie and Marcotte, 1985), that shows the convergence of Nash equilibrium to Wardrop equilibrium as the number of players goes to infinity.

- When we have only two players (formed by an infinity of individuals), i.e. $N = 2$, we obtain:

$$p^* = \frac{1}{6\delta}$$

which means that two groups are less aggressive than two standard players.

Figure 2.2 shows the probability of being aggressive p^* at the equilibrium, as a function of the number of players N in an infinite population of individuals, plotted for three different values of the cost δ (involved in an encounter between two aggressive individuals). We can observe that the equilibrium p^* is an increasing function of the number of groups. Note that when N increases, the probability of meeting a member of a different group also increases: we thus obtain that the level of aggressiveness is higher when the probability of interactions among individuals of different groups increases. Hence, if an individual is aggressive, it causes less damage to its group. This can explain the fact

that the equilibrium probability of being aggressive is increasing in N . As one may expect, we also observe that when the cost δ increases, at the equilibrium, the probability p^* of being aggressive decreases.

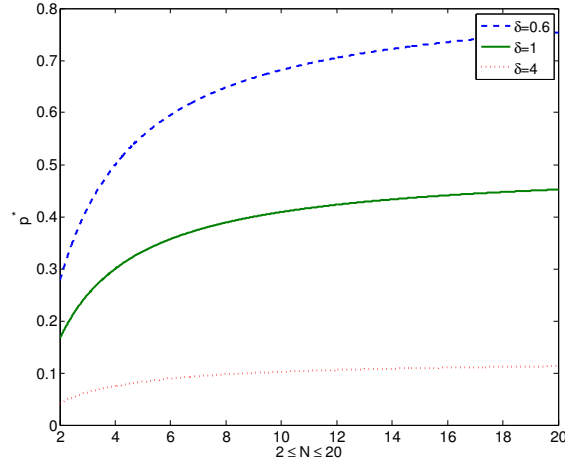


Figure 2.2: The value of p^* as a function of the number of groups $2 \leq N \leq 20$ for three different values of δ . The upper dashed curve is obtained with $\delta = 0.6$, the middle curve with $\delta = 1$ and the lower dotted one with $\delta = 4$.

Finite Population of Individuals

If the population of individuals has finite size NK , with payoff matrix (2.4). The fitness defined in (2.3) becomes:

$$\Gamma(p^i, \mathbf{p}^{-i}) = \frac{K-1}{NK-1} \left[(-\delta)(p^i)^2 + \frac{1}{2} \right] + \frac{K}{NK-1} \sum_{j=1, j \neq i}^N \left[(-\delta p^j + \frac{1}{2})p^i + \frac{1-p^j}{2} \right].$$

Proposition 3. *The Hawk-Dove game with a finite population divided into N groups admits a unique symmetric Nash equilibrium which is given by*

$$p^* = \frac{K(N-1)}{K(1+N)-2} \frac{1}{2\delta}. \quad (2.10)$$

Proof. It is easy to verify that, as in the infinite population case, Rosen's conditions for the existence and the uniqueness of the equilibrium introduced in (Rosen, 1965) are

satisfied. We thus compute the symmetric equilibrium:

$$\frac{\partial \Gamma(p^i, \mathbf{p}^{-i})}{\partial p^i} = -2 \frac{K-1}{NK-1} \delta p^i + \frac{K}{NK-1} \sum_{j=1, j \neq i}^N \left(\frac{1}{2} - \delta p^j \right)$$

and, by imposing $\frac{\partial \Gamma(p^i, \mathbf{p}^{-i})}{\partial p^i} = 0$, we obtain:

$$p^i = \frac{K}{2(K-1)} \sum_{j=1, j \neq i}^N \left(-p^j + \frac{1}{2\delta} \right).$$

If $p^i = p^j = p$, then:

$$p = \frac{K(N-1)}{2(K-1)} \left(\frac{1}{2\delta} - p \right).$$

The symmetric Nash equilibrium is thus given by:

$$p^* = \frac{K(N-1)}{K(1+N) - 2} \frac{1}{2\delta}.$$

□

In Figure 2.3 we plotted the value of the equilibrium p^* as a function of N , with $2 \leq N \leq 20$, for three different values of K and a fixed $\delta = 2$. As in the previous case, we can observe that p^* is an increasing function of N . In Figure (2.4) we plotted p^* as a function of the size of the groups. We can note that p^* rapidly decreases for small K ; when $K > 10$, p^* stabilizes and it is very slowly decreasing. The explanation for this behavior is that when K is small, then the probability of meeting an individual of one's own group is quite sensitive to K , which is not the case when K is large.

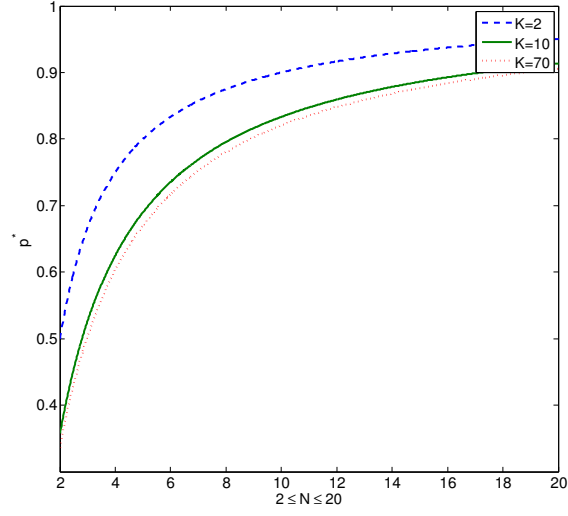


Figure 2.3: The value of p^* as a function of the number of groups $0 \leq N \leq 20$ for three different values of K and $\delta = 2$. Upper dashed line is obtained with $K = 2$, middle line with $K = 10$ and lower dotted line with $K = 70$.

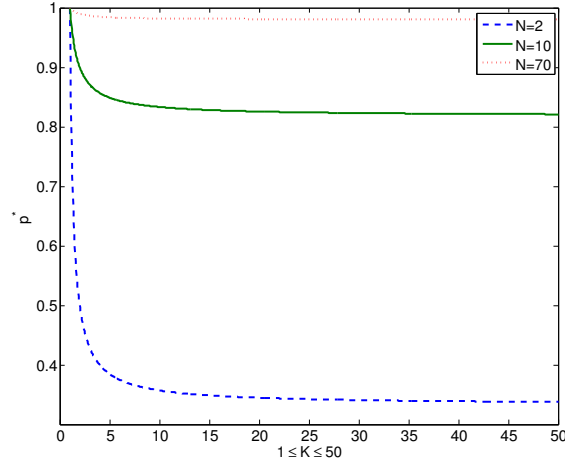


Figure 2.4: The value of p^* as a function of the size of groups $1 \leq K \leq 50$ for three different values of N . Upper dashed line is obtained with $N = 2$, middle line with $N = 10$, lower dotted line with $N = 70$.

2.3.2 Convergence to the Equilibrium

In evolutionary games, instead of interpreting the equilibrium as a static notion, it's preferable to suppose that individual agents can gradually adjust their choices to their current strategic environment, and to study whether the induced distribution of actions converges to a stable situation. To provide a dynamic foundation of the equilibrium of our game, we introduce actions' dynamics in our model, such that the Nash equilibria computed above are asymptotically stable for these dynamics.

We assume that each group-player changes its action in order to increase the payoff of the group. By following Rosen's approach (Rosen, 1965), we can describe the interval $\Delta(\mathcal{A}) = [0, 1]$ through the mapping $h_i(p^i) = (p^i, 1 - p^i)'$, so that $[0, 1] = \{p | h_i(p) \geq 0\}$. Let $h(\mathbf{p}) = (p^1, 1 - p^1, p^2, 1 - p^2, \dots, p^N, 1 - p^N)$ be the function representing all the $2N$ constraints, such that $[0, 1]^N = \{\mathbf{p} | h(\mathbf{p}) \geq 0\}$. We suppose that a group changes its action at a rate proportional to the gradient of its payoff function with respect to its action and subject to the $2N$ constraints. This leads to the action dynamics:

$$\dot{p}^i = r_i \nabla_i \Gamma(p^i, \mathbf{p}^{-i}) + \sum_{j=1}^{2N} u_j \nabla_i h_j(\mathbf{p}), \quad (2.11)$$

where $\Gamma(p^i, \mathbf{p}^{-i})$ is the fitness of G_i and the vector \mathbf{u} lies in a bounded subset $U(\mathbf{p}) \subset \mathbb{R}^{2N}$. The sum on the right hand of (2.11) serves to assure that p^i remains in $[0, 1]$. More precisely, define the $N \times 2N$ matrix $H = [\nabla_1 h(\mathbf{p}) \quad \nabla_2 h(\mathbf{p}), \dots, \nabla_{2N} h(\mathbf{p})]$, which is independent of \mathbf{p} :

$$H = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \dots & \dots \\ 0 & 0 & 1 & -1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \dots & 1 & -1 \end{bmatrix} \quad (2.12)$$

Define the function:

$$f(\mathbf{p}, \mathbf{u}, \mathbf{r}) = g(\mathbf{p}, \mathbf{r}) + H(\mathbf{p})\mathbf{u}.$$

From the definition of the pseudogradient, the system of N dynamic equations can be written as:

$$\dot{\mathbf{p}} = f(\mathbf{p}, \mathbf{u}, \mathbf{r}), \quad \mathbf{u} \in U(\mathbf{p}),$$

where the set $U(\mathbf{p})$ is given by:

$$U(\mathbf{p}) = \{\mathbf{u} \mid \|f(\mathbf{p}, \mathbf{u}, \mathbf{r})\| = \min_{\substack{v_j \geq 0, j \in J, \\ v_j = 0 \text{ otherwise}}} \|f(\mathbf{p}, \mathbf{u}, \mathbf{r})\|\}$$

and:

$$J = J(\mathbf{p}) = \{j \in \{1, \dots, 2N\} \mid h_j(\mathbf{p}) \leq 0\}.$$

According to Rosen's theorem, the negative definiteness of the Jacobian $G(\mathbf{p}, r)$ guarantees the global asymptotic stability of the equilibrium for the system (2.11). We thus use equations (2.11) as a counterpart of replicator dynamics in our groups evolutionary game. The global stability of the equilibrium point permits to determine the equilibrium point for any concave game by appropriate mathematical programming computational methods. In particular, gradient methods for a concave nonlinear programming problem (Rosen, 1965) can be modified to find the equilibrium point for a concave game.

2.4 Conclusion

In this chapter we proposed a new approach to evolutionary games, modeling those situations in which the player is not the individual involved in the interactions, as the fitness is associated with a whole group rather than to the single interacting individual. We studied the Hawk-Dove game, in this group-players framework in two different cases: for an infinite population of individuals and for a finite one. We verified that the obtained Hawk-Dove group-players game is convex, which allowed us to prove the existence and the uniqueness of the Nash equilibrium through Rosen's results (Rosen, 1965). We then explicitly computed it as a function of the number of groups. We finally defined a gradient-based dynamics, such that the equilibrium obtained is asymptotically stable for this dynamics. In the following chapter we define a different concept of equilibrium, related to a notion of local deviation within a group.

Chapter 3

A New Equilibrium Concept: Group Fitness Evolutionarily Stable Strategy

"Imagination is more important than knowledge. For knowledge is limited to all we now know and understand, while imagination embraces the entire world, and all there ever will be to know and understand."

Albert Einstein, 1929

Summary

Following the idea of group players presented in the previous chapter, we introduce here a new concept of equilibrium, the Group Fitness Evolutionarily Stable Strategy, based on the group fitness function and related to a notion of deviation within groups. We compare the GFESS to the standard ESS and we characterize it in the case of two pure actions games. We illustrate our results through some classical examples and we show a possible application in multiple access control framework.

3.1 Introduction

We pursue here the line of research presented in Chapter 2, studying a large population in which pairwise interactions among randomly selected individuals occur, but where the actual player of the game is the whole group. While in Chapter 2 we restrain to symmetric groups of the same size and we study the symmetric Nash equilibria of the game, we now consider groups with different sizes and we introduce a new concept of equilibrium, the **Group Fitness Evolutionarily Stable Strategy (GFESS)**. The stability required by the GFESS is related to a notion of deviation within a group, and thus of a “fraction” of the player (instead of a fraction of a population). We thus do not attempt to simply extend the standard definitions of EGT to a multipopulation setting, but we provide a new modeling framework, where the player is the group but the interactions occur among individuals.

We explore the relationship between GFESS and standard ESS and, for the particular case of two strategies games, we provide a characterization of the equilibria. We then extend the definition of group’s fitness to the case in which the payoff matrix associated with intragroup interactions is different from that associated with intergroups interactions, and we apply this model to study a problem of multiple access control.

State of Art

Garay and Varga ([Garay and Varga, 2000](#)) first define the *N-populations strict ESS* for multipopulation games with symmetric conflicts within the species (populations) and asymmetric conflicts between them. The player is the individual and the fitness is defined through a system of payoff matrices $\{A^{ij}\}$ where A^{ij} represents the payoff of species i in its conflict with species j . Then $\mathbf{p}^* = (\mathbf{p}^{1*}, \dots, \mathbf{p}^{N*})$ is a *N-populations strict ESS* if, for all $\mathbf{p} \neq \mathbf{p}^*$ and for $i \in \{1, \dots, N\}$ with $\mathbf{p}^i \neq \mathbf{p}^{i*}$, there exists $0 < \epsilon_p^i < 1$ such that, for all $0 < \epsilon^i < \epsilon_p^i$ we have:

$$\mathbf{p}^i \left(\sum_{j=1}^N A^{ij} \bar{\mathbf{p}}^j \right) < \mathbf{p}^{i*} \left(\sum_{j=1}^N A^{ij} \bar{\mathbf{p}}^j \right),$$

where $\bar{\mathbf{p}}^j = \epsilon^j \mathbf{p}^j + (1 - \epsilon^j) \mathbf{p}^{j*}$. The authors prove that this is equivalent to require that

$$\mathbf{p}^i \left(\sum_{j=1}^N A^{ij} \mathbf{p}^j \right) < \mathbf{p}^{i*} \left(\sum_{j=1}^N A^{ij} \mathbf{p}^j \right),$$

for all $\mathbf{p} \neq \mathbf{p}^*$ in some neighborhood of \mathbf{p}^* .

The static definition of strict N -populations ESS extends Maynard Smith's standard definition of evolutionary stability and that given by Hofbauer and Sigmund for asymmetric matrix games (J.Hofbauer and K.Sigmund, 1988). The authors also introduce a weaker notion of stability, the *evolutionary dynamical stability*: a state \mathbf{p}^* is said to be evolutionarily dynamically stable if, $\forall \mathbf{p} \neq \mathbf{p}^*$, the vector $(1, 1, \dots, 1) \in \mathbb{R}^N$ is an asymptotically stable rest point of the system

$$\dot{\epsilon}_i = \epsilon_i [\mathbf{p}^{i*} - \bar{\mathbf{p}}^i] [A^{ij} \bar{\mathbf{p}}^j], \quad i = 1, \dots, N.$$

A evolutionarily dynamically stable state is equivalent to the standard ESS for $N = 1$ and to the Cressman definition of two-species ESS for $N = 2$ (Cressman, 1992). In an analogous setting, Cressman et al. (Cressman et al., 2001) focuses on a dynamical approach to the ESS, assuming that selection acts much faster than mutation and thus the incumbent population has enough time between mutations to eliminate the less fit ones. They define a dynamics for the game and they provide a definition of a N -populations ESS related to a property of local asymptotically stability for the dynamics introduced. Their ESS is equivalent to the dynamically stable state in (Garay and Varga, 2000), and it is such that *at least one* of the species earns more than any mutant strategy.

While in (Garay and Varga, 2000) and (Cressman et al., 2001) it is assumed that all populations have the same number of individuals and thus the size of the populations is not considered, in our model the GFESS is defined as a function of the groups' (normalized) sizes, which allows us to study the impact of groups sizes on the equilibrium output.

Sandholm (Sandholm, 2010) compares the two rather different ways of extending the standard definition of ESS to multipopulation games. The *Taylor's condition* (see Chapter 2) for a N -populations state $\mathbf{x} = (x^1, \dots, x^N)$ to be an ESS requires that, for any multipopulation incumbent state $\mathbf{y} = (y^1, \dots, y^N) \neq \mathbf{x}$ in a neighborhood of \mathbf{x} , the aggregate payoff, i.e. the payoff of the society as a whole, of the invading society in state \mathbf{x} , exceeds the aggregate payoff of the society in state \mathbf{y} , and thus $(\mathbf{y} - \mathbf{x})' F(\mathbf{y}) < 0$.

For *Cressman ESS*, it is enough that for one population i in the society composed of N populations, action x^i earns a higher average payoff than the corresponding y^i and thus $(\mathbf{y}^i - \mathbf{x}^i)' F^i(\mathbf{y}^i) < 0$, where $F^i(\mathbf{y}^i)$ denotes the average payoff of population i . The more appropriate extension of the evolutionary stability studied by Maynard Smith and Price, with monomorphic populations of mixed strategists, is the ESS concept defined by Cressman, while Taylor ESS is more useful to understand the dynamics of behavior in polymorphic populations of pure strategists.

Note that for $N = 1$ these N -population games reduces to the standard EGT models whereas in our framework, if there's only one group, there's no game but an optimization problem, since the player is the group.

The chapter is structured as follows. We first introduce in Section 3.2 the definition of the GFESS and we analyze the relationship between GFESS standard ESS. In Section 3.3 we provide the characterization of the GFESS in the case of two actions game. Section 3.4 gives some numerical illustration through three classical examples, the Hawk-Dove game, the Stag Hunt game and the Prisoner's Dilemma. In section 3.5 we study an application in multiple access control in slotted Aloha.

3.2 Group Fitness Evolutionarily Stable Strategy

Consider a large population divided into N groups, denoted by $G_i, i = 1, 2, \dots, N$ where the normalized size of G_i is α_i , with $\sum_{j=1}^N \alpha_j = 1$. Suppose that each individual in group G_i interacts with a member of G_j with probability $\alpha_j, i, j = 1, \dots, N$. Let $\mathcal{A} = \{a_1, a_2, \dots, a_M\}$ be the finite set of pure actions and $\Delta(\mathcal{A})$ the corresponding set of mixed actions; as we did in the previous chapter, we suppose that all individuals in the same group adopt the same (mixed) action and thus, we associate to each G_i a vector $\mathbf{q}^i = (q_1^i, q_2^i, \dots, q_M^i) \in \Delta(\mathcal{A})$, where q_k^i is the probability that an individual in group G_i chooses an action $a_k \in \mathcal{A}$, $\sum_{l=1}^M q_l^i = 1$. Let $\mathbf{q} = (\mathbf{q}^1, \dots, \mathbf{q}^N) \in \Delta(\mathcal{A})^N$ denote the multi-action of the population.

The expected fitness of a group-player G_i is:

$$\Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) = \sum_{j=1}^N \alpha_j F(\mathbf{q}^i, \mathbf{q}^j), \quad (3.1)$$

where $\mathbf{q}^{-i} = (\mathbf{q}^1, \dots, \mathbf{q}^{i-1}, \mathbf{q}^{i+1}, \dots, \mathbf{q}^N) \in \Delta(\mathcal{A})^{N-1}$ and $F(\mathbf{q}^i, \mathbf{q}^j)$ denotes the immediate expected fitness of an individual adopting action \mathbf{q}^i against an opponent playing \mathbf{q}^j . $F(\cdot, \cdot)$ is assumed to be bilinear. Γ_i is thus a weighted sum of the possible individual fitness, where the weights are given by the normalized sizes of the players.

The definition of the GFESS is related to a robustness property against local deviations in a group. While in Chapter 2 we consider symmetric Nash equilibria among groups, which are stable against the deviation of one group-player, here we consider deviations of a 'fraction of the player', i.e. of a share of individuals in a group, and we define and characterize the equilibria (not necessarily symmetric) which are robust against these deviations. More precisely, an ϵ -deviation in group G_i consists in a (possibly) large

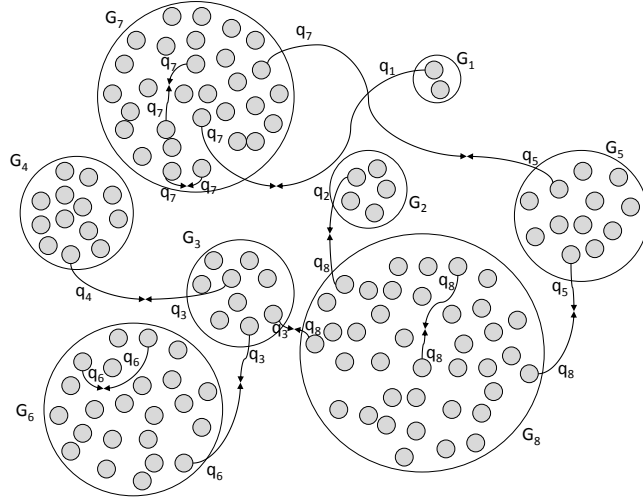


Figure 3.1: The population of individuals is divided into a finite number of groups $G_i, i = 1, \dots, N$. Each individual can interact with a member of its group or of a different one. All individuals in G_i adopts the same strategy q_i (no matter which individual they encounter).

deviation of a small fraction ϵ of individuals of the group from the incumbent action \mathbf{q}^i to a different (mutant) action \mathbf{p}^i . From the definition of the group fitness function 3.1, this is mathematically equivalent to a small deviation in the action by all members of G_i to the mixed action $\bar{\mathbf{p}}^i := \epsilon \mathbf{p}^i + (1 - \epsilon) \mathbf{q}^i \in \Delta(\mathcal{A})$. The population's state after the deviation becomes $(\mathbf{q}^1, \dots, \mathbf{q}^{i-1}, \bar{\mathbf{p}}^i, \mathbf{q}^{i+1}, \dots, \mathbf{q}^N)$ and the average payoff of the mutant group is given by:

$$\begin{aligned} \Gamma_i(\bar{\mathbf{p}}^i, \mathbf{q}^{-i}) &= \sum_{j=1}^N \alpha_j F(\bar{\mathbf{p}}^i, \mathbf{q}^j) = \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) + \epsilon(-\alpha_i \Omega(\mathbf{p}^i, \mathbf{q}^i) + \Gamma_i(\mathbf{p}^i, \mathbf{q}^{-i}) - \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i})) \\ &\quad + \epsilon^2 \alpha_i \Omega(\mathbf{p}^i, \mathbf{q}^i) \\ &= \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) + \epsilon \left(\alpha_i (F(\mathbf{p}^i, \mathbf{q}^i) + F(\mathbf{q}^i, \mathbf{p}^i) - 2F(\mathbf{q}^i, \mathbf{q}^i)) + \sum_{j \neq i} \alpha_j (F(\mathbf{p}^i, \mathbf{q}^j) - F(\mathbf{q}^i, \mathbf{q}^j)) \right) \\ &\quad + \epsilon^2 \alpha_i \Omega(\mathbf{p}^i, \mathbf{q}^i), \end{aligned} \tag{3.2}$$

where $\Omega(\mathbf{p}^i, \mathbf{q}^i) := F(\mathbf{p}^i, \mathbf{p}^i) - F(\mathbf{p}^i, \mathbf{q}^i) - F(\mathbf{q}^i, \mathbf{p}^i) + F(\mathbf{q}^i, \mathbf{q}^i)$.

Definition 4. A multiaction $\mathbf{q} = (\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^N)$ is a Group Fitness Evolutionarily Stable Strategy (GFESS) if for all $i = 1, \dots, N$ and for all $\mathbf{p}^i \neq \mathbf{q}^i \in \Delta(\mathcal{A})$, there exists some $\epsilon_{\mathbf{p}^i} \in (0, 1)$, which may depend on \mathbf{p}^i , such that for all $\epsilon \in (0, \epsilon_{\mathbf{p}^i})$

$$\Gamma_i(\bar{\mathbf{p}}^i, \mathbf{q}^{-i}) < \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}), \tag{3.3}$$

where $\bar{\mathbf{p}}^i = \epsilon \mathbf{p}^i + (1 - \epsilon) \mathbf{q}^i$.

Note that a GFESS is stable against local deviations in one group, i.e., we suppose that only one mutant group can be present in the population at the same time. We define:

$$\Psi_i(\mathbf{p}^i, \mathbf{q}) := \Omega(\mathbf{p}^i, \mathbf{q}^i) - \Gamma_i(\mathbf{p}^i, \mathbf{q}^{-i}) + \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}), \quad (3.4)$$

and we provide, in the following proposition, a characterization of the GFESS, which is equivalent to Definition (4).

Proposition 4. *A multiaction $\mathbf{q} = (\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^N)$ is a GFESS iff it satisfies the two following conditions:*

$$i. \forall \mathbf{p}^i \in \Delta(\mathcal{A}) \quad \Psi_i(\mathbf{p}^i, \mathbf{q}) \geq 0, \quad (3.5)$$

$$ii. \text{ if } \exists \mathbf{p}^i \neq \mathbf{q}^i \text{ such that:} \quad \Psi_i(\mathbf{p}^i, \mathbf{q}) = 0 \Rightarrow \Omega(\mathbf{p}^i, \mathbf{q}^i) < 0 \quad (3.6)$$

Proof. It follows from the definition that multiaction \mathbf{q} is a GFESS iff $\forall \epsilon \in (0, \epsilon_{\mathbf{p}^i})$ the difference $\Gamma_i(\bar{\mathbf{p}}^i, \mathbf{q}^{-i}) - \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i})$ is strictly negative; if we explicit this difference we get:

$$\begin{aligned} & \epsilon^2 \alpha_i \Omega + \epsilon [\alpha_i (F(\mathbf{p}^i, \mathbf{q}^i) + F(\mathbf{q}^i, \mathbf{p}^i) - 2F(\mathbf{q}^i, \mathbf{q}^i)) + \sum_{j \neq i} \alpha_j (F(\mathbf{p}^i, \mathbf{q}^j) - F(\mathbf{q}^i, \mathbf{q}^j))] < 0 \\ \Leftrightarrow & -\epsilon [\alpha_i (F(\mathbf{p}^i, \mathbf{q}^i) + F(\mathbf{q}^i, \mathbf{p}^i) - 2F(\mathbf{q}^i, \mathbf{q}^i)) + \sum_{j \neq i} \alpha_j (F(\mathbf{p}^i, \mathbf{q}^j) - F(\mathbf{q}^i, \mathbf{q}^j))] - \epsilon^2 \alpha_i \Omega > 0. \end{aligned}$$

As ϵ is small, this is true if either the coefficient of ϵ is strictly positive or if it is null and the coefficient of ϵ^2 is strictly positive. As the coefficient of ϵ is:

$$\begin{aligned} & - [\alpha_i (F(\mathbf{p}^i, \mathbf{q}^i) + F(\mathbf{q}^i, \mathbf{p}^i) - 2F(\mathbf{q}^i, \mathbf{q}^i)) + \sum_{j \neq i} \alpha_j (F(\mathbf{p}^i, \mathbf{q}^j) - F(\mathbf{q}^i, \mathbf{q}^j))] \\ & = \alpha_i (\Omega(\mathbf{p}^i, \mathbf{q}^i) - F(\mathbf{q}^i, \mathbf{q}^i) + F(\mathbf{p}^i, \mathbf{p}^i)) + \sum_{j \neq i} \alpha_j (F(\mathbf{p}^i, \mathbf{q}^j) - F(\mathbf{q}^i, \mathbf{q}^j)) \\ & = \alpha_i \Omega(\mathbf{p}^i, \mathbf{q}^i) - \Gamma_i(\mathbf{p}^i, \mathbf{q}^{-i}) + \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) = \Psi(\mathbf{p}^i, \mathbf{q}), \end{aligned}$$

conditions (i) and (ii) straightforwardly follow. \square

3.2.1 Nash Equilibrium and Group-Players

Definition 5. The multiaction $\mathbf{q} = (\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^N)$ is a *Nash Equilibrium* of the N groups game if $\forall i \in \{1, \dots, N\}$

$$\Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) \geq \Gamma_i(\mathbf{p}^i, \mathbf{q}^{-i}), \quad \mathbf{p}^i \neq \mathbf{q}^i. \quad (3.7)$$

If it holds with strict inequality, then \mathbf{q} is a *strict Nash equilibrium*.

It trivially follows that any strict Nash (multiaction) equilibrium is a GFESS. Note that, if multiaction \mathbf{q} satisfies condition (3.6), it implies that there exists an action $\mathbf{p}^i \neq \mathbf{q}^i \in \Delta(\mathcal{A})$ such that:

$$\begin{cases} \alpha_i \Omega(\mathbf{p}^i, \mathbf{q}^i) = \Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) - \Gamma_i(\mathbf{p}^i, \mathbf{q}^{-i}) \\ \Omega(\mathbf{p}^i, \mathbf{q}^i) < 0 \end{cases},$$

and thus:

$$\Gamma_i(\mathbf{q}^i, \mathbf{q}^{-i}) > \Gamma_i(\mathbf{p}^i, \mathbf{q}^{-i}).$$

It follows that, if $\forall i = 1, \dots, N$, condition (3.6) is satisfied $\forall \mathbf{p}^i \neq \mathbf{q}^i \in \Delta(\mathcal{A})$, then multiaction \mathbf{q} is a strict Nash equilibrium. The GFESS we defined can be interpreted as a *local strict Nash equilibrium* in the groups game.

3.2.2 GFESS and Standard ESS

Since we consider pairwise interactions among individuals and the fitness of a group is defined as a weighted sum of individual fitness, we study here the relationship between the ESS in a standard game (with individual players) and GFESS in a group-players game.

Proposition 5. Suppose that the fitness function F related to pairwise interactions among individuals is *symmetric*, i.e. $F(\mathbf{p}, \mathbf{q}) = F(\mathbf{q}, \mathbf{p})$. Then, if $\mathbf{q}^* \in \Delta(\mathcal{A})$ is an ESS in the standard game, multiaction $\mathbf{q} = (\mathbf{q}^*, \dots, \mathbf{q}^*) \in \Delta(\mathcal{A})^N$ is a GFESS in the N group-players game.

Proof. Let $\mathbf{q}^* \in \Delta(\mathcal{A})$ be an ESS in a standard evolutionary game and let $\mathbf{q} = (\mathbf{q}^*, \dots, \mathbf{q}^*) \in \Delta(\mathcal{A})^N$ be a multiaction for the N group-players game. From the symmetry of the fit-

ness function:

$$\begin{aligned}\Psi_i(\mathbf{p}^i, \mathbf{q}^*) &= -\left(\alpha_i(F(\mathbf{p}^i, \mathbf{q}^*) + F(\mathbf{q}^*, \mathbf{p}^i) - 2F(\mathbf{q}^*, \mathbf{q}^*) + \sum_{j \neq i} \alpha_j(F(\mathbf{p}^i, \mathbf{q}^*) - F(\mathbf{q}^*, \mathbf{q}^*))\right) \\ &= -2\alpha_i(F(\mathbf{p}^i, \mathbf{q}^*) - F(\mathbf{q}^*, \mathbf{q}^*)) - (1 - \alpha_i)(F(\mathbf{p}^i, \mathbf{q}^*) - F(\mathbf{q}^*, \mathbf{q}^*)) \\ &= -(1 + \alpha_i)(F(\mathbf{p}^i, \mathbf{q}^*) - F(\mathbf{q}^*, \mathbf{q}^*)).^*\end{aligned}$$

From the definition of ESS, we obtain that $\Psi_i(\mathbf{p}^i, \mathbf{q}^*) \geq 0$, which means that multi-action \mathbf{q} satisfies the first condition (3.5). If there exists a $\mathbf{p}^i \neq \mathbf{q}^*$ such that $\Psi_i(\mathbf{p}^i, \mathbf{q}^*) = 0$, then $F(\mathbf{p}^i, \mathbf{q}^*) = F(\mathbf{q}^*, \mathbf{q}^*)$ and $\Omega(\mathbf{p}^i, \mathbf{q}) = F(\mathbf{p}^i, \mathbf{p}) - F(\mathbf{p}^i, \mathbf{p}^i)$. Since \mathbf{q}^* is an ESS, $F(\mathbf{p}^i, \mathbf{q}^*) = F(\mathbf{q}^*, \mathbf{q}^*)$, implies that $F(\mathbf{p}^i, \mathbf{p}^i) < F(\mathbf{q}^*, \mathbf{p}^i)$, and thus condition (3.5) is satisfied. This completes the proof. \square

In order to provide a characterization of the GFESSs, for simplicity of presentation, in the next section, we consider the case of two pure actions games.

3.3 Analysis of N Groups Games with Two Pure Actions

In this section we study the group-players game when the set of pure actions is given by $\mathcal{A} = \{X, Y\}$. The payoff matrix associated with pairwise interactions among individuals is:

$$A = \begin{matrix} & \begin{matrix} X & Y \end{matrix} \\ \begin{matrix} X \\ Y \end{matrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{matrix}, \quad (3.8)$$

where the entries of the matrix A^{ij} , $i, j \in \{X, Y\}$ give the payoff of the first (row) individual if it plays pure action i against the second (column) individual playing action j . The payoffs of the column (individual) player are given by the transposed of A . A mixed strategy $\mathbf{q}^i \in \Delta(\mathcal{A})$ is given by $\mathbf{q}^i = (q_X^i, q_Y^i)$, where $q_Y^i = 1 - q_X^i$. With some abuse of notation, in the following we identify a strategy simply as the probability of playing strategy A , i.e. $q^i = q_X^i \in [0, 1]$.

Conditions (3.5)-(3.6) can be rewritten as follows:

- $\forall p^i \in [0, 1], i = 1, \dots, N$:

$$\Psi_i(p^i, \mathbf{q}) = (q^i - p^i) \left(\alpha_i(F(q^i, 1) - F(q^i, 0)) + \sum_{j=1}^N \alpha_j(F(1, q^j) - F(0, q^j)) \right) \geq 0. \quad (3.9)$$

- If $\Psi_i(p^i, \mathbf{q}) = 0$ for some $p^i \neq q^i$, then:

$$\Delta < 0, \quad \text{where } \Delta = a - b - c + d. \quad (3.10)$$

We now give a characterization of the possible GFESS for two pure actions games, depending of the values of the payoff matrix and on the size of the groups. We distinguish between pure, mixed and *fully mixed* GEES. Without loss of generality, we reorder the groups according to their size, to have $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_N$.

Pure GFESS

We identify a pure multiaction with a number $N_X \in \{0, \dots, N\}$, which indicates that the N_X first groups use pure action X and the remaining $N - N_X$ groups play pure action Y . For example, $N_X = N$ (resp. $N_X = 0$) stands for the multiaction (X, \dots, X) (resp. (Y, \dots, Y)), where all groups choose pure action X (resp. Y).

Proposition 6. *For the two pure actions game, we have the following characterization of the pure GFESS.*

- i. If $\forall i = 1, \dots, N$:

$$a - c > \alpha_i(b - a) \quad \text{or} \quad \begin{cases} a - c \geq \alpha_i(b - a) \\ \Delta < 0 \end{cases}$$

then $N_X = N$ is a GFESS;

- ii. If $\forall i = 1, \dots, N$:

$$b - d < \alpha_i(d - c) \quad \text{or} \quad \begin{cases} b - d \leq \alpha_i(d - c) \\ \Delta < 0 \end{cases}$$

then $N_X = 0$ is a GFESS;

- iii. Let $1 \leq N_X \leq N - 1$ and define $H(N_X) := \sum_{j=1}^{N_X} \alpha_j(a - c) + \sum_{j=N_X+1}^N \alpha_j(b - d)$. If:

$$\begin{cases} H(N_X) > \alpha_i(b - a) & i = 1, \dots, N_X \\ H(N_X) < \alpha_i(d - c) & i = N_X + 1, \dots, N \end{cases} \quad \text{or} \quad \begin{cases} H(N_X) \geq \alpha_i(b - a) & i = 1, \dots, N_X \\ H(N_X) \leq \alpha_i(d - c) & i = N_X + 1, \dots, N \\ \Delta < 0 \end{cases}$$

then N_X is a GFESS.

Proof.

- i. The pure multiaction $N_X = N$ (corresponding to (X, \dots, X)) is a GFESS in the two actions game if, either it satisfies condition (3.9) with strict inequality for all i , or, if the inequality is not strict, if for some i it holds with equality, for these, the condition (3.10) is verified. Note that $N_X = N$ means that $q^i = 1$ for all $i = 1, \dots, N$. We can then rewrite condition (3.9) as:

$$(1 - p^i) \left(\alpha_i(a - b) + \sum_{j=1}^N \alpha_j(a - c) \right) = (1 - p^i) \left(\alpha_i(a - b) + a - c \right) \geq 0$$

$\forall p^i \neq 1, \forall i = 1, \dots, N$. Since $(1 - p^i) > 0$, the latter inequality is satisfied iff: $a - c \geq \alpha_i(b - a) \quad \forall i$. We thus have that, either condition (3.9) is satisfied with strict inequality for all i , or, if the equality holds for some $i \in \{1, \dots, N\}$, then condition (3.10) needs to be satisfied, i.e. $\Delta < 0$, which completes the proof of (i).

- ii. Following the same line of the proof of (i), for the pure multiaction $N_X = 0$ (corresponding to $q^i = 0$ for all $i = 1, \dots, N$) condition (3.9) can be rewritten as:

$$-p^i \left(\alpha_i(c - d) + \sum_{j=1}^N \alpha_j(b - d) \right) = -p^i \left(\alpha_i(c - d) + b - d \right) \geq 0$$

$\forall p^i \neq 0, \forall i = 1, \dots, N$, which leads to:

$$\alpha_i(c - d) + b - d \leq 0 \quad \forall i = 1, \dots, N,$$

and thus, either $b - d < \alpha_i(c - d)$ for all i or $\alpha_i(c - d) + b - d \leq 0$ and $\Delta < 0$.

- iii. We now consider the pure multiaction N_X , with $N_X \in \{1, \dots, N - 1\}$, (i.e. of the kind $(X, \dots, X, Y, \dots, Y)$). Following the line of (i) and (ii), for N_X , condition (3.9) becomes

$$\begin{cases} \left(\alpha_i(a - b) + \sum_{j=1}^{N_X} \alpha_j(a - c) + \sum_{j=N_X+1}^N (b - d) \right) \geq 0 & i = 1, \dots, N_X \\ \left(\alpha_i(c - d) + \sum_{j=1}^{N_X} \alpha_j(a - c) + \sum_{j=N_X+1}^N (b - d) \right) \geq 0 & i = N_X + 1, \dots, N. \end{cases}$$

By defining $H(N_X) := \sum_{j=1}^{N_X} \alpha_j(a - c) + \sum_{j=N_X+1}^N \alpha_j(b - d)$, from the considerations in (i) and (ii), we obtain (iii), which completes the proof.

□

Fully Mixed GFESS

We say that a multiaction \mathbf{q} is **fully mixed** if each group i uses a mixed action $\mathbf{q}^i = (q_1^i, \dots, q_M^i) \in \Delta(\mathcal{A})$ which assigns positive probability to each pure action, i.e., $0 < q_l^i < 1$ for $l = 1, \dots, M, i = 1, \dots, N$.

Proposition 7. Define, for $i = 1, \dots, N$:

$$q^{i*} := \frac{d - b + ((1 + N)\alpha_i - 1)(d - c)}{(N + 1)\alpha_i \Delta}. \quad (3.11)$$

If $\Delta < 0$ and $0 < q^{i*} < 1, \forall i = 1, \dots, N$, then the multiaction $\mathbf{q}^* = (q^{1*}, \dots, q^{N*})$ is the unique **fully mixed GFESS** equilibrium of the N group-players two pure actions game.

Proof. From the definition of the GFESS for the two actions game, we can observe that a fully mixed multiaction $\mathbf{q} = (q_1, \dots, q_N) \in]0, 1[^N$ can satisfy condition (3.9) only with equality, i.e.:

$$\alpha_i(F(q^i, 1) - F(q^i, 0)) + \sum_{j=1}^N \alpha_j(F(1, q^j) - F(0, q^j)) = 0 \quad \forall i = 1, \dots, N,$$

This leads to

$$\alpha_i \Delta q^i + b - d + \alpha_i(c - d) + \Delta \sum_{j=1}^N \alpha_j q^j = 0.$$

By imposing it for all $i = 1, \dots, N$, we get:

$$q^{i*} = \frac{d - b + ((1 + N)\alpha_i - 1)(d - c)}{(N + 1)\alpha_i \Delta}.$$

$\mathbf{q} = (q^{1*}, \dots, q^{N*})$ is an admissible fully mixed multiaction if $q^{i*} \in]0, 1[$ and, if $\Delta < 0$ it is a GFESS. \square

Remark 2. Note that every fully mixed GFESS is a strict Nash equilibrium since the condition (3.10) is satisfied for all p^i and all i (see Subsection 3.2.1).

Mixed GFESS

We now look for equilibria with both mixed and pure actions. We represent a mixed multiaction by (N_X, N_Y, \mathbf{q}) , where group i for $i = 1, \dots, N_X$ (resp. $i = N_X + 1, \dots, N_X + N_Y$) uses pure action A (resp. B) and the remaining $N - (N_X + N_Y)$ groups adopt a strictly mixed action $q^i \in]0, 1[$.

Proposition 8. *Mixed multiaction in the form (N_X, N_Y, \mathbf{q}) is a GFESS iff:*

$$\begin{cases} \Delta < 0 \\ \alpha_i \Delta + d - b + \alpha_i(c - d) + \Delta(\alpha_{N_X} + y) \geq 0, & i = 1, \dots, N_X \\ d - b + \alpha_i(c - d) + \Delta(\alpha_{N_X} + y) \leq 0, & i = N_X + 1, \dots, N_X + N_Y \\ q^i = \frac{d - b + \alpha_i(d - c) - y\Delta}{\Delta\alpha_i}, & i = N_X + N_Y + 1, \dots, N \end{cases} \quad (3.12)$$

$$\text{where } y = \frac{(N - N_X - N_Y)(d - b) + \Delta\alpha_{N_X} + (d - c) \sum_{j=N_X+N_Y+1}^N \alpha_j}{\Delta(N - N_X - N_Y + 1)}.$$

Proof. Let us consider mixed multiaction (N_X, N_Y, \mathbf{q}) . From the condition (3.9), it is a GFESS iff:

$$\begin{cases} \alpha_i \Delta + b - d + \alpha_i(c - d) + \Delta y > 0, & i = 1, \dots, N_X \\ b - d + \alpha_i(c - d) + \Delta y < 0, & i = N_X + 1, \dots, N_X + N_Y \\ \alpha_i \Delta q^i + b - d + \alpha_i(c - d) + \Delta y = 0, & i = N_X + N_Y + 1, \dots, N \end{cases} \quad (3.13)$$

where $y = \sum_{i=1}^N \alpha_i q^i$. To compute q^i , $i = N_X + N_Y + 1, \dots, N$, we add together the $N - N_X - N_Y$ last equations' left hand sides in (3.13), which gives:

$$\Delta y - \Delta\alpha_{N_X} + (N - N_X - N_Y)(b - d) + (c - d) \cdot \sum_{j=N_X+N_Y+1}^N \alpha_j + \Delta(N - N_X - N_Y)y = 0, \quad (3.14)$$

and thus we obtain:

$$y = \frac{(N - N_X - N_Y)(d - b) + \Delta\alpha_{N_X} + (d - c) \sum_{j=N_X+N_Y+1}^N \alpha_j}{\Delta(N - N_X - N_Y + 1)} \quad (3.15)$$

This agrees with (3.12), completing the proof of the proposition. \square

3.4 Some Examples

In this section we study some examples of games with two group-players and two pure actions. The normalized sizes of the two groups are $\alpha_1 = \alpha$ and $\alpha_2 = 1 - \alpha$, with $0 < \alpha \leq 0.5$. We define three different games, where the payoff associated with a pairwise interaction among individuals belongs, respectively, to the antcoordination class, coordination class and pure dominance class of (standard) games. We compute

the equilibria as a function of α , in order to understand the impact of the relative size of groups at the equilibrium.

3.4.1 Hawk-Dove Game

We first consider the Hawk-Dove game, defined in (1.8), which belongs to the anticoordination class. We fix the values of the parameters $V = 1$ and $C = 2$ and we compute the GFESSs as a function of $\alpha \in [0, 0.5]$. Note that, as $\Delta = -1 < 0$, the second condition (3.10) is always satisfied, and thus GFESSs and strict NEs coincide. We find that:

- the asymmetric pure multiaction (H, D) is a GFESS for $0 < \alpha < 0.2$;
- the asymmetric mixed multiaction (H, q^{2*}) is a GFESS for $0 < \alpha < 0.3$;
- the fully mixed $(q^{1*}(\alpha), q^{2*}(\alpha))$ is a GFESS for $0.3 < \alpha < 0.5$ where, from (3.11):

$$q^{1*}(\alpha) = \frac{2 - 3\alpha}{6\alpha},$$

$$q^{2*}(\alpha) = \frac{3\alpha - 1}{6\alpha}.$$

For $\alpha \in [0, 0.2]$ the game thus admits two equilibria (H, D) and (H, q^{2*}) , whereas, for the other values of α the equilibrium is unique. When $\alpha = 0.5$, we have $q^{1*}(\alpha) = q^{2*}(\alpha) = 1/6$, which corresponds to the symmetric Nash equilibrium obtained in the previous chapter (with $\delta = 1$).

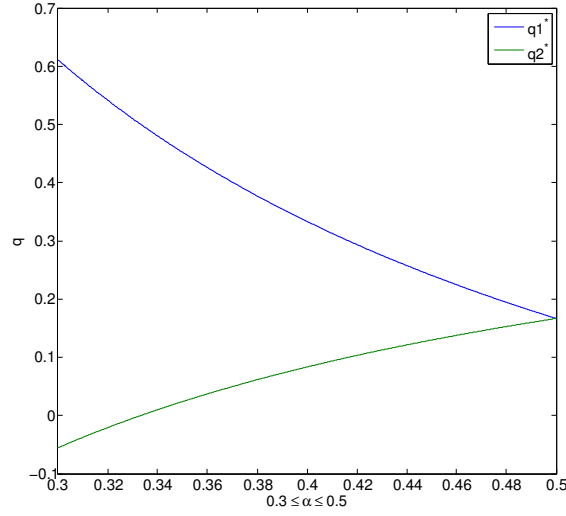


Figure 3.2: The fully mixed equilibrium actions q^{1*} (upper line) and q^{2*} (lower line), respectively of the first and the second group for $0.3 < \alpha < 0.5$ for the two group-players Hawk-Dove game.

3.4.2 Stag Hunt Game

We now consider a well-known example in classical GT, the Stag Hunt game, which belongs to the *coordination class*. The story behind has been described by Rousseau. Two individuals go out on a hunt; if they cooperate they can hunt a stag, otherwise, hunting alone, a hunter can only get a hare. The game represents those situations in which collaboration is rewarding for the players and social cooperation is in conflict with safely one. The general payoff matrix of this game is:

$$\begin{array}{cc} & \begin{matrix} S & H \end{matrix} \\ \begin{matrix} S \\ H \end{matrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \end{array} \quad \text{with } a > c \geq d > b,$$

where S and H stand respectively for Stag and Hare. In standard game theory, coordination games admit two strict pure action NEs (which are thus also ESSs), and a non-strict symmetric mixed NE. For the Stag Hunt game theses are, respectively, the risk dominant equilibrium (H, H) , the payoff dominant one (S, S) and the mixed NE, $q^{1*} = q^{2*} = \frac{d-b}{a-b-c+d}$, which is not evolutionarily stable.

To study the two group-players game with the above payoff matrix associated with pairwise interactions, we set $a = 2$, $b = 0$, $c = 1$, $d = 1$. Since $\Delta = 2 > 0$ the

second condition (3.10) for the GFESS is never satisfied, and thus condition (3.9) must be satisfied with strict inequality. This implies that the group-players Stag Hunt game do not admit the fully mixed GFESS. We obtain the following GFESSs:

- the pure symmetric (S, S) and (H, H) , for all values of α ;
- the pure non-symmetric (S, H) for $0.25 < \alpha < 0.5$;

We observe that the ESSs of the standard game are GFESSs of the group-players Stag Hunt game for any value of α , while, in a specific interval of values of α , our game also admits the non symmetric pure GFESS (S, H) .

3.4.3 Prisoner's Dilemma

We consider another classical example in game theory, the Prisoner's Dilemma, which belongs to the *pure dominance class*. The story beyond the game is the following: two criminals are arrested and separately interrogated. They can either accuse the other, either remain silent. If both of them accuse the other (defect), they will be both imprisoned for two years. If only one accuse the other, the accused is imprisoned for three years while the other is free. If both remain silent (cooperate), each of them will serve one year in jail. The general payoff matrix is the following:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \end{array} \quad \text{with } c > a > d > b,$$

where C and D stand respectively for cooperation and defection. In standard GT, pure dominance class games admit a unique pure, strict and symmetric NE, which is also the unique ESS. For the Prisoner's Dilemma this equilibrium is (D, D) . We set $a = 2, b = 0, c = 3, d = 1$ and we compute the corresponding GFESSs. As in the previous example, condition (3.10) is never satisfied, since $\Delta = 0$. We obtain that the non-symmetric pure multi-action (D, C) is the only GFESSs for all values of $\alpha \in [0, 0.5]$, which means that the smaller group defects and the bigger cooperates.

3.5 GFESSs in a Multiple Access Control Game

In this section we introduce a refinement of our model and we apply it to a MAC problem. We modify the group fitness function defined in (3.1) by supposing that the immediate payoff matrix differs if the interacting individuals belong to the same

group or to two different ones. The expected fitness function of group G_i playing (possibly mixed) action $\mathbf{q}^i \in \Delta(\mathcal{A})$ in a population of N groups playing multi-action $\mathbf{q} = (\mathbf{q}^1, \dots, \mathbf{q}^N) \in \Delta(\mathcal{A})^N$, is defined as follows:

$$\Gamma'_i(\mathbf{q}^i, \mathbf{q}^{-i}) := \alpha_i J(\mathbf{q}^i, \mathbf{q}^i) + \sum_{j \neq i} \alpha_j F(\mathbf{q}^i, \mathbf{q}^j), \quad (3.16)$$

where $J(\cdot, \cdot)$ indicates the immediate expected fitness related to an interaction within a group and $F(\cdot, \cdot)$ is the immediate expected fitness associated with interactions among individuals of different groups. After an ϵ -deviation of group i to action \mathbf{p}^i the fitness of player G_i can be expressed as follows:

$$\begin{aligned} \Gamma'_i(\bar{\mathbf{p}}^i, \mathbf{q}^{-i}) &= \Gamma'_i(\mathbf{q}^i, \mathbf{q}^i) + \alpha_i \epsilon^2 \Omega(\mathbf{p}^i, \mathbf{q}^i) + \epsilon [\alpha_i (J(\mathbf{p}^i, \mathbf{q}^i) + J(\mathbf{q}^i, \mathbf{p}^i) - 2K_i(\mathbf{q}^i, \mathbf{q}^i)) \\ &\quad + \sum_{j \neq i} \alpha_j (J_i(\mathbf{p}^i, \mathbf{q}^j) - J_i(\mathbf{q}^i, \mathbf{q}^j))], \end{aligned}$$

where $\bar{\mathbf{p}}^i = \epsilon \mathbf{p}^i + (1 - \epsilon) \mathbf{q}^i$ and $\Omega'(\mathbf{p}^i, \mathbf{q}^i) := J(\mathbf{p}^i, \mathbf{p}^i) - J(\mathbf{p}^i, \mathbf{q}^i) - J(\mathbf{q}^i, \mathbf{p}^i) + J(\mathbf{q}^i, \mathbf{q}^i)$. We define

$$\begin{aligned} \Psi'_i(\mathbf{p}^i, \mathbf{q}) &:= \alpha_i (2J(\mathbf{q}^i, \mathbf{q}^i) - J(\mathbf{p}^i, \mathbf{q}^i) - J(\mathbf{q}^i, \mathbf{p}^i)) - \sum_{j \neq i} \alpha_j (F(\mathbf{p}^i, \mathbf{q}^j) - F(\mathbf{q}^i, \mathbf{q}^j)) \\ &= \alpha_i \Omega'(\mathbf{p}^i, \mathbf{q}^i) - \Gamma'_i(\mathbf{p}^i, \mathbf{q}^{-i}) + \Gamma'_i(\mathbf{q}^i, \mathbf{q}^{-i}), \end{aligned} \quad (3.17)$$

and, as in Proposition 4, we determine the general conditions for a multi-action $\mathbf{q} = (\mathbf{q}^1, \dots, \mathbf{q}^N)$ to be a GFESS:

- $\forall \mathbf{p}^i \in \Delta(\mathcal{A})$:

$$\Psi'_i(\mathbf{p}^i, \mathbf{q}) \geq 0,$$

- If $\exists \mathbf{p}^i \neq \mathbf{q}^i \in \Delta(\mathcal{A})$ such that:

$$\Psi'_i(\mathbf{p}^i, \mathbf{q}) = 0 \Rightarrow \Omega'(\mathbf{p}^i, \mathbf{q}^i) < 0$$

3.5.1 A MAC Game

We now apply this group-players model to a particular MAC problem. We consider a population of mobiles forming a sparse ad-hoc network, where mobiles compete with their neighbors on the access to a radio channel. We suppose that mobiles are randomly placed over a plane, and that they are divided into a finite number N of groups G_i , with normalized size α_i , $i = 1, \dots, N$. Mobiles are matched through (both inter-groups and intra-groups) pairwise interactions, where each mobile decides either to transmit

(T) or to not transmit (S) a packet to a receiver when they are within transmission range of each other. Interferences occur as in the Aloha protocol: if more than one neighbor of a receiver transmits a packet at the same time then collision occurs and the transmission fails. The channel is ideal for transmission and all errors are due to collisions. Let μ be the probability that a mobile k has its receiver $R(k)$ within its range. When a mobile k transmits, all mobiles within a circle of radius R centered at node $R(k)$ cause interference to k for its transmission to $R(k)$, so that more than one transmission within a distance R of the receiver in the same slot causes a collision and the loss of mobile's k packet at $R(k)$. All mobiles in group G_i transmits with probability $p^i \in [0, 1]$, $i = 1, \dots, N$. If a mobile transmits a packet, it occurs a transmission cost of δ .

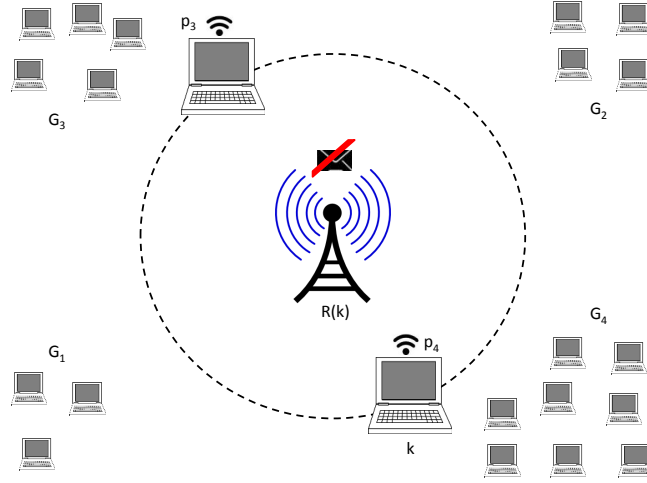


Figure 3.3: When a mobile k transmits, all mobiles within a circle of radius R centered at node $R(k)$ cause interference to k for its transmission to $R(k)$, so that more than one transmission within a distance R of the receiver in the same slot causes a collision and the loss of mobile's k packet at $R(k)$. A transmitter in group G_i transmits with probability $p^i \in [0, 1]$, $i = 1, \dots, N$.

The packet transmission is successful if the other users do not transmit (stay quiet) in that given time slot. If a mobile transmits successfully a packet, it gets a reward of V . We suppose that the payoff V is greater than the cost of transmission, i.e., $\delta < V$. We denote by γ the probability that a mobile is alone in a given local interaction; the tagged mobile does not know whether there is another transmitting mobile within its range of transmission. Let P_1 (resp. P_2) be the matrix representing the immediate fitness of a group when two mobiles belonging to the same group (resp. of two different ones) interact:

$$P_1 \equiv \begin{matrix} & \begin{matrix} T & S \end{matrix} \\ \begin{matrix} T \\ S \end{matrix} & \begin{pmatrix} -2\delta & 1-\delta \\ 1-\delta & 0 \end{pmatrix} \end{matrix}, \quad P_2 \equiv \begin{matrix} & \begin{matrix} T & S \end{matrix} \\ \begin{matrix} T \\ S \end{matrix} & \begin{pmatrix} -\delta & 1-\delta \\ 0 & 0 \end{pmatrix} \end{matrix}.$$

The definition of P_1 implies that when two mobiles of the same group G_i interact, any successful transmission is equally rewarding for group i . When two of mobiles in group G_i interact and play $q^i \in [0, 1]$, the expected fitness of G_i is:

$$\begin{aligned} J(q^i, q^i) &= \mu \left(q^i [\gamma(1 - \delta) + (1 - \gamma)((1 - \delta)(1 - q^i) - 2\delta q^i)] + (1 - \gamma)(1 - \delta)(1 - q^i)q^i \right) \\ &= \mu q^i [(1 - \delta)(2 - \gamma) - 2(1 - \gamma)q^i]. \end{aligned} \quad (3.18)$$

If a mobile in G_i interacts with a mobile in a different group G_j playing $q^j \in [0, 1]$, its expected payoff is the following:

$$\begin{aligned} F(q^i, q^j) &= \mu q^i [\gamma(1 - \delta) + (1 - \gamma)((1 - \delta)(1 - q^j) - \delta q^j)] \\ &= \mu q^i [1 - \delta - (1 - \gamma)q^j] \end{aligned} \quad (3.19)$$

The total expected payoff of G_i is then given by:

$$\Gamma'_i(q^i, \mathbf{q}^{-i}) = \mu q^i [1 - \delta + (1 - \gamma)(\alpha_i(1 - \delta - q^i) - \sum_{j=1}^N \alpha_j q^j)] \quad (3.20)$$

The multiaction \mathbf{q} is a GFESS of the group-players MAC game if $\forall i = 1, \dots, N$:

$$\Psi'_i(p^i, q) \equiv (q^i - p^i)[1 - \delta + (1 - \gamma)(\alpha_i(1 - \delta - 2q^i) - \sum_{j=1}^N \alpha_j q^j)] \geq 0.$$

Note that, in this game, the condition $\Omega'(\mathbf{p}^i, \mathbf{q}^i) < 0$ reduces to $(p^i - q^i)^2(1 - \gamma)\alpha_i > 0$. Since the latter inequality is always true, the second condition is always satisfied and thus the first condition is sufficient to guarantee the existence of a GFESS. In the following proposition we give a characterization of the GFESSs of the presented MAC game. Without loss of generality, we reorder the groups so that $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_N$.

Proposition 9. *We find that:*

- i. *The pure symmetric multiaction (S, \dots, S) is never a GFESS.*
- ii. *If a fixed group G_i adopts pure action T , then if $i < N$, at the equilibrium all smaller groups transmit. If $i = N$, i.e. the bigger group G_N uses action T at the equilibrium, then (T, \dots, T) is an equilibrium iff $\gamma > \bar{\gamma}$.*
- iii. *If a fixed group G_i adopts pure action S , then at the equilibrium, all smaller groups also use S .*

iv. The game admits a unique fully mixed GFESS $\mathbf{q}^* = (q^{1*}, \dots, q^{N*})$, given by:

$$q^{i*} = \frac{(1-\delta)(1+\gamma+(1-\gamma)(2+N)\alpha_i)}{2(N+2)(1-\gamma)\alpha_i} \quad (3.21)$$

under the condition: $\gamma < \underline{\gamma}$.

v. If a fixed group G_i adopts a mixed action $q^i \in]0, 1[$, then if $q^i > \frac{1-\delta}{2}$, at the equilibrium all smaller groups may use pure action T , whereas if $q^i < \frac{1-\delta}{2}$, smaller groups may play S .

The thresholds $\underline{\gamma}$ and $\bar{\gamma}$ are defined as follows:

$$\underline{\gamma} \equiv \min_{\alpha_i} \frac{\alpha_i(N+2)(1+\delta) - (1-\delta)}{\alpha_i(N+2)(1+\delta) + (1+\delta)},$$

$$\bar{\gamma} \equiv \max_{\alpha_i} \left(1 - \frac{1-\delta}{\alpha_i(\delta+1)+1} \right).$$

Proof. The multiaction $\mathbf{q} = (q^1, \dots, q^N) \in [0, 1]^N$ is a GFESS if $\forall i = 1, \dots, N$, the condition $\Psi'_i(\mathbf{p}^i, \mathbf{q}) \geq 0$ is verified $\forall p^i \in [0, 1]$.

i. If, $\forall i = 1, \dots, N$ $q^i = 0$, then

$$\Psi'_i(\mathbf{p}^i, \mathbf{0}) = -p^i[1-\delta+(1-\gamma)(1-\delta)\alpha_i] < 0 \quad \forall p^i \in [0, 1],$$

which proves that (S, \dots, S) is never a GFESS.

ii. Let \mathbf{q} be a GFESS such that $q^i = 1$ for a fixed i . This implies that

$$1-\delta-(1-\gamma)(\alpha_i(1+\delta)+Y) \geq 0,$$

with $Y = \sum_{j=1}^N \alpha_j q^j$. Then, if $\alpha_j < \alpha_i$ we have that

$$1-\delta-(1-\gamma)(\alpha_j(1+\delta)+Y) \geq 1-\delta-(1-\gamma)(\alpha_i(1+\delta)+Y) \geq 0$$

and thus $q^j = 1$ satisfy the GFESS condition. If all the groups transmit, then the condition for the GFESS is satisfied iff:

$$1-\delta-(1-\gamma)((1+\delta)\alpha_i-1) \geq 0 \quad \forall i = 1, \dots, N$$

and thus $\gamma \geq 1 - \frac{1-\delta}{\alpha_i(\delta+1)+1}, \forall i$.

iii. Let \mathbf{q} be a GFESS such that $q^i = 0$ for a fixed i . This implies that

$$1 - \delta + (1 - \gamma)(\alpha_i(1 - \delta) - Y) \leq 0.$$

If $\alpha_j < \alpha_i$, then:

$$1 - \delta + (1 - \gamma)(\alpha_j(1 - \delta) - Y) \leq 1 - \delta + (1 - \gamma)(\alpha_i(1 - \delta) - Y) \leq 0$$

and thus $\Psi'_j(\mathbf{p}^i, \mathbf{q}) \geq 0$.

- iv. Let \mathbf{q} be a fully mixed GFESS. Then, $\forall i: 1 - \delta + (1 - \gamma)(\alpha_i(1 - \delta - 2q^i) - Y) = 0$. After some algebra we thus obtain that $Y = \frac{(1 - \delta)(N + 1 - \gamma)}{(1 - \gamma)(N + 2)}$, and by substituting it in the previous equations we obtain the expressions of q^i . By imposing that $0 < q^i < 1 \forall i$ we obtain the condition $\gamma < \underline{\gamma}$.
- v. Let \mathbf{q} be a GFESS such that $q^i \in]0, 1[$ for a fixed i . Then, if for a $j < i$, $q^j = 0$ (resp. 1), $\Psi'_j(\mathbf{p}^i, \mathbf{q}) \geq 0$ iff $q^i > \frac{1 - \delta}{2}$ (resp. $q^i > \frac{1 - \delta}{2}$).

□

As an example, we consider a two groups MAC game, in which we fix a low value of the cost of transmission, $\delta = 0.2$, and we study the equilibria as a function of the value of the parameter γ , which is the probability that the transmitter is alone. We first set $\alpha = 0.5$ and we obtain that, for $\gamma < \underline{\gamma} = 0.35$ the game admits a fully mixed symmetric GFESS (q^*, q^*) . In figure 3.4 we compare q^* to the value of the equilibrium in the corresponding individual-players game: $q_{std}^* := \min(1, \frac{1}{1 - \gamma} - \Delta)$, and we observe that q^* is lower than q_{std}^* . Let $p_S(\mathbf{q})$ be the *probability of a successful transmissions* in a population under profile \mathbf{q} . For $N = 2$, it equals:

$$p_S(\mathbf{q}) := \mu[\gamma(\alpha q^1 + (1 - \alpha)q^2)] + (1 - \gamma)(2\alpha^2 q^1(1 - q^1) + \alpha(1 - \alpha)((1 - q^2)q^1 + (1 - q^1)q^2) + 2(1 - \alpha)^2 q^2(1 - q^2)).$$

In figure 3.5 we plot the probability $p_S^* = p_S(q^*)$ of a successful transmission in a population under the equilibrium (q^*, q^*) as a function of γ for $\alpha = 0.5$. We note that the probability to transmit successfully is higher for the two group players game than in the standard case.

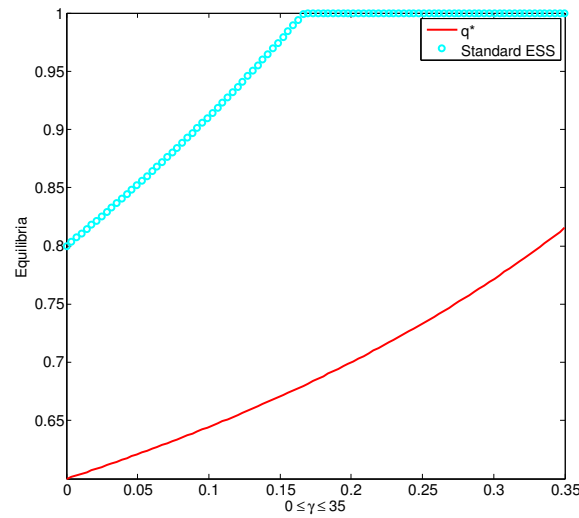


Figure 3.4: The value of the equilibrium strategy q^* in a two symmetric groups MAC game as a function of γ for $\alpha = 0.4$, compared to the ESS q_{std}^* of the standard game.

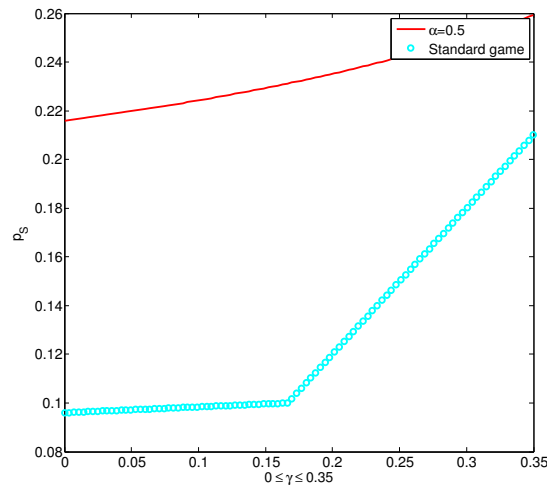


Figure 3.5: The probability $p_s(\mathbf{q}^*)$ of a successful transmission at the equilibrium for the fully mixed GFESS in a two symmetric groups MAC game and in the standard game, as a function of γ .

We then set $\alpha = 0.4$, to consider an asymmetric case, where the two groups have different sizes. For $\gamma < \underline{\gamma} = 0.3$ the game admits a fully mixed GFESS (q^{1*}, q^{2*}) , and for $\gamma = \bar{\gamma} > 0.53$, the game admits a pure symmetric equilibrium (T, T) . These equilibria are plotted in 3.6. In the interval $0 \leq \gamma < 0.4$ we also have a GFESS in the form (T, q^{T*}) . In figure 3.7 we plot the equilibrium actions of the second group q^{T*} and q^{2*} , and we

observe that the latter is lower than q^{T*} , i.e. the probability that the second (bigger) group transmits at the equilibrium is higher if the first group plays pure action T than if it plays mixed action q^{1*} .

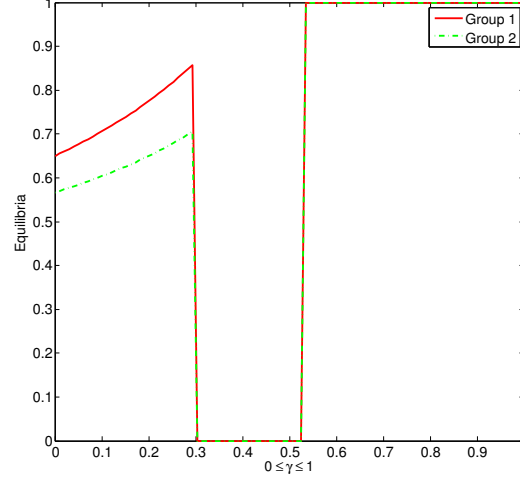


Figure 3.6: The value of the equilibria of the two groups in a two groups MAC game as a function of γ for $\alpha = 0.4$

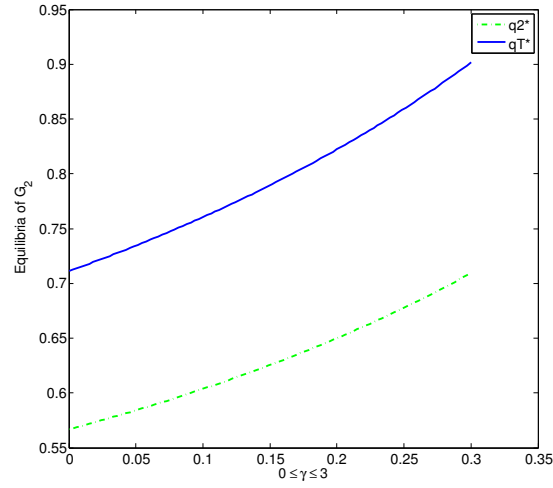


Figure 3.7: The value of the equilibrium strategy of the second group in the fully mixed equilibrium (q^{1*}, q^{2*}) and in the mixed equilibrium (T, q^{T*}) as a function of γ for $\alpha = 0.4$.

3.6 Conclusion

In this chapter we presented a new concept of evolutionary stability in the group-players framework, the GFESS, implying stability against local deviations within each group. While the ESS in standard EGT is a refinement of the Nash equilibrium, the GFESS can be seen as a strict local Nash (in the group context), which is stable against local deviations within a group. For the particular case of two pure actions games, we provided a characterization of the GFESSs, distinguishing between pure, fully mixed and mixed GFESS, and we studied three classical examples considering group players: the Hawk-Dove game, the Stag Hunt game and the Prisoner's Dilemma. By computing the equilibria as functions of the size of the groups, we saw how the presence of groups impacts the equilibrium output. We then introduced a slightly different situation by redefining the fitness of a group, where the immediate payoff associated with the pairwise interactions among individuals in the same group differs from that of inter-group interactions and we applied this model to a MAC problem.

Part II

Individual State and Policy Dynamics in Evolutionary Games

Markov Decision Evolutionary Games

"A mathematician, like a painter or a poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made with ideas. A painter makes patterns with shapes and colours, a poet with words. [...] A mathematician, on the other hand, has no material to work with but ideas, and so his patterns are likely to last longer, since ideas wear less with time than words. The mathematician's patterns, like the painter's or the poet's must be beautiful; the ideas like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics."

G.H. Hardy, A Mathematician's Apology, 1940

Summary

In this chapter we briefly introduce Markov Decision Evolutionary Games (MDEG), which combine Markov Decision Processes and Evolutionary Game Theory. These games involve a large population of players characterized by an individual state and randomly matched in pairwise interactions. The fitness of an individual depends on the actions played in the interaction and on the distribution of the individual states in the population. The action taken by a player also determines the transition probabilities to its next individual state. Players aim at maximizing the average sum of their immediate expected fitness during their finite life time. Under certain assumptions, it is possible to transform this game into a standard evolutionary game and find its equilibria.

4.1 Introduction

In this chapter we present a particular class of *stochastic evolutionary games*, Markov Decision Evolutionary Games (MDEG), as defined by Altman and Hayel ([Altman and Hayel, 2010](#)). In MDEG games, each player belonging to a large population is associated with an individual state. During its finite life time the player meets several times other users through random pairwise interactions and it may move among different states. The actions played by an individual determine not only its immediate fitness but also the transition probabilities to its next state. The objective of a player is the maximization of its expected immediate fitness during its life time, which depends also on the distribution of individual states in the populations. Following a method similar to that used in ([Filar and Raghavan, 1984](#)) for one controller, the authors transform this game into an equivalent standard evolutionary game and they study its ESSs.

State of Art

Stochastic games have been first introduced by Lloyd Shapley ([Shapley, 1953](#)). He models a play which proceeds by steps from "position" to "position" (state to state), according to transition probabilities jointly controlled by two players. Each state (or position) corresponds to a specific matrix game and the transition probabilities from one position to the other are determined by the current matrix game and by the actions used by the two players. The transition function considered is time homogeneous and players' evaluation of sequences of payoffs have a stationary structure. The author proves the existence of optimal *stationary strategies*.

Stochastic games belongs to the wider class of *sequential games*, which are processes generating a sequence of one-shot, non-cooperative games played by the same set of players. The transition probabilities which determine the game to be played at any time $t + 1$ are functions of t , the game played at t , and the actions chosen at t . A stochastic game is thus a sequential game where these transition functions are stationary and the players' evaluations of the sequences of payoffs have a stationary structure. Jovanovic and Rosenthal ([Jovanovic and Rosenthal, 1988](#)) define *anonymous sequential games* as processes of one-shot, non-cooperative games with a continuum of players, where individuals affect their opponents only when aggregated, and not at an individual level. While in stochastic games the notion of state is associated with the system, Jovanovic and Rosenthal relate it to individuals. Furthermore, they look at equilibria composed of "distributional strategies" (which are sequences of state-action distributions) and therefore they do not restrain to the stationary state of the system. After proving that every

anonymous sequential game admits an equilibrium, they also consider the stationary case, where the utility and the transition probabilities do not depend on time.

Flesch et al. (Flesch et al., 2013) unify stochastic games as defined by Lloyd Shapley and EGT into *stochastic evolutionary games*, which are described as *irreducible* two-person stochastic games with symmetric payoff matrices and symmetric transition probabilities. At each stage of the game, two players belonging to a large population interact with each other in one of finitely many *environments*. The transition probabilities between the environments determine the impact of each of these environments on the fitness of the individuals. Under the *irreducible transition law* assumption, the authors extend the notion of evolutionary stability in terms of stationary strategies and define the replicator dynamics in this multi-state framework. In contrast with (Flesch et al., 2013), the notion of state in MDEG has an individual connotation, as in (Jovanovic and Rosenthal, 1988), but transition probabilities and policies are assumed to be stationary.

Stochastic games and Markov Decision Processes (MDP), are both two suitable mathematical structures for the classes of dynamic, stochastic, decision models. Filar and Vrieze (Filar and Vrieze, 1997) present the theories of stochastic games and MDP in a complete unified fashion, and by defining *Competitive Markov Decision Processes*, they emphasize the importance of the link between these two topics extensively studied by mathematicians, engineers and economists.

The mutations in a population modeled by standard EGT are random deviations, which means that the notion of randomness is somehow implied by the requirement of evolutionary stability. However, since the population considered is large, evolutionary games can be thought as continuous deterministic approximations of discrete stochastic games, where a finite set of players may take random actions.

Tembine et al. (Tembine et al., 2009) analyze the connection between stochastic and deterministic evolutionary games, studying the *mean-field asymptotic* of MDEG. More precisely, they define a *finite population game*, called *mean field interaction*, in which, at each stage, each individual interacts with other randomly selected players, and thus interactions may involve more than two individuals. The states and the actions of each player in an interaction determine the instantaneous payoff for all the players involved. Actions also determine the transition probabilities to the next state of the players involved in the interactions. The authors provide a rigorous derivation of the asymptotic behavior of this system when the size of the population grows to infinity and they show that the large population asymptotic of the mean field interaction is equivalent to a MDEG in which a local interaction is described by a single player against a population profile.

Applications

MDEG can find a wide range of applications in different fields. Altman and Hayel present a simple application of MDEG to mobile communications in (Altman and Hayel, 2008), where mobile terminals transmit packets occasionally and their destination may occasionally receive simultaneously a transmission from another terminal which results in a collision. When packets collide, one of the packets can still be received correctly if transmitted at a higher power. Successful transmissions is rewarded, but energy consumption is penalizing. The action corresponds to the choice of power transmission level, and the set of available actions depend on the individual state of the player, which consists in the level of the battery. The state can be either full, in which case there are two transmitting power levels available, or almost empty in which only the weak power level is available. Transmission at high power is costly in terms of energy consumption and thus it results in a higher probability of moving to the lower energy level. The authors define the fitness of a terminal as the total number of packets successfully transmitted during its lifetime and characterize the ESS in this context.

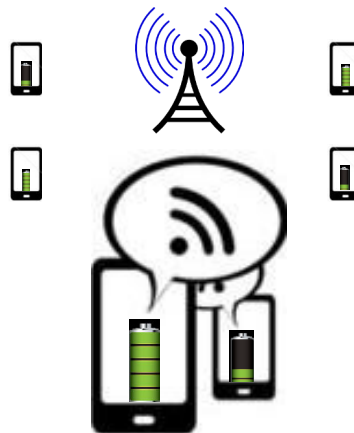


Figure 4.1: Each mobile can be in full or in almost empty battery state. In state full, it can choose to transmit at high power or at low power, while, if the battery is almost empty, it can only transmit at low power. If a destination receives simultaneously a transmission from two terminals, packets collide, but one of the packets can still be received correctly if transmitted at a higher power.

In Chapter 5 we'll develop the dynamical aspect related to MDEG theory, studying the interdependent dynamics of individual states and policies, where players are supposed to control a continuous time MDP and we illustrate other possible application fields. In Chapter 6 we'll see a particular dynamic version of the Hawk-Dove game in a MDEG kind of framework.

In what follows, we first give some basic notions of Markov Decision Process (MDP) and we then introduce MDEG as presented in (Altman and Hayel, 2010), focusing on the notion of *occupation measure* and that of *equivalence* between policies.

4.2 Markov Decision Process

A Markov Decision Process (MDP) can be identified by the tuple $\{S, \mathcal{A}, Q\}$, where:

- S is the finite set of individual **states**, with $|S| = K$;
- \mathcal{A} is the set of available actions. For each state $s \in S$, $\mathcal{A}_s \subseteq \mathcal{A}$ is the subset of available **actions** for a player in state s .
- Q is the set of **transition probabilities**. Given the individual states $s, s' \in S$ and the actions $a, a' \in \mathcal{A}$, $Q_{s'}^t(s, a, a')$ denotes the probability to move from state s to state s' taking action a when interacting with an individual that takes action a' at time t .

In what follows we restrain to *time homogeneous* transition probabilities, which do not depend on t , and we also suppose that they only depend on the player's action (a) and not on that of its opponent (a'). A **policy** $u = \{u_0, u_1, u_2, \dots\}$ of a player is a sequence of probability measures on \mathcal{A} chosen as functions of the present state of the player and on the history of its previous states and actions. That is, $u_0 = u_0(\eta)$, where η is the initial state and $u_{t+1} = u_{t+1}(h^t, s_{t+1}, \cdot)$, where s_{t+1} is the state at time $t + 1$ and h^t is the **state-action history** of the player until time t :

$$h^t := (s_l, a_l, \quad l = 0, \dots, t) \quad t = 0, 1, 2, \dots \quad (4.1)$$

Let \mathbf{H} be the set of all possible states-actions histories which can be observed, and \mathcal{U} the set of all policies. We further define the following particular classes of policies:

- \mathcal{U}_M is the set of **Markov policies** of a player such that, at every moment t of decision making, the probability measure u_t on \mathcal{A} is chosen as a function of the current moment of time and the current state of the player (and not on the past history);
- \mathcal{U}_S is the set of **stationary policies**, which are such that at any moment t , the probability measure u_t on \mathcal{A} is chosen only as a function of the current state of the player (and not of the current moment of time);
- \mathcal{U}_D is the set of **deterministic policies**, (also called **pure stationary policies**) which

assign an action to each state.

4.2.1 Occupation measures

In MDEG each player belonging to a large population is associated with a MDP and the actions played are supposed to determine the fraction of time the player spends in each state. This implies that the share of the population at a given state may depend on the distribution of strategies in the population. In order to model this dependence, we need to define the concept of occupation measure, which describes the expected amount of time that an individual spends in a given state under a given policy.

Denote by $f_{\eta,u}(s, a)$ the expected number of time units during which a player is in state s and chooses action a under policy u , given the initial distribution η . The **occupation measure** corresponding to policy u is the quantity:

$$f_{\eta,u} := \{f_{\eta,u}(s, a)\}.$$

Let us now define the probability that an individual is in state s using action a under policy u at local time t , given the initial state probability distribution η :

$$p_t(\eta, u; s, a) := \mathbb{P}_\eta^u(X_t = s, A_t = a).$$

Note that, since the population is supposed to be totally symmetric, this initial distribution, at local time 0, is the same for all individuals. Define:

$$p_t(\eta, u; s) := \sum_{a \in \mathcal{A}} p_t(\eta, u; s, a).$$

Note that it's a sub-probability measure, as $\sum_s p_t(\eta, u; s)$ may be smaller than one. We have that:

$$f_{\eta,u}(s, a) = \sum_{t=0}^{\infty} p_t(\eta, u; s, a), \quad f_{\eta,u}(s) = \sum_{t=0}^{\infty} p_t(\eta, u; s). \quad (4.2)$$

The lifetime of an individual is identified as the time interval before X_t leaves S ; the **expected life-time** (corresponding to initial state η and policy u) is thus defined as:

$$T_{\eta,u} = \sum_s f_{\eta,u}(s).$$

For a given initial state η , $\sup_{u \in \mathcal{U}} T_{\eta,u}$ is assumed to be finite. From (Kallenberg, 1983),

$\sup_{u \in \mathcal{U}} T_{\eta,u} = \max_{u \in \mathcal{U}_D} T_{\eta,u}$, which means that this assumption is equivalent to requesting that the expected lifetime is finite for all deterministic policies.

From the theory of MDP, if all players use an equilibrium stationary policy $u_s \in \mathcal{U}_S$, then no player can benefit by a unilateral deviation to any policy, including non-stationary ones. Hence u_s is an equilibrium among all policies and, with no loss of generality, it is possible to restrain to the set \mathcal{U}_S .

4.3 Markov Decision Evolutionary Games

MDEG combine MDP and EGT into a new game theoretic framework. Consider, as in standard EGT, a large population of players, randomly matched in pairwise interactions. Each player is characterized by an individual state, such that the fitness of a player does not depend only on the actions chosen in the interactions, but also on this individual state. The action taken by a player determines the transition probability to its next individual state. Each player is thus associated with a MDP, where the transition probabilities \mathcal{Q} are time homogeneous and depend only on the action of the player (and not on the action chosen by its opponent).

Let $r(s, a, s', a')$ be the **immediate fitness** that a player receives when it is in state s and uses action a in an interaction with a player who is in state s' and uses action a' . When the whole population uses a stationary policy $u \in \mathcal{U}_S$, then at any time t (which is either fixed or is an individual time of an arbitrary player) the state of the system is independent of t . For all $(s, a) \in \mathcal{S} \times \mathcal{A}_S$, the fraction of the population in stationary regime in individual state s , that uses action a when all the population uses stationary policy u is given by:

$$\alpha(u; s, a) = \frac{f_{\eta,u}(s, a)}{T_{\eta,u}}. \quad (4.3)$$

Then, the **stationary system state** is:

$$\alpha(u) = \{\alpha(u; s, a)\}. \quad (4.4)$$

Denote by $r(u; s, a)$ the **immediate reward** that a player receives when it is in state s and it uses action a while interacting with a player whose policy is u . We have that:

$$r(u; s, a) = \sum_{(s', a')} \alpha(u; s', a') r(s, a, s', a').$$

Consider an arbitrary tagged player and let X_t and A_t be respectively its state and

action at time t (as measured on its individual clock). Then its **expected immediate reward** at that time is given by

$$R_t(u) = r(u; X_t, A_t).$$

Assume now that a player arrives in the system at time 0. The **global expected fitness** of a player choosing a policy v in a population whose policy is u is then:

$$F_\eta(v, u) = \sum_{t=0}^{\infty} \mathbb{E}_{\eta, v}[R_t(u)]$$

When η is concentrated on state s we write with some abuse of notation $F_\eta(v, u) = F_s(v, u)$, and when η is fixed, the index is omitted. When the system is in its **stationary state**, the global expected fitness simplifies to:

$$F_\eta(v, u) = \sum_{t=0}^{\infty} \mathbb{E}_{\eta, v}[R_t(u)] = \sum_{(s, a)} \frac{f_{\eta, v}(s, a)}{T_{\eta, u}} \sum_{(s', a')} f_{\eta, u}(s', a') r(s, a, s', a') \quad (4.5)$$

Remark 3. Equation (4.5) would not hold if the policy of a player could depend on the absolute time or on the behavior (i.e. on the actions) of other players. On the other hand, since players are not distinguishable, and since the lifetime distribution of a mobile depends only on its local time, we may expect equation (4.5) to hold.

Two policies v and v' are said to be **equivalent** if the corresponding occupation measures coincide. We shall write $v =_e v'$. Note that if v and v' are equivalent policies for a given player, then for any policy u used by the rest of the population, the fitness expressed by equation (4.5) under v and under v' are the same. Indeed, the fitness depends only of the policy used through the occupation measures.

By considering the expression (4.5) for the fitness, Altman and Hayel (Altman and Hayel, 2010) define the ESS in a MDEG framework. The authors further define the modified global expected fitness function:

$$\tilde{F}_\eta(v, u) = \sum_{(s, a)} f_{\eta, v}(s, a) \sum_{(s', a')} f_{\eta, u}(s', a') r(s, a, s', a'),$$

which is bilinear in the occupation measures. They then define the *weak* ESS, related to the notion of equivalent class in terms of occupation measure. They show how to transform a MDEG into a standard evolutionary game and to compute the equilibria.

While Altman and Hayel focus on a static definition of the ESS in MDEG framework, in the next chapter we consider MDEG from a dynamical point of view. We associate

each player to continuous time MDP and we model the interdependent dynamics of individual states and policies.

In Chapter 6 we study a simple Hawk-Dove game in MDEG framework, where, in contrast to what we presented here, the transition probabilities of a player may depend on the action of the opponent it interacts with and not on its action.

State Policy Couple Dynamics

“We repeat most emphatically that our theory is thoroughly static. A dynamic theory would unquestionably be more complete and preferable.”

Von Neumann and Morgenstern, Theory of Games and Economic Behavior,
1953

Summary

In this chapter, we present a new dynamical approach to MDEG, where individual states are supposed to evolve in time according to a continuous time MDP. We introduce the concept of interdependent dynamics of states and policies in a general framework, and we then define the State Policy coupled Dynamics (SPcD) in a simple case, in order to find closed-form solutions and to establish the relation between the equilibria of the defined system of differential equations and the equilibria of the game. These solutions are obtained by assuming that the processes of states and policies move with different velocities, which allows us to apply two different approximation techniques: the singular perturbation method, and a matrix approach.

5.1 Introduction

In the previous chapter we presented MDEG models, introduced by Altman and Hayel ([Altman and Hayel, 2010](#)), where each player is associated with an individual state. During its finite life time the player meets several times other users through random pairwise interactions and it may move among different states. The actions played by an individual determine its immediate fitness and the transition probabilities to the next state. The evolution of states is thus described by a discrete time MDP. In contrast with Altman and Hayel, who transform the problem into an equivalent standard evolutionary game and look for its ESS, we study here the dynamics involved in the game. We consider both the local dynamics of individual states and the dynamics intrinsically related to the global evolution of the distribution of policies in the population, which are described by interdependent differential equations. Individuals are thus assumed to control a CT-MDP instead of a discrete time one. We give some general results about the convergence of the coupled dynamical system to an equilibrium of the population game. We deepen the analysis of the convergence of the coupled dynamics in a particular case with two states and two actions, for which we propose two different approaches to the problem.

State of Art

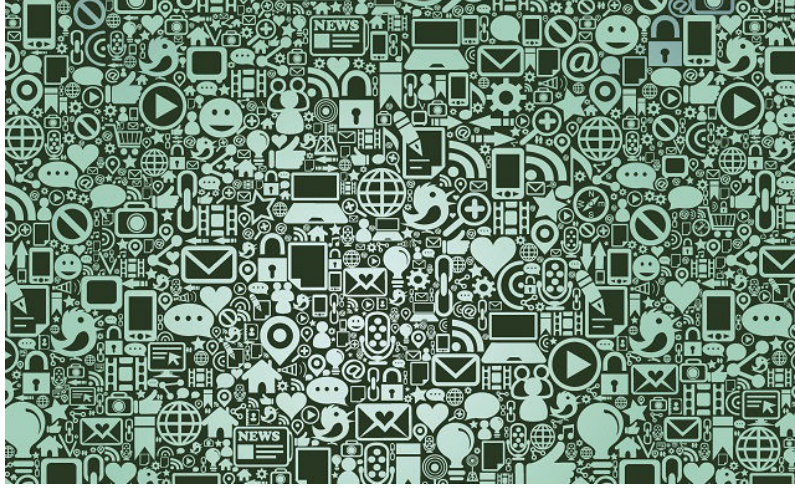
Evolutionary games can be interpreted as a branch of dynamical systems through the replicator equations (see e.g. ([Hofbauer and Sigmund, 1998](#))), which represents an important foundations for understanding individuals' behaviors in a population. In the last years the relationship between evolutionary dynamics and *learning algorithms* has been investigated. In ([Borgers and Sarin, 1997](#)), for example, the authors analyze the "Cross' learning process", which is a simple learning model where a player updates its action only on the basis of its own action and on the fitness it gets (and not on the others' actions). They prove that, in a particular case, this dynamics converges to the replicator dynamics. The authors also provide a discussion on the relation between learning and evolution, comparing the interpretation of the stochastic game underlying Cross' learning process model, and the deterministic population game underlying the replicator dynamics. In ([Beggs, 2005](#)), the author investigates the convergence of fitness and strategies in another similar model of reinforcement learning, the "Erev and Roth's model". The author shows that in a two-person constant-sum game, when both players learn according to Erev and Roth's rule, the long-run behavior of the system is related to a system of equation similar to the 'adjusted replicator dynamic' introduced

in (Maynard Smith, 1982).

In general, a learning procedure describes how each individual adapts its action based on the information it has, like its own fitness, average fitness, historical actions of the others, probabilistic beliefs on the other actions, etc. An analysis of level of information under learning processes in games is given in (Fudenberg and Levine, 1998). As already mentioned in Section 1.3, many different learning algorithms/dynamics have been proposed in the game theory literature, like Brown-Nash-Von Neumann, logit dynamics, etc. All these dynamics can be generalized to the notion of *revision protocols* (Sandholm, 2010) which define a general rule (“compare and innovate”, “target and innovate”, “compare and non-innovate”, etc.) followed by individuals (see Section 1.3 for more details).

Motivations and Applications

Our model finds its first motivation in the study of optimal power control policies in wireless networks (Altman and Hayel, 2008) (see the Introduction to the previous chapter). In Information and Communications Technology, it can be used to study social networks applications, crowd sourcing and Internet of Things (IoT). Emerging applications in engineering such as crowd-sourcing and (mis)information propagation, for example, involve a large population in a complex network of heterogeneous users or agents, who strategically make dynamic decisions. These agents interact with each other in a complex environment, in which each individual takes strategic and dynamic decisions in response to the agents it interacts with. In all these applications, the action set of each agent depends on a local state. In social networks, for example, each agent may decide to add/remove friends/news based on its own current status. Its decision impacts its own status dynamics but also the interaction with other agents. In IoT, a sensor has to determine when to upload its information to the fusion center. This decision impacts its battery level but also the communication quality as collisions may occur for example. As pointed out in several references cited above, the replicator dynamics equations are related to several learning algorithms that can be implemented in such sensors or actuators in IoT. Then, by studying these equations, we can understand the convergence behavior of decentralized algorithms that can be used in such applications.



This chapter is organized as follows: in Section 5.2 we introduce our evolutionary game model that takes into account an individual state dynamics coupled to the policies ones. A complete characterization of the coupled dynamical system is performed in Section 5.3, for a particular two states and two pure actions game. By assuming that the two dynamic processes evolve with different velocities, in Subsection 5.3.6 and 5.3.7 we find the equilibria of the game through two different techniques: the singular perturbation method and by rewriting the problem as a matrix game. We then compare the two solutions obtained and we prove that they are equivalent in terms of fitness and average sojourn time. Some applications in network systems are proposed in Section 5.3.9.

5.2 General Model

5.2.1 The Individual State Dynamical Model

We consider here a population game in which each individual controls a **Continuous Time Markov Decision Process (CT-MDP)** (Guo and Hernandez-Lerma, 2009). Let \mathcal{S} be the finite individual state space of players, with $|\mathcal{S}| = K$. We suppose that each player disposes of the same finite set of actions, \mathcal{A} , with $|\mathcal{A}| = M$. CT-MDP are defined through the set of **transition rates** \mathcal{R} , which describe the rate of the process' transitions from one state to another. More precisely, we denote by $\mathcal{R}_s(s', a)$ the transition rate from state s' to state s given action a , which satisfies $\mathcal{R}_s(s', a) \geq 0$ for all $s' \in \mathcal{S}$, $s' \neq s$, and $a \in \mathcal{A}$. These quantities are *conservative*, i.e.

$$\forall s' \in \mathcal{S}, \forall a \in \mathcal{A}, \quad \sum_s \mathcal{R}_s(s', a) = 0.$$

Also the transition rates are *stable*, i.e.

$$\sup_{a \in \mathcal{A}} \mathcal{R}_{s'}(a) < \infty, \quad \forall s' \in \mathcal{S}$$

with $\mathcal{R}_{s'}(a) := -\mathcal{R}_{s'}(s', a) \geq 0$. Since the set of actions is finite, the stability of the transition rates is guaranteed. Note that the transition rates of a player only depend on its actions, and not on the others' ones.

In standard evolutionary games, each individual plays a pure action, whereas in our framework, individuals choose a deterministic policy in the finite set $\mathcal{U}_D = \{u_1, \dots, u_D\}$ (see the definition in Section 4.2 for the definition of \mathcal{U}_D).

The choice of a policy determines the time spent by each individual in each state. Indeed, for any state s' and action a , the sojourn time in state s' is a random variable which follows an exponential distribution with parameter $\mathcal{R}_{s'}(a) = \sum_{s \neq s'} \mathcal{R}_s(s', a)$. Then, under a given deterministic policy $u_j \in \mathcal{U}_D$, as there is a unique action $a = u_j(s')$ associated with each state s' , such that $u_j(a|s') = 1$, the time spent in any state s' for any individual choosing this policy, follows an exponential distribution with parameter $\mathcal{R}_{s'}(a) = \mathcal{R}_{s'}(u_j(s'))$. This implies that the fraction of individuals in a given state depends on the distribution of policies over the population. This fraction is important to define the fitness obtained for each individual at each pairwise interaction. Given a deterministic policy $u_j \in \mathcal{U}_D$ and a state $s \in \mathcal{S}$, we denote by $T_s(u_j)$ the **average sojourn time** that an individual playing policy u_j spends in state s , which is given by:

$$T_s(u_j) = \frac{1}{\mathcal{R}_s(u_j(s))}.$$

Note that a general sojourn time distribution could also be considered, but we consider the exponential distribution, in order to obtain closed-form solutions of the equilibrium Markovian policy of the defined game.

5.2.2 Population of Players

Let us consider a fixed population of N players, where each one controls a CT-MDP, with \mathcal{S} , \mathcal{A} and \mathcal{R} respectively the finite set of states, actions, and transition rates. Let $\Delta(\mathcal{A})$ be the set of distributions over \mathcal{A} . We define the proportion of individuals (decision makers) that are in state $s \in \mathcal{S}$ at time t as $\mathbf{w}_s^N(t) := \frac{1}{N} \sum_{l=1}^N \mathbb{1}_{\{s^l(t)=s\}}$, where $\mathbb{1}$ is the indicator function, i.e. $\mathbb{1}_{\{s^l(t)=s\}} = 1$ if the state of player l at time t is s , and it equals

zero otherwise. For each state s , we denote by $Y_s^l(t)$ the probability that an individual l is in state s at time t under a deterministic policy $u \in \mathcal{U}_D$, i.e. $Y_s^l(t) = \mathbb{P}_{s_0}^u(s^l(t) = s)$, where s_0 is the initial state. Then, from Assumption 1, $\mathbb{P}_{s_0}^u$ does not depend on s_0 and, from the law of large numbers, when the size of the population grows to infinity, then $Y_s^l(t)$ can be approximated by the proportion of individuals in state s at time t , given by $\mathbf{w}_s(t) = \lim_{N \rightarrow \infty} \mathbf{w}_s^N(t)$. This means that the individual state dynamics, corresponds to the dynamics of the proportion of individuals in state s in the global population. Let $\mathcal{S} = \{s_1, \dots, s_K\}$ and $w_i = w_{s_i}$. We further suppose that the individual dynamics also depends on the policies and that, for any state $s_i \in \mathcal{S}$, there exists a Lipschitz function h_i which describes the dynamics of w_i as follows:

$$\dot{w}_i(t) = h_i(\mathbf{w}(t), \mathbf{q}(t)), \quad \forall s_i \in \mathcal{S}, \quad (5.1)$$

where $\mathbf{w}(t) = (w_1(t), \dots, w_K(t))$ is the vector of state probabilities and $\mathbf{q}(t) = (q_1(t), \dots, q_D(t))$ is the vector of distribution over the deterministic policies in the population, such that $q_j(t)$ indicates the proportion of individuals playing deterministic policy $u_j \in \mathcal{U}_D = \{u_1, u_2, \dots, u_D\}$ at time t . Indeed, we have seen previously that the choice of a such policy has an impact on the state dynamics of this individual.

5.2.3 Policies Dynamics

We assume that the proportion of individuals choosing each deterministic policy is evolving over time as a dynamical process. Without specifying any revision protocol, we define the dynamics of policies through a set of Lipschitz continuous functions $G := \{g_1, \dots, g_D\}$, such that:

$$\dot{q}_j(t) = g_j(\mathbf{w}(t), \mathbf{q}(t)) \quad \forall u_j \in \mathcal{U}_D. \quad (5.2)$$

Then, the dynamical evolution of states and policies distributions in the population is represented by a system of $K + D$ equations:

$$\begin{cases} \dot{w}_1 = h_1(\mathbf{w}(t), \mathbf{q}(t)) \\ \vdots \\ \dot{w}_K = h_K(\mathbf{w}(t), \mathbf{w}(t)) \\ \dot{q}_1(t) = g_1(\mathbf{w}(t), \mathbf{q}(t)) \\ \vdots \\ \dot{q}_D(t) = g_D(\mathbf{w}(t), \mathbf{q}(t)) \end{cases} \quad (5.3)$$

5.2.4 Two Time-Scales Behavior

We assume that the state and the policy dynamics move with different velocities. By supposing that the individual state dynamics move much faster than the slow updating policies processes, we can consider the singular perturbation model (Kokotovic et al., 1986), briefly presented in what follows, to find the rest points of the system (5.3). We first introduce Assumption 1, which assures that, under any deterministic policy, there exists a stationary distribution over the individual states which does not depend on the initial state.

Assumption 1. *Under any deterministic policy, the stochastic process of the individual states forms an ergodic Markov chain.*

Singular Perturbation Method

The singular perturbation model of finite dimensional systems has been extensively studied in mathematical literature (see e.g. (Levinson, 1950) (Hoppensteadt, 1967) (O'Malley, 1971)) and it was the first method used in control theory as a tool to simplify dynamic models. Following Kokotovic's approach (Kokotovic et al., 1986), we provide here the basic concepts of singular perturbation asymptotics and time-scale modeling that will be used in the next section. Consider the interdependent dynamics:

$$\dot{\mathbf{x}} = g(\mathbf{x}, \mathbf{z}, \mathbf{u}, \epsilon, t), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad \mathbf{x} \in \mathbb{R}^n, \quad (5.4)$$

$$\epsilon \dot{\mathbf{z}} = h(\mathbf{x}, \mathbf{z}, \mathbf{u}, \epsilon, t), \quad \mathbf{z}(t_0) = \mathbf{z}_0, \quad \mathbf{z} \in \mathbb{R}^m, \quad (5.5)$$

where \mathbf{u} is the control vector, ϵ is a small scalar and g and h are sufficiently many times continuously differentiable functions of their arguments. To reduce the order of the system (5.4)-(5.5), the parameter ϵ is perturbed: when ϵ equals zero, the order of the system reduces from $n + m$ to n , since equation (5.5) degenerates into the *transcendental equation*

$$0 = g(\bar{\mathbf{x}}, \bar{\mathbf{z}}, \bar{\mathbf{u}}, 0, t). \quad (5.6)$$

The system is said to be in *normal* or *standard form* if and only if the following assumption is satisfied:

Assumption 2. *In a domain of interest, equation (5.6) has $k \geq 1$ distinct roots:*

$$\bar{\mathbf{z}} = \varphi_i(\bar{\mathbf{x}}, \bar{\mathbf{u}}, t), \quad i = 1, 2, \dots, k \quad (5.7)$$

This assumption assures that a well defined (n-dimensional) model correspond to each root. To obtain the i reduced model, one need to substitute $\bar{\mathbf{z}}$ into (5.4) to get the *quasi-steady-state* model:

$$\dot{\bar{\mathbf{x}}} = g(\bar{\mathbf{x}}, \varphi_i(\bar{\mathbf{x}}, \bar{\mathbf{u}}, t), \bar{\mathbf{u}}, 0, t) = g(\bar{\mathbf{x}}, \bar{\mathbf{u}}, t). \quad (5.8)$$

The velocity of \mathbf{z} , given by g/ϵ , is large when ϵ is small and thus \mathbf{z} may rapidly converge to a root of (5.6), which is the quasi-steady-state of (5.5). Singular perturbation thus generates a two time-scale behavior of the dynamic system (5.4)-(5.5): the slow process is approximated by the reduced model (5.8), while the discrepancy between the original system and (5.8) is the fast transient. We now investigate the relation between the original variable \mathbf{z} and $\bar{\mathbf{z}}$. Note that the initial value of $\bar{\mathbf{z}}$ may be far from \mathbf{z}_0 , as it is given by

$$\bar{\mathbf{z}}(t_0) = \varphi(\bar{\mathbf{x}}(t_0), \bar{\mathbf{u}}(t_0), t_0) \neq \mathbf{z}_0.$$

We can expect that \mathbf{z} is well approximated by $\bar{\mathbf{z}}$ in a sub-interval $[t_1, T]$, where $t_1 > t_0$, i.e.:

$$\mathbf{z} = \bar{\mathbf{z}}(t) + O(\epsilon), \quad \forall t \in [t_1, T]. \quad (5.9)$$

However, it is possible to constrain the quasi-steady state $\bar{\mathbf{x}}$ to start from x_0 to have:

$$\mathbf{x} = \bar{\mathbf{x}}(t) + O(\epsilon), \quad \forall t \in [t_0, T]. \quad (5.10)$$

We introduce a new time variable to study the behavior of the variable \mathbf{z} . Let

$$\epsilon \frac{dz}{dt} = \frac{dz}{d\tau} \Rightarrow \frac{d\tau}{dt} = \frac{1}{\epsilon},$$

and $\tau = 0$ when $t = t_0$. The new time variable

$$\tau = \frac{t - t_0}{\epsilon}, \quad \tau = 0 \text{ at } t = t_0$$

is stretched, that is, τ goes to infinity when ϵ goes to zero. We define the *boundary layer system* which describes \mathbf{z} as a function of τ :

$$\frac{d\hat{\mathbf{z}}}{d\tau} = g(x_0, \hat{\mathbf{z}}(\tau), u, 0, t_0), \quad \hat{\mathbf{z}}(0) = \mathbf{z}_0, \quad (5.11)$$

where x_0 and t_0 are fixed parameters. We thus obtain that:

$$\mathbf{z} = \bar{\mathbf{z}}(t) + \hat{\mathbf{z}}(\tau) - \bar{\mathbf{z}}(t_0) + O(\epsilon), \quad (5.12)$$

where $\bar{\mathbf{z}}(t)$ is the slow and $\hat{\mathbf{z}}(\tau) - \bar{\mathbf{z}}(t_0)$ the fast transient of \mathbf{z} . Two assumptions are needed to guarantees the validity of the approximations (5.9) and (5.12).

Assumption 3. The equilibrium $\bar{\mathbf{z}}(t_0)$ of (5.11) is asymptotically stable uniformly in x_0 and t_0 , and \mathbf{z}_0 belongs to its domain of attraction, so $\hat{\mathbf{z}}(\tau)$ exists for $\tau \geq 0$.

This implies that $\lim_{\tau \rightarrow \infty} \hat{\mathbf{z}}(\tau) = \bar{\mathbf{z}}(t_0)$ uniformly in x_0 and t_0 and thus \mathbf{z} is close to its quasi-steady-state $\bar{\mathbf{z}}$ for $t_1 > t_0$. The interval $[t_0, t_1]$ can be made arbitrarily short by making ϵ sufficiently small.

Assumption 4. The eigenvalues of $\partial g / \partial \mathbf{z}$ evaluated along $\bar{\mathbf{x}}(t)$, $\bar{\mathbf{z}}(t)$, $\bar{\mathbf{u}}(t)$ for all $t \in [t_0, T]$ have real parts smaller than a fixed negative number, i.e.:

$$\Re \lambda \left\{ \frac{\partial g}{\partial \mathbf{z}} \right\} \leq -c < 0.$$

Note that, if \mathbf{z}_0 is assumed to be sufficiently close to $\bar{\mathbf{z}}(t_0)$, then Assumption 4 is stronger than Assumption 3. Furthermore, Assumption 4 implies that the root $\bar{\mathbf{z}}(t)$ is distinct as required by Assumption 2. We can now state the main result, often called *Tikhonov theorem*.

Theorem 4. [Theorem 3.1 (Kokotovic et al., 1986)] If Assumptions 3 and 4 are satisfied, then the approximations (5.10) and (5.12) are valid for all $t \in [t_0, T]$, and there exists a t_1 such that (5.9) is valid for all $t \in [t_1, T]$.

For the proof of this result, see e.g. (Levinson, 1950), (Hoppensteadt, 1967).

Fast state and slow policy processes

In order to describe the two time-scales behavior of the system (5.3), we introduce the small positive parameter $\epsilon > 0$:

$$\begin{cases} \epsilon \dot{w}_1 = h_1(\mathbf{w}(t), \mathbf{q}(t)) \\ \vdots \\ \epsilon \dot{w}_K(t) = h_K(\mathbf{w}(t), \mathbf{q}(t)) \\ \dot{q}_1(t) = g_1(\mathbf{w}(t), \mathbf{q}(t)) \\ \vdots \\ \dot{q}_D(t) = g_D(\mathbf{w}(t), \mathbf{q}(t)) \end{cases} \quad (5.13)$$

Then, the velocity of the state process, $\dot{w}_i = h_i(\mathbf{w}, \mathbf{q}) / \epsilon$, is fast when ϵ is small, which means that the states dynamics may rapidly converge to its steady-state. From the singular perturbation theory, if the assumptions introduced above are satisfied, one can solve the reduced model and easily find a good approximation of the solution of the original system (5.3). When supposing the two time-scales behavior of the states and

policies dynamic system, an alternative technique to find a solution of (5.3) consists in considering the stationary distribution of states and then solve the policies dynamics. Let $r(s, a; s', a')$ be the immediate fitness that a player gets when it is in state s and plays action a against an individual in state s' using action a' . When the distribution of the individual states is stationary, we denote by $F(u_i, u_j)$ the immediate expected fitness of a player using deterministic policy $u_i \in \mathcal{U}_D$ against a population playing $u_j \in \mathcal{U}_D$. It can be defined as a function of the average sojourn times as follows:

$$F(u_i, u_j) = \sum_{s, s' \in \mathcal{S}} \frac{T_s(u_i)}{\sum_{s'} T_{s'}(u_i)} r(s, u_i(s); s', u_j(s')) \frac{T_{s'}(u_j)}{\sum_{s'} T_{s'}(u_j)}. \quad (5.14)$$

Remark 4. In discrete time MDEG (see Chapter 4) introduced in (Altman and Hayel, 2010), stationary distributions can be computed considering occupation measures (see Section 4.2.1), whereas in our continuous time setting we deal with average sojourn time expressions. It is possible to establish a relation between discrete and continuous time MDEG by transforming the CT-MDP into an equivalent discrete time MDP through the uniformization technique (see e.g. (Puterman, 1994) (Ibe, 2008)), which makes the total transition rate from a state the same for all states and allows to consider the discrete time MDP embedded at transition epochs of each event. But, since we focus here on the replicator dynamics and not on the equilibria of the game, we study the MDEG in continuous time, where each individuals control their transition rates.

Once we have defined the expected fitness function, we assume that each individual chooses its deterministic policy in order to maximize its immediate expected fitness and we then define a stationary equilibrium u^* policy for our continuous-time MDEG problem.

Definition 6. Under the Assumption 1 and assuming that the distribution over the states is time homogeneous, a policy $u^* \in \mathcal{U}_D$ is a **deterministic equilibrium policy** if, $\forall u_j \in \mathcal{U}_D$:

$$F(u^*, u^*) \geq F(u_j, u^*).$$

Note that, from Assumption 1, if a deterministic policy u^* is optimal, than no player can benefit from unilateral deviations to any policy, including non deterministic ones. Hence, if all population uses an equilibrium among stationary policy, then no player can benefit by a unilateral deviation to any policy. We have seen in Section 1.3.2 that, in standard evolutionary game theory, there is a relation between equilibria of a game and rest points of the replicator dynamics. We can establish a similar relationship in our setting. Given the vector of policies distributions at time t , $\mathbf{q}(t) = (q_1(t), \dots, q_D(t))$, we denote by $F(u_k, \mathbf{q}) := \sum_{u_j \in \mathcal{U}_D} F(u_k, u_j) q_j(t)$ the immediate expected fitness of an

individual using policy $u_k \in \mathcal{U}_D$ against a populations in state \mathbf{q} , and by $\bar{F}(\mathbf{q}) := \sum_{i,j \in \mathcal{U}_D} q_i(t) F(u_i, u_j) q_j(t)$ the average expected fitness of the population.

Based on the expression (5.14), we can define the dynamics of the evolution of $q_k(t)$ over time, where $u_k \in \mathcal{U}_D$:

$$\frac{\dot{q}_k(t)}{q_k(t)} = (F(u_k(t), \mathbf{q}(t)) - \bar{F}(\mathbf{q}(t))). \quad (5.15)$$

Every interior rest point of the replicator dynamics (5.15) is a stationary equilibrium policy for the game we defined. Note that the dynamics proposed in equation (5.15) assumes that the distribution over the individual states is already stationary, as the state dynamics update very quickly compared to the policy dynamics. Getting results about equilibrium policies in the general case may be complicated, as the number of deterministic policies \mathcal{U}_D explodes and it becomes very difficult to properly write a clear mathematical analysis to obtain closed-form solutions of the equilibrium. In the next section, we present a complete analysis and characterization of the equilibrium policy for a particular game with two states and two strategies, considering the coupled dynamical system.

5.3 Complete Characterization of the Game with Two States and Two Actions

5.3.1 Individual State and Its Dynamics

In this section we study a particular case of the state-policy coupled dynamics model. We suppose that each player can be in one of two possible states, with $\mathcal{S} = \{1, 0\}$. Every individual goes through a cycle that starts in state 1 and moves to states 0 after some random time, at a rate that depends on its policy. After some exponentially distributed time it returns to state 1 and so on. At each pairwise interaction, the set of available actions of a player depends on its state: in state 1, $\mathcal{A}_1 = \{x, y\}$, whereas in state 0 an individual can only use y . We consider the set of deterministic policies $\mathcal{U}_D := \{u_x, u_y\}$, where u_x (resp. u_y) is the deterministic policy which consists in always playing action x (resp. y) in state 1. Under both policies, in state 0 an individual plays y . Each player chooses one deterministic policy and we denote by $q_i(t)$ the proportion of individuals in the population that play the deterministic policy u_i at time t . Note that $q_y(t) = 1 - q_x(t)$. We suppose that the policy chosen impacts the fitness of the player interacting with

another individual and also the time it spends in state 1. We define by μ_i the rate of decay from state 1 to state 0 when using policy u_i , $i \in \{x, y\}$, where $\mu_x > \mu_y$, and by μ the rate of change from state 0 to state 1.

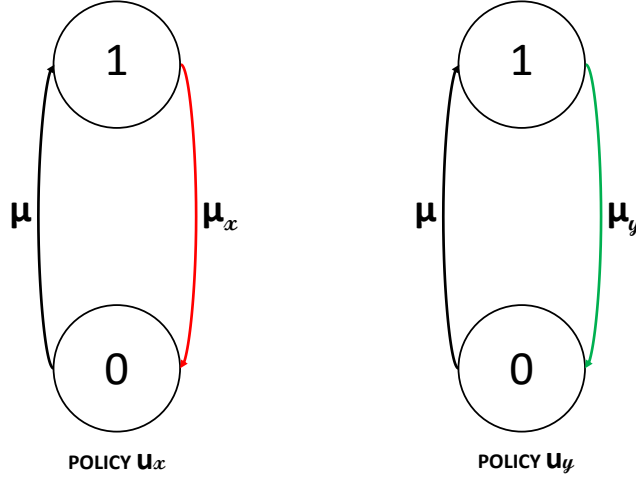


Figure 5.1: A player can choose between policy u_x and u_y . Accordingly with this choice, in state 1 the individual plays, respectively x or y and move to state 0, at a rate μ_x or μ_y . In state 0, under both policies, the only available action is y and the rate at which the player returns to state 1 is μ .

As stated in Section 5.2.2, since the population considered is large, from the law of large numbers, the individual state dynamics can be approximated by the population state dynamics. Let $w_1(t)$ denote the share of individuals in state 1, which is equivalent to the probability that an individual in the population is in state 1 at time t . We define the dynamics of $w_1(t)$ as follows:

$$\dot{w}_1(t) = -\mu_x w_1(t) q_x(t) - \mu_y w_1(t) (1 - q_x(t)) + \mu (1 - w_1(t)). \quad (5.16)$$

The first (resp. the second) term on the right side of the equation indicates that the share of individuals in state 1 which choose policy u_x (resp. u_y) at time t , given by $\mu_x w_1(t)$ (resp. $w_1(t)(1 - q_x(t))$), leaves state 1 at a rate μ_x (resp. μ_y), whereas the third term indicates that individuals in state 0 move to state 1 at rate μ (independently on the policy chosen).

5.3.2 Individual Fitness

At each pairwise interaction, the immediate fitness obtained by an individual, which depends on its current action and the current action of its opponent, is given by the

following fitness bimatrix:

$$A := \begin{matrix} & \begin{matrix} x & y \end{matrix} \\ \begin{matrix} x \\ y \end{matrix} & \begin{pmatrix} (a, a) & (b, c) \\ (c, b) & (d, d) \end{pmatrix} \end{matrix}, \quad (5.17)$$

where x and y are the available actions and the matrix entry A_{ij} indicates the payoff respectively of the first (row) and the second (column) player. The expected fitness of a player interacting at time t , depends on the population profile at time t , which is now expressed by the couple $\zeta(t) := (w_1(t), q_x(t))$. We denote by $r_x(\zeta(t))$ (resp. $r_y(\zeta(t))$) the **expected fitness** of an individual playing pure action $x \in \mathcal{A}$, against a population whose profile is $\zeta(t)$. By considering payoff matrix (5.17), we obtain the following expressions:

$$r_x(\zeta(t)) := w_1(t)(q_x(t)a + (1 - q_x(t))b) + (1 - w_1(t))b,$$

$$r_y(\zeta(t)) := w_1(t)(q_x(t)c + (1 - q_x(t))d) + (1 - w_1(t))d.$$

We can now define the fitness of an individual choosing deterministic policy $u_i \in \mathcal{U}_D$ at time t , denoted by $F_i(\zeta(t))$, $i = x, y$. The fitness $F_i(\zeta(t))$ depends on the population profile $\zeta(t)$. As we are dealing with a large system, from the law of large numbers, we can assume that the probability that any individual is in state 1 at time t equals $w_1(t)$ and it is the same for all players, as explained in Section 5.2.2. Then, an individual choosing policy u_x will be in state 1 (resp. 0) at time t with probability $w_1(t)$ (resp. $1 - w_1(t)$), and it will get an immediate fitness $r_x(\zeta(t))$ (resp. $r_y(\zeta(t))$). The expected immediate fitness of an individual choosing policy u_x at time t is thus given by

$$F_x(\zeta(t)) = w_1(t)r_x(\zeta(t)) + (1 - w_1(t))r_y(\zeta(t)).$$

If an individual plays policy u_y , whatever its state is, it plays pure action y , which leads to:

$$F_y(\zeta(t)) = w_1(t)r_y(\zeta(t)) + (1 - w_1(t))r_y(\zeta(t)) = r_y(\zeta(t)).$$

The average expected fitness of a population whose profile at time t is $\zeta(t) = (w_1(t), q_x(t))$ is

$$\bar{F}(\zeta(t)) = q_x(t)F_x(\zeta(t)) + (1 - q_x(t))F_y(\zeta(t)). \quad (5.18)$$

5.3.3 Equilibrium Profile

We study the properties of stability of the population profile, supposing that individuals play deterministic policies in \mathcal{U}_D . Let $\mathbf{q}^* = (q_x^*, q_y^*)$ and define $\text{supp}(\mathbf{q}^*) = \{u_i \in \mathcal{U}_D | q_i^* > 0, \text{ given } \zeta^*\}$.

Definition 7. A population profile $\xi^* = (w_1^*, q_x^*)$ is an equilibrium profile iff $\forall u_i \in \text{supp}(q^*)$ we have that:

$$F_i(\xi^*) \geq F_j(\xi^*) \quad \forall j \neq i, \quad i, j \in \{x, y\}$$

An equilibrium profile is a particular population profile $\xi^* = (w_1^*, q_x^*)$ which is stable in the sense of robustness against a deviation of the proportion of individuals playing the deterministic policy u_x . In other words, this definition says that no individual has an interest in changing its deterministic policy, considering this population profile. Note that an equilibrium policy, if adopted by the whole population, determines a stationary individual state.

Remark 5. It can be easily proved that if the population profile $\xi^* = (p_1^*, q_x^*)$ satisfies the *indifference principle*, i.e.:

$$F_x(\xi^*) = F_y(\xi^*),$$

then it is an equilibrium profile.

5.3.4 Policy Based Replicator Dynamics

As we consider here policies instead of actions, we introduce a policy based replicator dynamics (PbRD), to study the evolution of the share of individuals using deterministic policy u_x at time t , represented by $q_x(t)$. The PbRD is given by the following equation:

$$\dot{q}_x(t) := q_x(t)(F_x(\xi(t)) - \bar{F}(\xi(t))). \quad (5.19)$$

Then, the growth rate of the population share using the deterministic policy u_x at time t is:

$$\frac{\dot{q}_x(t)}{q_x(t)} = F_x(\xi(t)) - \bar{F}(\xi(t)), \quad (5.20)$$

By substituting (5.18) into (5.19):

$$\begin{aligned} \dot{q}_x(t) &= q_x(t)[F_x(\xi(t)) - q_x(t)F_x(\xi(t)) - (1 - q_x(t))F_y(\xi(t))] \\ &= q_x(t)(1 - q_x(t))(F_x(\xi(t)) - F_y(\xi(t))), \\ &= g(w_1(t), q_x(t)). \end{aligned}$$

We can investigate the dynamics of actions in this framework, where the fitness is a function of the population profile depending on policies and states and establish its relation with the dynamics of policies. If we pick one random individual in the population at time t , the probability that it plays pure action x , denoted by $q(t)$, is given by

the product $q_x(t)w_1(t)$. From this and from definition (5.19), we obtain:

$$\begin{aligned}\dot{q}(t) &= \dot{q}_x(t)w_1(t) + q_x(t)\dot{w}_1(t) = q_x(t)[w_1(t)(F_x(\xi(t)) - \bar{F}(\xi(t))) + \dot{w}_1(t)] \\ &= \frac{q(t)}{w_1(t)}[w_1(t)(F_x(\xi(t)) - \bar{F}(\xi(t))) + \dot{w}_1(t)],\end{aligned}$$

which leads to the following equation for the growth rate of the proportion of individuals playing pure action x in the population at time t :

$$\frac{\dot{q}(t)}{q(t)} = (F_x(\xi(t)) - \bar{F}(\xi(t))) + \frac{\dot{w}_1(t)}{w_1(t)}. \quad (5.21)$$

Equation (5.21) shows how the evolution of states impacts the dynamics of actions in our context. The growth rate of action x is increasing in the growth rate of state 1. We observe that a sufficiently high growth rate of state 1 can lead to a growing rate of action x even if policy u_x is non-optimal.

5.3.5 State-Policy Coupled Dynamics

We define the system of State-Policy Coupled Dynamics (SPcD) which combines the dynamics of the individual state and the dynamics of the policies used in the population:

$$(S) \begin{cases} \dot{w}_1 = h(w_1(t), q_x(t)) \\ \dot{q}_x = g(w_1(t), q_x(t)) \end{cases}$$

where $\xi(t) = (w_1(t), q_x(t))$ is the population profile at time t . Note that, since h and g are continuously differentiable functions of ξ , (i.e. the partial derivatives $\partial h/\partial w_1$, $\partial h/\partial q_x$, $\partial g/\partial w_1$, $\partial g/\partial q_x$ are continuous) they are locally Lipschitz continuous functions with respect to ξ in the compact space $[0, 1]^2$, which guarantees the existence of a solution of the system (S).

The rest point of the SPcD is $\xi^* = (w_1^*, q_x^*)$ satisfying:

$$\begin{cases} h(\xi^*) = 0 \\ g(\xi^*) = 0. \end{cases} \quad (5.22)$$

Lemma 1. *Any interior rest point of the SPcD (S) is an equilibrium profile of the state-policy game.*

Proof. Trivially, if ξ^* is internal, it satisfies the indifference principle $F_x(\xi^*) = F_y(\xi^*)$, so ξ^* is an equilibrium profile. \square

Remark 6. Note that the converse does not necessarily hold. Any equilibrium profile is a rest point of the PbRD in (5.19), but it's not necessarily a rest point of the individual state dynamics.

Lemma 2. Any Lyapunov stable rest point of the SPcD (S) is an equilibrium profile of the state-policy game.

Proof. Suppose that $\xi^* = (w_1^*, q_x^*)$ is a stable rest point. If $\text{supp}(q^*) = \{u_x, u_y\}$ then, ξ^* is an interior rest point and by Lemma 1, ξ^* is an equilibrium profile. Let $\text{supp}(q^*) = \{u_x\}$ and let us suppose that ξ^* is not an equilibrium. This implies that $F_y(\xi^*) > \bar{F}(\xi^*)$, and, from the continuity of the fitness function, there exists a neighborhood U_{ξ^*} of ξ^* such that $\forall \xi \in U_{\xi^*}, \xi \neq \xi^*, F_y(\xi) > \bar{F}(\xi)$. This implies that, for this profile, the component q_y increases exponentially, which contradicts the Lyapunov stability of ξ^* . The proof is analogous if $\text{supp}(q^*) = \{u_y\}$, which complete the proof of the lemma. \square

5.3.6 Singular Perturbations Approximation Method

As introduced in Section 5.2, we consider the existence of a small parameter $\epsilon > 0$, such that:

$$\epsilon \dot{w}_1 := h(w_1, q_x).$$

We then rewrite the system of the two coupled differential equations as follows:

$$(S_\epsilon) \begin{cases} \epsilon \dot{w}_1 = h(w_1, q_x), \\ \dot{q}_x = g(w_1, q_x). \end{cases}$$

We can thus approximate the solution of (S_ϵ) using the standard singular Perturbation Model (Kokotovic et al., 1986) introduced in the previous section. The quasi-steady-state-model is obtained by first solving in w_1 the transcendental equation $0 = h(w_1, q_x)$ and then rewriting the differential equation \dot{q} as a function of the obtained roots. As the transcendental equation has a unique real solution $\bar{w}_1 := \pi_1(q_x)$, our system satisfies Assumption 2 and we can thus solve the quasi-steady-state equation:

$$\dot{q}_x = g(\pi_1(q_x), q_x). \quad (5.23)$$

If the Assumption 4 is satisfied, the reduced model is a good approximation of the original system. In our case, this assumption simplifies to the following condition: $\frac{\partial h}{\partial w_1}(w_1, q_x) < 0$. Since $\frac{\partial h}{\partial w_1}(w_1, q_x) = -\mu_x q_x - \mu_y(1 - q_x) - \mu < 0$, we can apply the singular perturbation method to solve (S_ϵ) .

The two-time-scale behavior of $w_1(t)$ and $q_x(t)$ also has a geometric interpretation, as

trajectories in \mathbb{R}^2 . If we define the manifold sets $M_\epsilon := \{\varphi \text{ s.t. } w_1 = \varphi(q_x, \epsilon) \text{ \& } \epsilon = h(q_x, \varphi(q_x, \epsilon))\}$, it is possible to rewrite the problem in terms of invariant manifolds. When the parameter $\epsilon = 0$, then M_0 is an **equilibrium manifold** which corresponds to the quasi steady state model. As Assumption 4 is satisfied, the equilibrium manifold M_0 is stable (attractive) (Kokotovic et al., 1986). An important result states that the existence of a conditionally stable manifold M_0 for $\epsilon = 0$ guarantees the existence of an invariant manifold M_ϵ satisfying the following convergence for all $\epsilon \in [0, \epsilon^*]$:

$$\varphi(\epsilon, q_x) \rightarrow \varphi(0, q_x), \quad \text{and} \quad M_\epsilon \rightarrow M_0 \quad \text{as} \quad \epsilon \rightarrow 0.$$

The positive constant ϵ^* is determined by imposing the **manifold condition**:

$$\epsilon \frac{\partial \varphi}{\partial x} g(\varphi(q_x, \epsilon), q_x) = h(\varphi(q_x, \epsilon), q_x),$$

for all q_x and $\epsilon \in [0, \epsilon^*]$. We illustrate in Figure 5.2 the attractiveness of the slow manifold M_0 for a numerical example.

Let us now compute the solution of the system (S_0) , i.e. the system obtained with $\epsilon = 0$. We suppose that the distribution of individual states is stationary (expressed by Equation (5.16)). By imposing $\dot{w}_1 = 0$, we obtain the following slow manifold $M_0 := \{\varphi \text{ s.t. } w_1 = \varphi(q_x, 0) \text{ \& } 0 = h(q_x, \varphi(q_x, 0))\}$:

$$\varphi(q_x, 0) = \frac{\mu}{\mu + \mu_x q_x + \mu_y (1 - q_x)} := \varphi_1(q_x). \quad (5.24)$$

We can now rewrite the PbRE (5.19) as:

$$\dot{q}_x(t) = q_x(t)(1 - q_x(t)) [F_x(\pi_1(q_x(t)), q_x(t)) - F_y(\pi_1(q_x(t)), q_x(t))].$$

Proposition 10. *For ϵ sufficiently small, the solution of the system (S_ϵ) can be approximated by the solution of S_0 . This is given by the population profile $\zeta^* = (w_1^*, q_x^*)$, such that:*

$$w_1^* = \frac{\mu - s^*(\mu_x - \mu_y)}{\mu + \mu_y} \quad \text{and} \quad q_x^* = \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)}, \quad (5.25)$$

where s^* is the equilibrium of the replicator dynamics (1.6) for the standard evolutionary game whit payoff matrix given by (5.17):

$$s^* = \frac{d - b}{\Delta} \quad \text{with} \quad \Delta = a - b - c + d.$$

Proof. Let us first study the equation $\dot{q}_x = 0$ before substituting the stationary equation of the state dynamics. To solve this equation is equivalent to find the population profile

$\xi = (\pi_1, q_x)$ such that:

$$F_x(\pi_1, q_x) = F_y(\pi_1, q_x).$$

By replacing the expressions of the fitness (5.18) and (5.3.2) in the latter equality we get

$$\begin{aligned} & \pi_1 [\pi_1(q_x a + (1 - q_x)b) + (1 - \pi_1)b] + (1 - \pi_1) - [\pi_1(q_x c + (1 - q_x)d) + (1 - \pi_1)d] \\ &= \pi_1(q_x c + (1 - q_x)d) + (1 - \pi_1)d. \end{aligned}$$

After some manipulations,

$$\pi_1 a \pi_1 q_x + \pi_1 b (1 - \pi_1 q_x) + (1 - \pi_1) c \pi_1 q_x + (1 - \pi_1) d (1 - \pi_1 q_x) = c \pi_1 q_x + d (1 - \pi_1 q_x).$$

This equality corresponds to

$$\pi_1 q_x [\pi_1 a - \pi_1 b + (1 - \pi_1) c - (1 - \pi_1) d - c + d] = d - \pi_1 b - (1 - \pi_1) d.$$

Thus

$$\pi_1 q_x [\pi_1 a - \pi_1 b - \pi_1 c + \pi_1 d] = \pi_1 d - \pi_1 b.$$

We finally obtain

$$\pi_1 q_x = \frac{d - b}{\Delta} := s^*.$$

The stationary condition of the first differential equation (5.16) leads to the following relation between w_1 and q_x :

$$w_1 = \pi_1(q_x) = \frac{\mu}{\mu + \mu_x q_x + \mu_y (1 - q_x)},$$

then we have to solve now: $\pi_1(q_x) q_x = s^*$. The latter equation is equivalent to:

$$\frac{\mu q_x}{\mu + \mu_x q_x + \mu_y (1 - q_x)} = s^*.$$

After some simple manipulations we obtain

$$q_x = \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)} := q_x^*.$$

Finally, as we have $\pi_1(q_x) q_x = s^*$ then $w_1^* := \pi_1(q_x^*) = \frac{s^*}{q_x^*}$ which leads to

$$w_1^* = \frac{\mu - s^*(\mu_x - \mu_y)}{\mu + \mu_y}.$$

□

Note that the rest point q_x^* of the PbRE (5.19) verifies the following relation:

$$q_x^* \pi_1(q_x^*) = s^*.$$

This result says that the equilibrium probability that any individual picked out randomly in the population is playing action x , is equal to s^* . This value is the mixed equilibrium of the standard matrix game given by matrix A . It means that, if we consider a state dependent action game, the equilibrium is obtained under conditional probability over the state.

We have the following necessary and sufficient condition under which the solution obtained is a strict interior point.

Lemma 3. *The solution q_x^* obtained in proposition (10) is a strict interior point if and only if:*

$$\mu > \mu_x \frac{s^*}{1 - s^*}.$$

Proof. The solution obtained in proposition (10) is:

$$q_x^* = \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)}.$$

This solution is a strict interior point if and only if:

$$0 < q_x^* < 1.$$

First, let's look at the positivity condition $q_x^* > 0$. This is equivalent to:

$$0 < q_x^* \iff \mu > s^*(\mu_x - \mu_y).$$

After some basic algebras, the second condition is:

$$q_x^* < 1 \iff \mu > \mu_x \frac{s^*}{1 - s^*}.$$

We have clearly that for all $s^* \in]0, 1[$, μ_x and μ_y :

$$\frac{s^*}{1 - s^*} \mu_x > s^* \mu_x > s^*(\mu_x - \mu_y).$$

Then if $\mu > \mu_x \frac{s^*}{1 - s^*}$ the solution is a strict interior point, and the converse is true. This

concludes the proof. \square

Note that this condition does not depend on the rate μ_y .

Numerical Illustration

We illustrate the theoretical results obtained in this section through a numerical example. We fix the values of the transition rates, $\mu = 10$, $\mu_x = 1.5$ and $\mu_y = 1$, and the payoffs of the matrix game: $a = -0.3$, $c = 0$, $b = 1$ and $d = 0.5$. These values yield to the following equilibrium of the standard evolutionary game: $s^* = \frac{5}{8} = 0.625$.

In figure 5.2 we plot the trajectories of the system (S_ϵ) of the coupled differential equations for different initial conditions and for $\epsilon = 0.01$. We simulate a discrete time version of the differential equations. We plot also the invariant manifold M_0 and we observe that it is an attractor of the trajectories.

Proposition (10) gives the following solution of the system (S_0) :

$$q_x^* = 0.7097, \quad \text{and} \quad w_1^* = 0.8807.$$

This couple corresponds exactly to the attractor of the trajectories on figure 5.2 and then our simulation validates the result of this proposition.

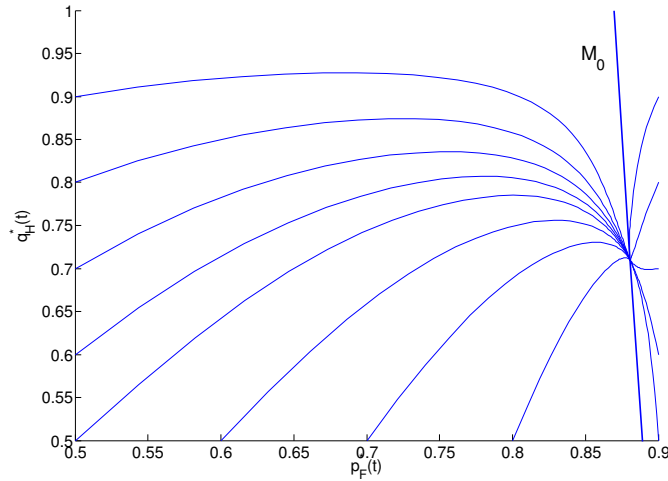


Figure 5.2: Trajectories of the system (S_ϵ) from different starting points and the slow manifold M_0 for $\epsilon = 0.01$.

In the next section, we present an alternative method based on rewriting our game into a matrix game considering only pure policies.

5.3.7 Matrix Game Approximation Technique

In alternative to the singular perturbation method, when assuming the two time-scales behavior of the SPcD system, one can consider the distribution over individual states to be stationary and then solve the obtained standard normal form game. In particular, by following the approach used in (Altman and Hayel, 2010), we can rewrite the two states and two actions model as a matrix game, where individuals play deterministic policies instead of actions. We get the following payoff bimatrix:

$$\begin{array}{cc} & \begin{array}{c} u_y \\ u_x \end{array} \\ \begin{array}{c} u_y \\ u_x \end{array} & \begin{pmatrix} (F(u_y, u_y), F(u_y, u_y)) & (F(u_y, u_x), F(u_x, u_y)) \\ (F(u_x, u_y), F(u_y, u_x)) & (F(u_x, u_x), F(u_x, u_x)) \end{pmatrix} \end{array} \quad (5.26)$$

where $F(u_i, u_j)$ is the expected fitness of an individual playing pure policy u_i against an individual using u_j , with $i, j \in \{x, y\}$. Note that, as showed in (Altman and Hayel, 2010), it is possible to apply this matrix game approach to the general model with M actions and K states but, in this case, the size of the matrix would be of dimension $D \times D$, where D is the number of deterministic policies. We restrict here our analysis to the two actions/states case in order to compare the equilibrium obtained with the two approaches.

The stationary distributions in states 1 and 0 are given respectively by the following average sojourn times:

$$T_1(i) = \frac{\frac{1}{\mu_i}}{\frac{1}{\mu} + \frac{1}{\mu_i}} = \frac{\mu}{\mu + \mu_i},$$

$$T_0(i) = \frac{\frac{1}{\mu}}{\frac{1}{\mu} + \frac{1}{\mu_i}} = \frac{\mu_i}{\mu + \mu_i},$$

where $i \in \mathcal{A}$ denotes the choice of policy u_i . The expected fitness $F(u_i, u_j)$ can be expressed as a function of these average sojourn times as follows:

$$F(u_i, u_j) = \sum_{s, s' \in \mathcal{S}} \sum_{a, a' \in \mathcal{A}} T_s(a) T_{s'}(a') r(a, a'),$$

where $r(a, a')$ is the immediate fitness of a player using action a against an opponent playing a' . From the payoffs matrix (5.17), we obtain:

$$\begin{aligned} F(u_y, u_y) &= d, \\ F(u_x, u_y) &= T_1(x)b + T_0(x)d, \\ F(u_y, u_x) &= T_1(x)c + T_0(x)d, \\ F(u_x, u_x) &= T_1(x) [T_1(x)a + T_0(x)b] + T_0(x) [T_1(x)c + T_0(x)d]. \end{aligned} \quad (5.27)$$

By considering this matrix game as a representation of a standard evolutionary game, we can write the replicator equation. Let $\delta_x(t) \in [0, 1]$ be the probability that a player plays u_x at time t . We have that:

$$\begin{aligned} \dot{\delta}_x(t) &= \delta_x(t)(1 - \delta_x(t))(F(u_x, \delta_x(t)) - F(u_y, \delta_x(t))) \\ &= \delta_x(t)(1 - \delta_x(t)) [F(u_x, u_y) - F(u_y, u_y) \\ &\quad + \delta_x(t)(F(u_x, u_x) - F(u_y, u_x) + F(u_y, u_y) - F(u_x, u_y))]. \end{aligned} \quad (5.28)$$

We can compute the mixed equilibrium δ_x^* by solving the indifference principle equation $F(u_y, \delta_x^*) = F(u_x, \delta_x^*)$, with $F(u_i, q) = (1 - q)F(u_i, u_y) + qF(u_i, u_x)$ with $i \in \mathcal{A}$. We obtain:

$$\delta_x^* = \frac{F(u_y, u_y) - F(u_x, u_y)}{F(u_x, u_x) - F(u_y, u_x) + F(u_y, u_y) - F(u_x, u_y)}.$$

If $0 \leq \delta_x^* \leq 1$, then it is an admissible equilibrium for the matrix game and it corresponds to a rest point of the replicator dynamics (5.28). We resume our result in the following proposition.

Proposition 11. *If the distribution of the individual states is stationary, the equilibrium policy of the game can be computed by applying the matrix game approximation technique, which leads to the equilibrium*

$$\delta_x^* = \frac{s^*}{T_1(x)}. \quad (5.29)$$

under the condition $0 \leq s^* \leq T_1(x)$.

Proof. As we have mentioned above, the equilibrium can be computed by imposing the indifference principle, which leads to:

$$\delta_x^* = \frac{F(u_y, u_y) - F(u_x, u_y)}{F(u_x, u_x) - F(u_y, u_x) + F(u_y, u_y) - F(u_x, u_y)}.$$

By substituting the values of the fitnesses (5.27) into the latter equation and by carrying

out the values of the time ratios $T_1(x)$ and $T_0(x)$, we get:

$$\delta_x^* = \frac{\frac{\mu(d-b)}{\mu+\mu_x}}{\frac{\mu^2 a + \mu\mu_x b + \mu_x \mu c + \mu_x^2 d}{(\mu+\mu_x)^2} + d - \frac{\mu b + 2\mu_x d + \mu c}{(\mu+\mu_x)}} = \frac{\frac{\mu(d-b)}{\mu+\mu_x}}{\frac{\mu^2 a + \mu\mu_x b + \mu_x \mu c + \mu_x^2 d + d(\mu+\mu_x)^2 - (\mu+\mu_x)(\mu b + 2\mu_x d + \mu c)}{(\mu+\mu_x)^2}}.$$

After some algebra:

$$\delta_x^* = \frac{\mu(d-b)}{\mu+\mu_x} \cdot \frac{(\mu+\mu_x)^2}{\mu^2(a+d-b-c)} = \frac{s^*}{T_1(x)}.$$

In order for δ_x^* to be an admissible equilibrium, it must satisfy $\delta_x^* \in [0, 1]$, which completes the proof. \square

In figure 5.3 we illustrate the convergence of the replicator equation (5.28) to the equilibrium $\delta_x^* = 0.71875$, which is obtained by setting, as in the previous numerical example, $\mu = 10$, $\mu_x = 1.5$ and $\mu_y = 1$, and $a = -0.3$, $c = 0$, $b = 1$ and $d = 0.5$, starting from $q_x(0) = 0.2$ and $q_x(0) = 0.9$.

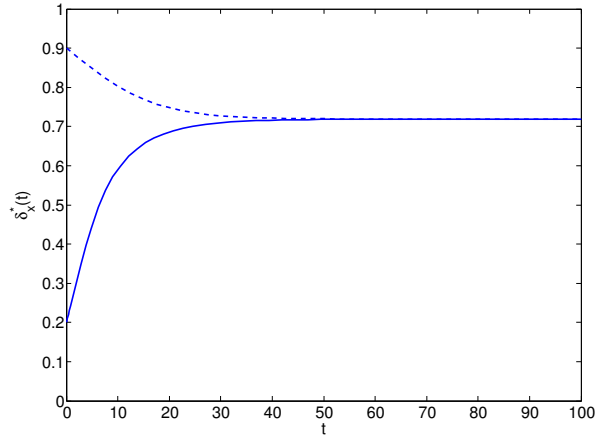


Figure 5.3: Convergence of the replicator dynamics to the equilibrium $\delta_x^* = 0.71875$, obtained by setting $\mu = 10$, $\mu_x = 1.5$ and $\mu_y = 1$, and $a = -0.3$, $c = 0$, $b = 1$ and $d = 0.5$, starting from $q_x(0) = 0.2$ and $q_x(0) = 0.9$.

5.3.8 Relation Between the Equilibria

In section 5.3.6, we suppose that each individual plays a deterministic policy $u_i \in \mathcal{U}_D$, which consists in always choosing action i in state 1 and action y otherwise and, by applying the singular perturbation method, we determine the equilibrium profile $\zeta^* =$

(w_1^*, q_x^*) of such game. In section 5.3.7, we assume that the distribution of individual states is already stationary which allows us to rewrite the game as a standard evolutionary game and to compute its mixed equilibrium δ^* . While q_x^* represents the share of the population choosing policy u_x at the equilibrium, δ^* represents the probability of choosing the deterministic policy u_x . This means that, in the first case, the population considered is *polymorphic* (with a fraction q^* of the population choosing policy u_x and the remaining $1 - q_x^*$ choosing u_y), while in the latter case we have a *monomorphic* population, where all individuals play the same mixed action (which consists in choosing u_x with probability δ_x^*). We can compare the equilibria q_x^* and δ_x^* obtained with these two different approaches.

Proposition 12. *The relation between the equilibrium δ_x^* and the equilibrium q_x^* is the following:*

$$q_x^* < \delta_x^*.$$

Proof. We evaluate the difference between the equilibria δ_x^* , obtained by solving our model as a matrix game in (5.29) with the value of the equilibrium obtained through the singular perturbation method, and q_x^* in (10):

$$\begin{aligned} \delta_x^* - q_x^* &= \frac{s^*(\mu + \mu_x)}{\mu} - \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)} \\ &= s^* \frac{(\mu_x - \mu_y)(\mu - s^*(\mu + \mu_x))}{\mu(\mu - s^*(\mu_x - \mu_y))}. \end{aligned}$$

□

If we consider the same values of the parameters chosen for the previous numerical example, the matrix game approach gives the following equilibrium:

$$\delta_x^* = \frac{s^*(\mu + \mu_x)}{\mu} = 0.71875 > q_x^* = 0.7097,$$

which verifies the proposition 12.

We now compare the two equilibria in terms of the average fitness of the population, i.e. $\bar{F}(\delta_x^*)$ and $\bar{F}(\xi^*)$, with $\xi^* = (w_1^*, q_x^*)$, and we verify that, as expected, the fitnesses are equal.

Proposition 13. *The average fitness of the population at the equilibria points obtained with the two approaches are equals, i.e. $\bar{F}(\xi^*) = \bar{F}(\delta_x^*)$.*

Proof. Considering the first approach, based on the singular perturbations method, we

have:

$$\begin{aligned}\bar{F}(\xi^*) &= q_x^* F_x(\xi^*) + (1 - q_x^*) F_y(\xi^*) \\ &= F_y(\xi^*) + q_x^* (F_x(\xi^*) - F_y(\xi^*)),\end{aligned}$$

with $\xi^* = (w_1^*, q_x^*)$. At the equilibrium state, we have $F_x(\xi^*) = F_y(\xi^*)$ and thus $\bar{F}(\xi^*) = F_y(\xi^*) = r_y(\xi^*) = w_1^* (q_x^* c + (1 - q_x^*) d) + (1 - w_1^*) d$. Knowing that $q_x^* w_1^* = s^*$, we obtain:

$$\bar{F}(\xi^*) = s^* c + (1 - s^*) d.$$

By rewriting the game into a matrix game, we obtain the following equilibrium profile: $\delta_x^* = \frac{s^*}{T_1(x)}$. The average fitness of the population in this case is:

$$\begin{aligned}\bar{F}(\delta_x^*) &= \delta_x^* F(u_x, \delta_x^*) + (1 - \delta_x^*) F(u_y, \delta_x^*) \\ &= F(u_y, \delta_x^*) + \delta_x^* (F(u_x, \delta_x^*) - F(u_y, \delta_x^*)).\end{aligned}$$

At the equilibrium, we have the following equality $F(u_x, \delta_x^*) = F(u_y, \delta_x^*)$ and then the average fitness of the population becomes simply:

$$\bar{F}(\delta_x^*) = F(u_y, \delta_x^*) = \delta_x^* F(u_y, u_x) + (1 - \delta_x^*) F(u_y, u_y).$$

The average fitness of the population is: $\bar{F}(\delta_x^*) = \delta_x^* (T_1(x) c + T_0(x) d) + (1 - \delta_x^*) d$. Since $\delta_x^* T_1(x) = s^*$,

$$\bar{F}(\delta_x^*) = s^* c + (1 - s^*) d,$$

which completes the proof. \square

Finally, we prove that the two equilibria obtained with the two approaches are in the same equivalent class in terms of average sojourn times. Let $T_1(q)$ be the average sojourn time in state 1 for an individual in a polymorphic population whose profile is ξ^* , and let $T_1(\delta_x^*)$ be the average sojourn time in state 1 for an individual in a monomorphic population playing mixed action δ^* . For the monomorphic population case, we obtain:

$$T_1(\delta_x^*) = \delta_x^* T_1(x) + (1 - \delta_x^*) T_1(y) = \delta_x^* \frac{\mu}{\mu + \mu_x} + (1 - \delta_x^*) \frac{\mu}{\mu + \mu_y}. \quad (5.30)$$

For the case of the polymorphic population:

$$T_1(q_x^*) = \pi_1(q_x^*) = \frac{\mu}{\mu + \mu_x q_x^* + \mu_y (1 - q_x^*)}. \quad (5.31)$$

The equivalence between these average sojourn times is proved in the following propo-

sition.

Proposition 14. *The mixed equilibrium δ_x^* and the equilibrium obtained by the singular perturbation approach yield to the same average sojourn times, i.e.*

$$T_1(\delta_x^*) = T_1(q_x^*).$$

Proof. We first rewrite δ_x^* as a function of the immediate payoffs in matrix (5.17):

$$\delta_x^* = \frac{(\mu_x + \mu)(d - b)}{\mu\Delta}.$$

where $\Delta := a - b - c + d$. We substitute it in (5.30), and we get:

$$T_1(\delta_x^*) = \frac{\mu\Delta + (\mu_x + \mu_y)}{\mu\Delta(\mu + \mu_y)}.$$

Analogously, we substitute the expression of s^* in q_x^* in proposition(10), and we rewrite it in (5.31), which leads to:

$$T_1(q_x^*) = \frac{\mu\Delta + (\mu_x + \mu_y)}{\mu\Delta(\mu + \mu_y)}.$$

which proves that $T_1(q_x^*) = T_1(\delta_x^*)$. □

The previous results show that we can define two equivalent classes for deterministic policies that yield same average fitness and average sojourn times.

5.3.9 Applications in Network Systems

Energy Control in Wireless Network

The two-states two pure actions model can be applied to describe a particular problem that arises in dynamic power control in mobile networks, which has been presented in (Altman and Hayel, 2008). The underlying idea is based on the fact that battery life is a very critical issue in wireless systems, and then, defining optimal transmission policies based on battery levels is very important. Moreover, this energy management problem is even more important when interactions occurs between the devices, complicating the analysis of such control systems. We then consider a system in which the action of each device impacts the lifetime of its battery or its battery level, and also impacts its transmission rate. A large number of mobiles transmit packets occasionally. Each transmitter can be in Full (F) or Almost empty (A) battery state. When a mobile is in F state it can choose to transmit packets using high (h) or low (l) power, whereas if it is in state A , it can only transmit packets using l power. In general, several mobiles

try to join a common receiver at the same time and interferences occur between the received signals. We suppose that transmissions are sparse so that the probability that more than two mobiles transmit simultaneously is negligible. We also assume that a transmission is successful either if the mobile is the only transmitter during a slot or if it transmits at higher power than the other transmitter. The time spent in state F depends on the action chosen by the mobile. Then the state of the mobile changes to the lower battery state A . After an exponentially distributed time, its battery state becomes empty. We assume that the battery is immediately recharged, so that the mobile goes back to state F . When transmitting at high power, the mobile's battery is consumed faster, and thus the transition rate from F to A is faster. Then, in this framework, the state space corresponds to $\mathcal{S} := \{A, F\}$, the action space is $\mathcal{A} := \{h, l\} = \mathcal{A}_F$ and the restricted action space for state A is $\mathcal{A}_A := \{l\}$. The set of deterministic policies $\mathcal{U}_D := \{u_h, u_l\}$ is composed of the policy u_h such that $u_h(A) = l$ and $u_h(F) = h$; the policy u_l such that $u_l(A) = l$ and $u_l(F) = l$. Then, the system (S) of coupled dynamics describes the time evolution of the proportion of mobiles in each state A and F , and at the same time the proportion of mobiles using policy u_h and u_l . By assuming that the state dynamic is highly faster than the policy dynamics (the change of policy has to be reimplemented into the mobiles by manufacturer or designers), then our analysis describes the equilibrium situation which corresponds to the long term evolution of this system.

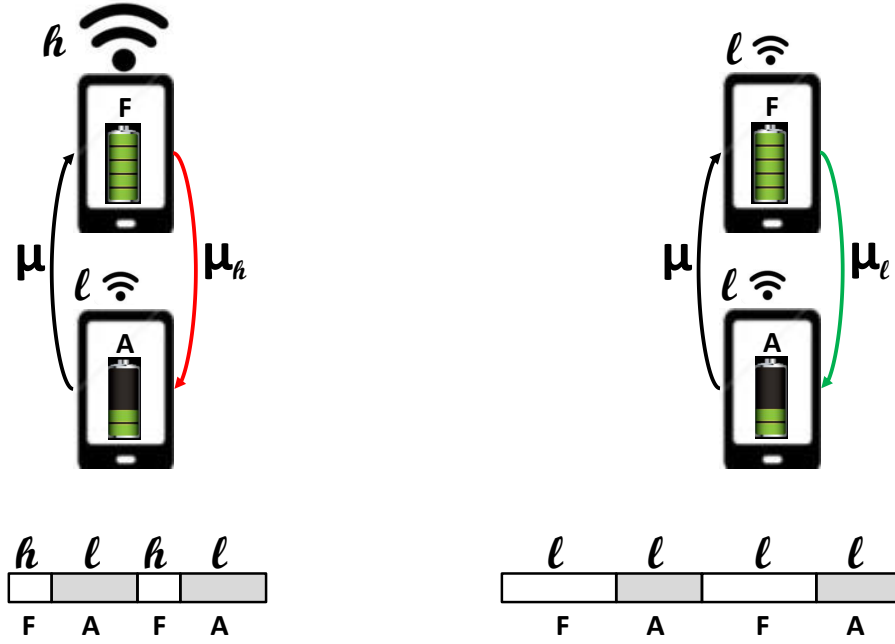


Figure 5.4: Energy control in wireless networks. Each mobile continuously moves from full battery state (F) to almost empty battery state (A) and then back to F. If the mobile chooses policy u_h (left side), in state F it transmits at high power (h), otherwise, if it chooses policy u_l it transmits at low power (l) (right side). In state A, it always transmits at low power. Since high power transmission is costly in terms of energy consumption, when transmitting at high power, the transition rate to state A is faster, and thus the time spent in state F is shorter.

Network Formation Games

Another application of the proposed model can be found in network formation games (Jackson, 2005). We consider a large number of nodes where each node is in one of two possible states: Infected or Susceptible, so that $S = \{I, S\}$. Nodes interact through pairwise interactions, during which both nodes exchange contents. If a node is in state S, it determines the type of unidirectional link to the node it is interacting with. The type of link can be charged at a price (p) or for free (f); if a node is in the infected state (state I), it can only create free links. Pay connection is safer, so that when a link is not a free one, the probability for a node to be infected is lower, independently of the choice of the other node to pay or not and also independent of the state of the other node. After some random time in I state, a node becomes susceptible again. This application into networks formation games could ask more assumptions on the model, especially if the transition rate depends on the state of the opponent. In this case we should define a more general game framework considering interactive MDPs, like anonymous sequen-

tial games ([Wiecek and Altman, 2014](#)). This generalized framework has a highly more complicated internal structure. We thus let its analysis as an extension for future works. It has to be noted that the singular perturbation approach, proposed in Subsection [5.3.6](#) is valid for this application, by considering a more complicated dynamics of individual state, which depends on the action also of the opponent.

5.4 Conclusion

In this chapter we considered a particular type of evolutionary game in which the action of the individual not only determines its immediate fitness but it also impacts the transition rates of the Markov process of the individual state. We defined the interdependent dynamics of the individual state and of the policy, where the evolution of policies distribution in the population dynamics is assumed to follow the well-known replicator dynamics. After introducing these combined dynamics in a general framework, we analyzed a particular case for which we proved the correspondence between stable rest points of the dynamics and the equilibrium profiles of the evolutionary game. Under the assumption that the two dynamics evolve with different timescales, we proposed two methods to obtain the rest points. We gave a complete characterization of these equilibrium profiles and we showed that these equilibria are equivalent in terms of average sojourn times and expected fitness. Finally, we illustrated our framework with two application scenarii in network systems.

Chapter 6

A Dynamic Approach for the Study of a Hawk-Dove Problem

“Mathematics without natural history is sterile, but natural history without mathematics is muddled.”

J. Maynard Smith, *Games, Sex and Evolution*, 1987

Summary

We develop in this chapter a dynamic model for the Hawk-Dove game in order to study the impact of the aggressive behavior of adults on the evolution of young individuals. As in MDEG presented in Chapter 4, players are associated with a MDP, but transition probability here do not depend on the player's action but on the action of the individual it meets. We define a Hawk-Dove game with four possible individual states, where the individual state determines the set of available actions. By considering the stationary distribution over the states, we transform the game into a standard evolutionary game, and we compute its equilibria. We then consider this Hawk-Dove dynamic game with group players as presented in Chapter 2.

6.1 Introduction

In this chapter we present a simple example of a Hawk-Dove game in a MDEG type of framework, where, in contrast with the theory presented in Chapter 4, we assume that transition probabilities do not depend on the action of the player but may depend on the action taken by those it encounters. The Hawk-Dove game is one of the most studied example to model the level of aggressiveness in a population and it finds many applications in different fields (see Section 1.4). The aim of our MDEG version of the game is to study the impact of the aggressive behavior of adults on the evolution of young individuals.

State of Art

Dynamical models of the Hawk-Dove game can be found in EGT literature. Houston and McNamara (Houston and McNamara, 1988) (Houston and McNamara, 1991) study a repeated version of the Hawk-Dove game, where each bird is associated with a state variable representing animal's level of energy evolving in time as a Markov Process. The action played by an animal depends on its reserve of energy and the fitness function depends on the average level of aggressiveness in the population. Birds are supposed to choose their policy in order to minimize the probability of dying during a period of time, and the ESS is in the form: *"play Hawk if and only if reserves are below a critical level of the energy resources $c^*(t)$ "*. McNamara et al. (McNamara et al., 1991) give an analytic justification of the computations performed in (Houston and McNamara, 1988) in a simpler setting, where they obtain interesting structural properties of the equilibrium using dynamic programming tools.

Some of the features considered in the Hawk-Dove model that we define can be found in (Altman and Hayel, 2008) (see the Introduction to the previous chapter), where the authors present a simple application of MDEG to mobile communications. As in our case, individuals can choose among an aggressive and a non-aggressive action only in one of the possible individual states, while in the others, only non-aggressive one is allowed.

In what follows, we consider a Hawk-Dove game where players are associated with one of four possible states, representing the age and the strength of the individual, and we suppose that the aggressive behavior is possible only in one of these states. Players aim at maximizing their expected immediate fitness during their lifetime. The system is assumed to be in its stationary state, which means that the distribution over the individ-

ual states is stationary, and the game is then transformed into a standard evolutionary game. We identify the stationary equilibrium of the game and we compute its value. In Section 6.3 we consider the case of group players presented in Chapter 2, for this Hawk-Dove MDEG.

6.2 The Model

Let us consider an infinitely large population of players matched in pairwise random interactions. We define a **four state model**, such that each individual is born ‘young’ (Y) and after each interaction with another randomly selected individual can become an adult or remain in young state. Adults can be aggressive (play Hawk) or non-aggressive (play Dove). If a young meets an aggressive adult either it evolves as a ‘weak adult’ (A_W) or it becomes a ‘weak young’ (Y_W). A weak young, when evolving, can only become a weak adult, whereas a young who has never been attacked may evolve into a ‘strong adult’ (A_S). The tuple $\{S, \mathcal{A}, \mathcal{Q}\}$ describing our game is defined as follows:

- The set of states is $S = \{Y, Y_W, A_S, A_W\}$, where Y, Y_W correspond to ‘young’ and ‘weak young’, A_S to ‘strong adult’ and A_W to ‘weak adult’.
- The set of actions is $\mathcal{A} = \{H, D\}$, with $\mathcal{A}_{A_S} = \mathcal{A}$, and $\mathcal{A}_{A_W} = \mathcal{A}_Y = \mathcal{A}_{Y_W} = \{D\}$. Players can choose whether to behave aggressively (H) or not (D) only in A_S state, while in all the other states, players can only play D .
- We define the probability of remaining in the young state (either weak or not) for a young individual by $y \in [0, 1]$. Consequently, the probability for a young individual to evolve into an adult one is $1 - y$. Analogously, the probability for an adult to stay in adult state (either weak or strong) is denoted by $x \in [0, 1]$, while with probability $1 - x$ an adult dies and is replaced by a young individual. We assume that the transition probabilities $Q_{s'}^t(s, a, a')$ (defined in Section 4.2) are time homogeneous and do not depend on the action of the player (a), but they may depend on that of its opponent (a'). We describe the set \mathcal{Q} as follows.
 - If a **young individual meets an adult one**, the transition probabilities depend on the action of the player it meets (and not on its action):

$$\begin{array}{ll}
 Q_Y(Y, \cdot, D) = y & Q_Y(Y, \cdot, H) = 0 \\
 Q_{Y_W}(Y, \cdot, D) = 0 & Q_{Y_W}(Y, \cdot, H) = y \\
 Q_{A_W}(Y, \cdot, D) = 0 & Q_{A_W}(Y, \cdot, H) = 1 - y \\
 Q_{A_S}(Y, \cdot, D) = 1 - y & Q_{A_S}(Y, \cdot, H) = 0
 \end{array}$$

- If a **weak young individual meets an adult one**, the transition probabilities only depend on its state (and not on the adult's action):

$$\begin{aligned} \mathcal{Q}_Y(Y_W, \cdot, \cdot) &= 0 \\ \mathcal{Q}_{Y_W}(Y_W, \cdot, \cdot) &= y \\ \mathcal{Q}_{A_W}(Y_W, \cdot, \cdot) &= 1 - y \\ \mathcal{Q}_{A_S}(Y_W, \cdot, \cdot) &= 0 \end{aligned}$$

- Analogously, when an **adult individual of type i meets another adult**, the transition probabilities only depend on the player state:

$$\mathcal{Q}_{A_i}(A_i, \cdot, \cdot) = \begin{cases} x & i = j \\ 0 & i \neq j \end{cases}$$

$$\mathcal{Q}_Y(A_i, \cdot, \cdot) = 0 \quad \mathcal{Q}_{Y_W}(A_i, \cdot, \cdot) = 0$$

Without loss of generality, we restrict to the set of stationary policies \mathcal{U}_S . We assume that if the population uses some common (mixed) stationary policy (except for a tagged player) then the global state process before the tagged user is born forms a time-homogeneous Markov chain; when the tagged player is born, it finds the Markov chain in steady state. Since the choice of an action in \mathcal{A} is limited to state A_S , we can identify a (mixed) stationary policy $u \in \mathcal{U}_S$ with the rule "play action H with probability q_u when in state A_S , D otherwise", i.e. $u(H|A_S) = q_u$, $u(D|A_S) = 1 - q_u$ and $u(D|A_W) = u(D|Y) = u(D|Y_W) = 1$. Let u_H and u_D be the deterministic policies corresponding respectively to $q_u = 1$ and $q_u = 0$. Let $v \in \mathcal{U}_S$ be the common policy adopted by the population. The fraction of adults of type $A_i \in \{A_W, A_S\}$ in such a population is defined as:

$$\alpha_i(v) := \mathbb{P}(A_i|v)$$

The total proportion of adults in the population, denoted by $P(A)$, does not depend on the policy v , and it corresponds to the expected lifetime spent in adult state over the total expected lifetime of an individual:

$$P(A) = \frac{\frac{1}{1-x}}{\frac{1}{1-x} + \frac{1}{1-y}}. \quad (6.1)$$

The probability of being attacked when young is $\alpha_S(v)q_v$, which corresponds to the probability of finding a strong adult playing aggressively. Consequently $\alpha_W(v) = P(A)(\alpha_S(v)q_v)$ and $\alpha_S(v) = P(A)(1 - \alpha_S(v)q_v)$. We can then explicit the fraction of

strong and weak adults in the population as a function of $P(A)$:

$$\alpha_S(v) = \frac{P(A)}{1 + P(A)q_v}, \quad \alpha_W(v) = \frac{q_v P(A)^2}{1 + P(A)q_v}. \quad (6.2)$$

Since in classical EGT the fitness function is related to the rate of reproduction of an individual, we suppose that only adults reproduce and thus the fitness of a young individual is assumed to be zero. The immediate payoff matrix, describing the fitness of the row player when meeting the column player, is the following:

	$A_S(H)$	$A_S(D)$	A_W	Y	Y_W
$A_S(H)$	$\frac{1}{2} - \delta$	1	1	1	1
$A_S(D)$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
A_W	$-\Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$	$\frac{1}{2} - \Delta$

where $\delta, \Delta > \frac{1}{2}$. The parameter δ represents the cost of the fight, whereas Δ reflects the loss of fitness for individuals who have been attacked when young. We omitted the null rows for Y and Y_W .

The expected immediate fitness for an adult of type $i \in \{A_S, A_W\}$, choosing policy $u \in \mathcal{U}_S$ against a population playing $v \in \mathcal{U}_S$ is $F(i, u, v) := E_{u,v}[r(i, s, s', a')]$, where $r(i, a, s', a')$ is the immediate fitness of an adult of type $i \in \{A_S, A_W\}$ playing $a \in \mathcal{A}_i$ against an individual in state $s' \in \mathcal{S}$ playing $a' \in \mathcal{A}_{s'}$. As each player is born in young state, $F(i, u, v)$ does not depend on the initial state.

6.2.1 Deriving the Expected Fitness and the Equilibria

We suppose that a tagged individual chooses stationary policy u against a population playing stationary policy v . The **total expected fitness** of the tagged player during its lifetime is given by:

$$F(u, v) = \frac{1}{1-x} (\alpha_S(v)F(S, u, v) + \alpha_W(v)F(W, u, v)), \quad (6.3)$$

where $\frac{1}{1-x}$ is the expected lifetime in adult state. $F(S, u, v)$ denotes the **immediate expected fitness** of an adult strong (A_S) playing u against a population playing v , with:

$$F(S, u, v) = \left[q_u \left((1 - P(A)) + \alpha_S(v)(1 - (\frac{1}{2} + \delta)q_v) + \alpha_W(v) \right) + (1 - q_u) \left(\frac{1 - P(A)}{2} + \frac{1 - q_v}{2} \alpha_S(v) + \frac{\alpha_W(v)}{2} \right) \right], \quad (6.4)$$

and $F(W, u, v)$ is the immediate expected fitness of a weak adult (A_W):

$$F(W, u, v) = \left[(1 - P(A))(\frac{1}{2} - \Delta) - \alpha_S(v)(\frac{1}{2} - \Delta - \frac{q_v}{2}) + (\frac{1}{2} - \Delta)\alpha_W(v) \right] = F(W, v). \quad (6.5)$$

Note that the fitness of a weak adult $F(W, v)$ does not depend on the policy u , as a weak adult always plays D . By substituting the expressions (6.4) and (6.5) in (6.3) we obtain:

$$\begin{aligned} F(u, v) = & \frac{\alpha_S(v)}{1 - x} \left[q_u \left((1 - P(A)) + \alpha_S(v)(1 - (\frac{1}{2} + \delta)q_v) + \alpha_W(v) \right) + \right. \\ & \left. + (1 - q_u) \left(\frac{1 - P(A)}{2} + \frac{1 - q_v}{2} \alpha_S(v) + \frac{\alpha_W(v)}{2} \right) \right] \\ & + \frac{\alpha_W(v)}{1 - x} \left[(1 - P(A))(\frac{1}{2} - \Delta) - \alpha_S(v)(\frac{1}{2} - \Delta - \frac{q_v}{2}) + (\frac{1}{2} - \Delta)\alpha_W(v) \right], \end{aligned} \quad (6.6)$$

where $\alpha_S(v) = \frac{P(A)}{1 + P(A)q_v}$ and $\alpha_W(v) = \frac{P(A)^2 q_v}{1 + P(A)q_v}$. By definition we have that $\alpha_S(v) + \alpha_W(v) = P(A)$, which leads to:

$$\begin{aligned} F(u, v) = & \frac{\alpha_S(v)}{1 - x} \left[q_u \left(1 - P(A) + P(A) - \alpha_S(v)(\frac{1}{2} + \delta)q_v \right) + \right. \\ & \left. + (1 - q_u) \left(\frac{1 - P(A)}{2} + \frac{P(A)}{2} - \frac{q_v \alpha_S(v)}{2} \right) \right] + \frac{\alpha_W(v)}{1 - x} \left[\frac{1}{2} - \Delta - \alpha_S(v)\frac{q_v}{2} \right] \\ = & \frac{1}{1 - x} \left[\alpha_S(v) \left(\frac{1}{2} - \delta q_v \alpha_S(v) \right) q_u - \Delta \alpha_W(v) + P(A) \left(\frac{1}{2} - \frac{q_v}{2} \alpha_S(v) \right) \right]. \end{aligned} \quad (6.7)$$

In the following theorem we present the symmetric equilibrium of the game, which depends on the value of the cost of the fight δ . We obtain that, if the the cost of the fight is below a certain threshold δ_* the deterministic symmetric equilibrium policy is u_H . If the cost is strictly higher than δ_* , the symmetric equilibrium is given by a mixed policy v_* .

Theorem 5. *Given the game described in Section 6.2 we have that:*

- i. *if $\delta \leq \delta_*$ the unique symmetric equilibrium is the deterministic policy u_H ;*
- ii. *if $\delta > \delta_*$ the unique symmetric equilibrium is the stationary policy v_* such that*

$$q_{v_*} = \frac{1}{P(A)(2\delta - 1)}, \quad (6.8)$$

where:

$$\delta_* = \frac{1 + P(A)}{2P(A)}.$$

Proof. We first look for the equilibrium in pure actions. If the population's policy is $v = u_D$ (i.e. all individuals in the population plays pure action D in state A_S), we obtain that:

$$\alpha_S(u_D) = P(A) \quad \text{and} \quad \alpha_W(u_D) = 0.$$

By substituting these values and $q_v = 0$ in equation (6.7) we get:

$$F(u, u_D) = \frac{P(A)(q_u + 1)}{2(1 - x)}. \quad (6.9)$$

Note that $F(u, u_D)$ is increasing in q_u , and thus u_D (corresponding to $q_u = 0$) is never an equilibrium.

We consider the case $v = u_H$. We have that:

$$\alpha_S(u_H) = \frac{P(A)}{1 + P(A)} \quad \text{and} \quad \alpha_W(u_H) = \frac{P(A)^2}{1 + P(A)}.$$

By substituting these values, $q_v = 1$ and, respectively $q_u = 0$ and $q_u = 1$ in 6.7 we get:

$$F(u_D, u_H) = -\frac{P(A)(2\Delta P(A) - 1)}{2(1 + P(A))(1 - x)},$$

$$F(u_H, u_H) = \frac{P(A)(1 + P(A) - \delta P(A) - \Delta P(A) - \Delta P(A)^2)}{(1 - x)(1 + P(A))^2}.$$

The aggressive deterministic policy u_H is an equilibrium only if the inequality $F(u_H, u_H) \geq F(u_D, u_H)$ holds. We have that:

$$F(u_H, u_H) - F(u_D, u_H) \geq 0 \quad \Leftrightarrow \quad -\frac{P(A)(2\delta P(A) - 1 - P(A))}{2(1 + P(A))^2(1 - x)} \geq 0.$$

The latter inequality is satisfied if and only if

$$1 + P(A) - 2\delta P(A) \geq 0.$$

We thus obtain that the aggressive behavior u_H is an equilibrium in deterministic policies if and only if $\delta \leq \frac{1 + P(A)}{2P(A)}$, which proofs the first statement of the theorem.

We now apply the indifference principle to find the equilibrium in mixed policies. By substituting $q_u = 0$ in (6.7) we obtain:

$$F(u_D, v) = \frac{1}{1 - x} \left[P(A) \left(\frac{1}{2} - \frac{q_v \alpha_S(v)}{2} \right) - \Delta \alpha_W(v) \right] = \frac{1}{1 - x} \left[\frac{P(A)(1 - 2\Delta P(A)q_v)}{2(1 + P(A)q_v)} \right].$$

For policy u_H , equation (6.7) becomes:

$$F(u_H, v) = \frac{1}{1-x} \left[\alpha_S(v) \left(\frac{1}{2} - \delta q_v \alpha_S(v) \right) - \Delta \alpha_W(v) + P(A) \left(\frac{1}{2} - \frac{q_v}{2} \alpha_S(v) \right) \right].$$

We impose the indifference among the two possible pure policies of the first player, i.e. $F(u_D, v) = F(u_H, v)$ and we find the following value:

$$q_{v_*} = \frac{1}{P(A)(2\delta - 1)}.$$

If $0 < q_{v_*} < 1$, then the policy v_* s.t. $v^*(H|A_S) = q_{v_*}$ is an admissible mixed policy. By imposing these constraints on (6.8), we get $\delta > \frac{1 + P(A)}{2P(A)}$, which completes the proof. \square

In figure 6.1 we plot the equilibrium probability q_{v_*} , i.e. the probability of being aggressive in state A_S at the equilibrium defined in (6.8), as a function of the proportion $P(A)$ of adults in the population, for three different values of δ . The continuous higher line is obtained with $\delta = 4$, the dotted line with $\delta = 6$, the continuous lower one with $\delta = 10$. As expected we can observe that q_{v_*} is a decreasing function of $P(A)$: this means that the higher the proportion of adults in the population is, the lower is the probability of being aggressive. As a matter of fact, the aggressive behavior of strong adults is rewarding only when adopted against a non-aggressive individual, while it is costly if adopted against another aggressive adult. In figure 6.2, q_{v_*} is plotted as a function of the threshold δ_* . The continuous higher line is obtained with $P(A) = 0.15$, the dotted line with $P(A) = 0.5$, the continuous lower one with $P(A) = 0.88$. As the value of δ_* represents the cost of the fight between two aggressive adults, as expected, we can see that q_{v_*} is decreasing in δ_* .

Remark 7. We observe that the equilibrium value q_{v_*} , corresponding to the probability of playing aggressively in state A_S , does not coincide with the mixed Nash equilibrium $q^* = \frac{1}{2\delta}$ of the standard Hawk-Dove game, even when $x \rightarrow 1$, so that $P(A) \simeq 1$. As a matter of fact, even if there are almost only adult individuals in the population, the fraction of weak adults may not be negligible, and thus we still may have a populations with two types of adults.

6.3 The Dynamic Hawk-Dove Game and Group Players

We now reformulate this Hawk-Dove game, considering groups players, as presented in Chapter 2. States are associated with individuals within a group, which are randomly matched in pairwise interactions, but the rules that determine their policy are

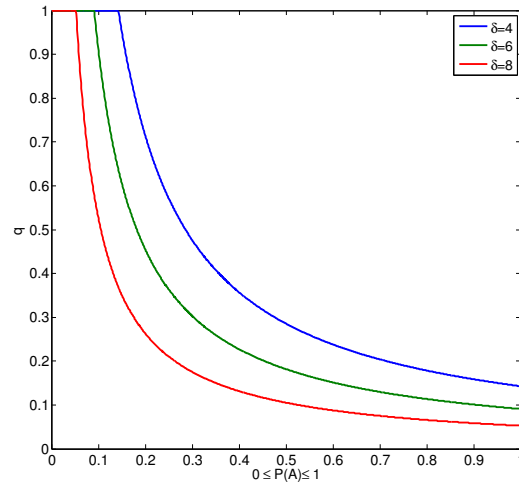


Figure 6.1: The equilibrium q_{v*} plotted as a function of the proportion $P(A)$ of adults in the population for different values of δ . The higher line is obtained with $\delta = 4$, the middle line with $\delta = 6$, the lower one with $\delta = 10$.

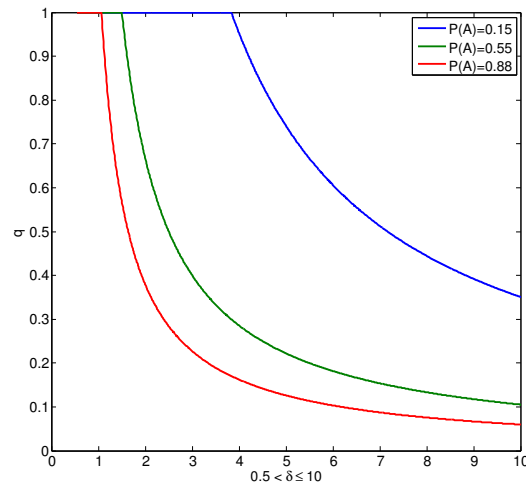


Figure 6.2: The equilibrium q_{v*} as a function of the value of δ for different values of $P(A)$. The higher line is obtained with $P(A) = 0.15$, the middle line with $P(A) = 0.5$, the lower one with $P(A) = 0.88$.

chosen in order to maximize the fitness of the group they belong to. We suppose that the large population of individuals is divided into N symmetric groups of the same size and that all players within a group choose the same policy. We further suppose that the probability of being a strong (or a weak) adult does not depend on the group the player belongs to, but it depends only on the average quantity $\bar{q}_u := \frac{\sum_{i=1}^N q_{u_i}}{N}$, which is the av-

erage probability of being aggressive in state A_S in the population. These probabilities are given, respectively by:

$$\alpha_S(\bar{u}) := \mathbb{P}(A_S|\bar{u}) = P(A) (1 - \alpha_S(\bar{u})\bar{q}_u),$$

$$\alpha_W(\bar{u}) := \mathbb{P}(A_W|\bar{u}) = P(A) (\alpha_S(\bar{u})\bar{q}_u).$$

where \bar{u} is the mixed policy such that $u(H|A_S) = \bar{q}_u$. These quantities can be rewritten as:

$$\alpha_S(\bar{u}) = \frac{P(A)}{1 + P(A)\bar{q}_u}, \quad (6.10)$$

$$\alpha_W(\bar{u}) = \frac{\bar{q}_u P(A)^2}{1 + P(A)\bar{q}_u}. \quad (6.11)$$

We suppose that the entire population chooses the same stationary policy $u \in \mathcal{U}_S$, except for a fixed group i playing $u_i \in \mathcal{U}_S$. We have that:

$$\bar{q}_u = \frac{q_{u_i} + (N-1)q_u}{N}.$$

6.3.1 Group Fitness and Equilibria

We associate to pairwise interactions among individuals the payoff matrix defined in Section 6.2. The expected fitness during the lifetime of a group-player i choosing policy $u_i \in \mathcal{U}_S$, in a population adopting policy $u \in \mathcal{U}_S$ is given by:

$$\begin{aligned} \Gamma(u_i, u) = \frac{1}{1-x} & \left[\alpha_S(\bar{u}) \left(\frac{1}{N} F(S, u_i, u_i) + \frac{N-1}{N} F(S, u_i, u) \right) \right. \\ & \left. + \alpha_W(\bar{u}) \left(\frac{1}{N} F(W, \cdot, u_i) + \frac{N-1}{N} F(W, \cdot, u) \right) \right], \end{aligned} \quad (6.12)$$

where:

$$\begin{aligned} F(S, u_i, u) &= q_{u_i} \left((1 - P(A)) + \alpha_S(\bar{u}) \left(1 - \left(\frac{1}{2} + \delta \right) q_u \right) + \alpha_W(\bar{u}) \right) \\ &+ (1 - q_u) \left(\frac{1 - P(A)}{2} + \frac{1 - q_v}{2} \alpha_S(\bar{u}) + \frac{\alpha_W(\bar{u})}{2} \right) \\ &= q_{u_i} \left(\frac{1 - P(A)}{2} + \alpha_S(\bar{u}) \left(\frac{1}{2} + \delta \right) q_u + \alpha_W(\bar{u}) \right) \\ &+ \left(\frac{1 - P(A)}{2} + \frac{1 - q_{u_i}}{2} \alpha_S(\bar{u}) + \frac{\alpha_W(\bar{u})}{2} \right), \end{aligned}$$

is the expected immediate fitness of a strong adult in group i , playing policy u_i against a population playing u and

$$F(W, \cdot, u) = (1 - P(A))\left(\frac{1}{2} - \Delta - \alpha_S(\bar{u})\left(\frac{1}{2} - \Delta - \frac{q_u}{2}\right) + \left(\frac{1}{2} - \Delta\right)\alpha_W(\bar{u})\right),$$

is the expected immediate fitness of a weak adult in a population whose policy is u . Quantities $\alpha_S(\bar{u})$ and $\alpha_W(\bar{u})$ are defined in (6.10) and (6.11) with $\bar{q}_u = \frac{q_{u_i} + (N-1)q_u}{N}$, while $1/N$ and $N - 1/N$ are, respectively, the probability for a player of meeting an opponent in the same group and in a different group.

In order to obtain insight on the impact of the groups on the equilibrium, we shall focus on some particular values of the parameters. We consider the case of two groups ($N = 2$) and we further fix $\Delta = 1$ and $P(A) = 0.5$, which means that there's an equal share of young and adult individuals in the population. We first look at symmetric equilibria in deterministic policies and we obtain that, if the cost of the fight is higher than a certain threshold, the deterministic policy u_D is an equilibrium, while, if the cost is low enough, the deterministic equilibrium policy is u_H . This result is presented in the following theorem.

Theorem 6. *Given the group-players dynamic Hawk-Dove model, with $N = 2$, $P(A) = 0.5$ and $\Delta = 1$, the equilibrium in deterministic policy is:*

- u_H for $\delta < 1.05$;
- u_D for $\delta > 1.333$.

Proof. Let us first consider a non aggressive population, i.e. $u = u_D$. We substitute $q_u = 0$ in (6.12) and we compare the results obtained for the two possible response of the tagged group i , $u_i = u_D$ and $u_i = u_H$.

$$\Gamma(u_D, u_D) = \frac{-1}{2P(A)(x-1)},$$

$$\Gamma(u_H, u_D) = \frac{-P(A)(4 + 2P(A) - 2P(A)\delta - P(A)^2 - 2P(A)\Delta)}{(x-1)(2 + P(A))^2 + 3P(A)^2\Delta + P(A)^3\Delta - P(A)^3\delta}$$

We compute the difference:

$$\Gamma(u_D, u_D) - \Gamma(u_H, u_D) = \frac{-P(A)(-4 + 3P(A)^2 + 4P(A)\delta + 4P(A)\Delta - 6P(A)^2\Delta - 2P(A)^3\Delta + 2P(A)^3\delta)}{2((x-1)(2 + P(A))^2)} \quad (6.13)$$

If the population is aggressive, i.e. if $u = 1$, we have that:

$$\Gamma(u_D, u_H) = \frac{P(A)(-2 - P(A) + P(A)^2 + 2P(A)\Delta - 3P(A)^2\Delta - P(A)^3\Delta + P(A)^3\delta)}{(x-1)(2 + P(A))^2}$$

$$\Gamma(u_H, u_H) = \frac{P(A)(-1 - P(A) + P(A)\delta + P(A)\Delta - P(A)^3\Delta - P(A)^2\Delta + P(A)^3\delta)}{(x-1)(1+P(A))^2}$$

In this case:

$$\begin{aligned} \Gamma(u_H, u_H) - \Gamma(u_D, u_H) &= \frac{P(A)(-2 - 3P(A) - 2P(A)^2 + 4P(A)\delta + 2P(A)\Delta - A^2\Delta)}{((x-1)(1+P(A))^2(2+P(A))^2)} \\ &+ \frac{P(A)(4P(A)^2\delta - 2P(A)^3\Delta + 4P(A)^3\delta - 2P(A)^3 + 2P(A)^4\delta - P(A)^4)}{((x-1)(1+P(A))^2(2+P(A))^2)} \end{aligned} \quad (6.14)$$

By substituting $\Delta = 1$ and $P(A) = 0.5$ in (6.13) and in (6.14), we obtain respectively:

$$\Gamma(u_D, u_D) - \Gamma(u_H, u_D) = -\frac{3(-4 + 3\delta)}{100(x-1)} > 0 \Leftrightarrow \delta > \frac{4}{3} = 1.333$$

$$\Gamma(u_H, u_H) - \Gamma(u_D, u_H) = \frac{(-61 + 58\delta)}{450(x-1)} > 0 \Leftrightarrow \delta < \frac{61}{58} \simeq 1.05$$

□

We now consider also symmetric mixed policies. We obtain that, if the cost is higher than a certain value, the group-players Hawk-Dove dynamic game admits a symmetric mixed (stationary) policy equilibrium u_* . We resume our results in the following theorem.

Theorem 7. *The group-players dynamic Hawk-Dove model, for the fixed values of the parameters $N = 2$, $P(A) = 0.5$, $\Delta = 1$, admits the following equilibria:*

i. *the deterministic policy u_H for $0.5 < \delta < 1.05$;*

ii. *the stationary policy u_* for $\delta \geq 0.8125$, with:*

$$q_{u_*} = -\frac{-4 + 7\delta - \sqrt{12 - 48\delta + 49\delta^2}}{(-1 + 2\delta)};$$

iii. *the deterministic equilibrium policy u_D for $\delta \geq 1.333$.*

Proof. By substituting the values of $N = 2$, $P(A) = 0.5$ and $\Delta = 1$ in 6.12, we obtain:

$$\Gamma(u_i, u) = \frac{(-16q_{u_i} - 5q_{u_i}^2 - 6q_{u_i}q_u + 9\delta q_{u_i}^2 - 16 + 10q_{u_i}\delta q_u - q_u^2 + \delta q_u^2)}{4((-1 + p)(4 + q_{u_i} + q_u)^2)}.$$

We compute the derivative:

$$\frac{\partial \Gamma(u_i, u)}{\partial u_i} = \frac{(-8 - 6q_{u_i} - 10q_u - q_{u_i}q_u - q_u^2 + 18\delta q_{u_i} + 2q_{u_i}\delta q_u + 10\delta q_u + 2\delta q_u^2)}{((x-1)(4 + q_{u_i} + q_u)^3)}$$

and, by imposing $\frac{\partial \Gamma(u_i, u)}{\partial u_i} = 0$, we find:

$$q_{u_i} = \frac{-8 - 10q_u - q_u^2 + 10\delta q_u + 2\delta q_u^2}{-6 - q_u + 18\delta + 2\delta q_u}$$

We now impose the symmetry $q_{u_i} = q_u$ and we obtain the two solutions:

$$q_{u^1} := -\frac{-4 + 7\delta - \sqrt{12 - 48\delta + 49\delta^2}}{(-1 + 2\delta)}$$

$$q_{u^2} := \frac{-(-4 + 7\delta + \sqrt{12 - 48\delta + 49\delta^2})}{(-1 + 2\delta)}$$

The second solution q_{u^2} is always negative and thus it is not acceptable, while the first solution satisfies $0 < q_{u^1} < 1$ iff $\delta > 0.8125$. We set $q_{u^*} = q_{u^1}$ and we define by u_* the policy corresponding to play H with probability q_{u^*} in state A_S . This completes the proof. \square

From Theorem 7, we can observe that it is possible to determine two intervals of the values of the cost δ (involved in a fight between two aggressive strong adults), in which we have two equilibria. More precisely, we can observe that:

- for $0.5 < \delta \leq 0.8125$ the game admits one deterministic equilibrium policy u_H ;
- for $0.8125 < \delta \leq 1.05$ the game admits two equilibria, the deterministic u_H and the mixed u_* ;
- for $1.05 < \delta < 1.333$ the game admits one mixed equilibrium u_* ;
- for $\delta \geq 1.333$ the game admits one deterministic equilibrium policy u_D .

In figure 6.3 we plot the probability of being aggressive in state A_S at the equilibrium, as a function of the cost δ . In figure 6.4 we compare the mixed equilibrium policies obtained for the dynamic Hawk-Dove game, respectively for group-players and for individual-players. We can observe that the probability of being aggressive (in state A_S) at the equilibrium is lower for group-players. This is coherent with what we found in chapter 2.

6.4 Conclusion

We revisited in this chapter the Hawk-Dove game in a MDEG framework. In this context an individual may be in one of several states, and the actions played by individuals

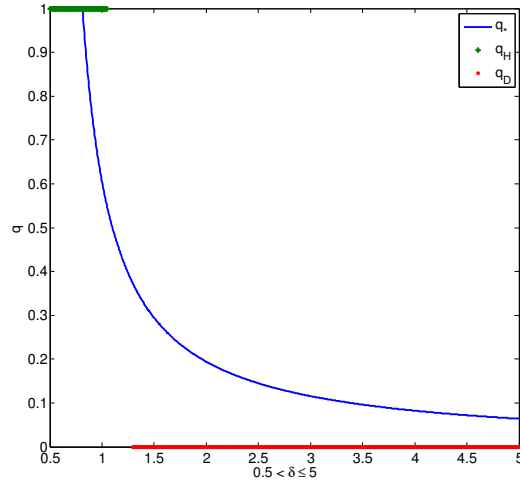


Figure 6.3: The probability of being aggressive in state A_S at the equilibria, obtained with $N = 2$, $P(A) = 0.5$ and $\Delta = 1$, as a function of the cost δ .

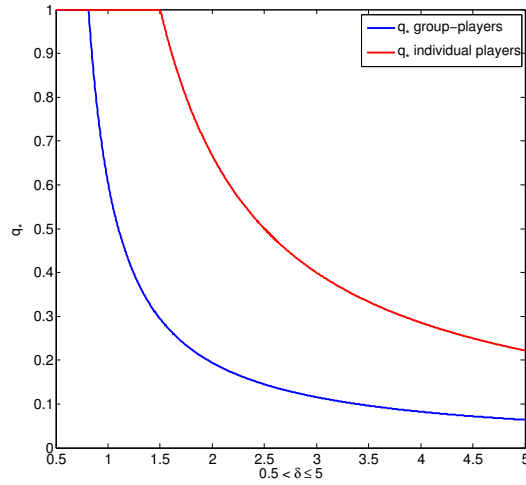


Figure 6.4: The equilibrium value q_{u*} as a function of the value of δ

involved in pairwise interactions determine not only the immediate fitness but also the transition probabilities of the players' individual state. We found that the described game has a unique symmetric equilibrium, which can be the pure aggressive policy u_H or a mixed equilibrium policy v_* , depending on the value of the cost of the fight between two aggressive individuals δ . We then further extended our dynamic Hawk-Dove game, by considering group-players as presented in Chapter 2. We fixed the number of groups $N = 2$, the share of adults $P(A) = 0.5$ and $\Delta = 1$, to compute the equilibria as a function of the cost of the fight δ . If the cost is very low, the aggres-

sive deterministic policy u_H is an equilibrium, otherwise, above certain values of the cost, we have a mixed equilibrium policy u_* and the deterministic policy u_H . In particular, for $\delta \in [0.8125, 1.05]$ the game admits the two equilibria u_H and u_* and, for $\delta \geq 1.333$ the equilibria u_D and u_* . The presence of groups in the dynamic Hawk-Dove game thus brings novel features with respect to the individual player game: the fact the non-aggressive policy u_D can be an equilibrium and the fact of having more than one equilibrium (for some values of δ). We also compared the mixed equilibria obtained respectively for the individual-players and the group-players games and we obtained that the probability of being aggressive in strong adult state is higher for individual-players. This is coherent with the results obtained in Chapter 2, where we showed that the presence of groups lowers the level of aggressiveness.

Part III

Stochastic Hybrid Dynamics

Hybrid Stochastic Systems

"Nature almost surely operates by combining chance with necessity,
randomness with determinism..."

Eric Chaisson, *Epic of Evolution: Seven Ages of the Cosmos*, 2007

Summary

In this chapter we extend the theory of control of an hybrid stochastic dynamical system to the case of a two players non-zero sum game. The system evolves in continuous time and it is subjected to abrupt changes of the parameters, determined by two (discrete time) Markov decision processes, each of which is controlled by a player that aims at minimizing its objective function. As we did in Chapter 5, we assume a two time scale behavior of the system: the lengths of the time intervals between the "jumps" of the parameters are assumed to be small, which means that parameters evolve faster than the state of the system. This allows us to approximate the hybrid game with a deterministic averaged dynamic game. We prove that an asymptotic Nash equilibrium of such hybrid game can be constructed on the basis of a Nash equilibrium of a deterministic averaged dynamic game.

7.1 Introduction

In this chapter, we study a non-zero sum dynamic game whose state is described as a hybrid dynamical system that evolves in continuous time and that is subjected to abrupt changes of the parameters. These changes are determined by two Markov decision processes, each of which is controlled by a player that aims at minimizing its objective function. The lengths of the time intervals between the “jumps” of the parameters are assumed to be equal to a small positive parameter ϵ , which means that parameters change their values frequently (the smaller is the parameter ϵ , the higher is the frequency). The MDPs are thus supposed to move faster than the state of the system. We then define an averaged dynamic game which allows us to approximate the hybrid system. More precisely, the main result of this chapter establishes that an asymptotic Nash equilibrium of the game defined by the solutions of the hybrid system (see Definition 1 in Section 7.2 below) can be constructed on the basis of a Nash equilibrium of the game defined by the solutions of the deterministic averaged system. The principal difference of our result from those obtained in the aforementioned works is in that the information structures of the hybrid and the averaged games need to be adjusted. In fact, we show that an asymptotic Nash equilibrium of the hybrid game, in which each of the players chooses its actions on the basis of the full information of its states/actions histories can be constructed on the basis of an open loop Nash equilibrium of the averaged game, where players have no information about their previous state/action history (nor about the state of the system).

State of Art

Hybrid dynamics control problems, with stochastic control and two time scales behavior, have been first introduced by Altman and Gaitsgory ([Altman and Gaitsgory, 1993](#)), in the case of one controller and linear dynamics and then extended to a two players zero-sum game ([Altman and Gaitsgory, 1995](#)). In ([Altman and Gaitsgory, 1997](#)) and ([Nguyen et al., 2001](#)) these models are generalized to non linear dynamics, respectively for one controller and for the zero-sum game.

The problem of optimization of a nonlinear hybrid system governed by a Markov decision process is close in nature to stochastic singular perturbed control problems intensively studied in the literature (see, for example, ([Abbad and Filar, 1992](#)), ([Bensoussan and Blankenship, 1987](#)), ([Bensoussan, 1989](#)), ([Bielecki and Filar, 1991](#)), ([Delebecque and Quadrat, 1978](#)), ([Kokotovic et al., 1986](#)), ([Kushner, 1990](#)), ([Philips and Kokotovic, 1981](#)), ([Pervozvansky and Gaitsgory, 1988](#)) and references therein). The main idea in

dealing with this category of problems is to optimize slow motions, assuming that the fast ones are approximated by their quasi stationary distributions of states obtained with “frozen” slow variables and controls, (see (Kokotovic et al., 1986), (Bensoussan, 1989), (O’Malley, 1974), (Kokotovic, 1984)). A common approach is an application of singular perturbations or averaging techniques to the Hamilton-Jacobi-Bellman (HJB) equation for problems in continuous time (as in (Bensoussan and Blankenship, 1987), (Bensoussan, 1989)) or to the dynamic programming equation for singularly perturbed discrete time MDPs (Abbad and Filar, 1992), (Bielecki and Filar, 1991), (Delebecque and Quadrat, 1978), (Philips and Kokotovic, 1981), (Pervozvansky and Gaitsgory, 1988). In contrast to this approach, here we continue the line of research started in (Altman and Gaitsgory, 1997) and (Nguyen et al., 2001), where an averaging method is applied directly to the “slow” stochastic equation.

Such type of systems arise in modeling admission control in telecommunication networks, in which the dynamics of the state variables (representing information packets transmission times at different nodes) is determined by MDPs describing the changes of the routes, their numbers and the type of sessions that are present in the networks (see e.g. (Shi et al., 1998)).

The chapter is organized as follows. In Sections 7.2 and 7.3, the hybrid and, respectively, averaged deterministic games are introduced. In Section 7.4, the main results are stated (see Propositions 15 and 16), and in Section 7.4.1, these are proved.

7.2 Hybrid Game

Let the dynamics of the state vector $Z(t) \in \mathbb{R}^N$ be described by the equation

$$\dot{Z}(t) = f^1(Z(t), Y^1(t)) + f^2(Z(t), Y^2(t)), \quad Z(0) = z_0, \quad t \in [0, 1], \quad (7.1)$$

where $f^i(\cdot, \cdot) : \mathbb{R}^N \times \mathbb{R}^{N_i} \rightarrow \mathbb{R}^N$, $i = 1, 2$, are continuous functions satisfying Lipschitz conditions in z (see Assumption 6 below). The functions $Y^i(t) \in \mathbb{R}^{N_i}$, $i = 1, 2$, are “controls” defined by two players. These controls are not chosen directly by the players. They are obtained as the result of the players controlling the transition probabilities of two associated stochastic discrete event systems described as follows.

The system i ($i = 1, 2$) has a finite state space \mathcal{S}^i and it changes its states at discrete moments of time $t_j = j\epsilon$, $j = 0, 1, \dots, \lfloor \epsilon^{-1} \rfloor$, where $\epsilon > 0$ is a small parameter representing the time unit and $\lfloor b \rfloor$ stands for the greatest integer which is smaller then or equal to b . The player i has a finite action space A^i , and if it chooses an action $a \in A^i$, then,

provided that the current state of system i is $s \in \mathcal{S}^i$, its next state will be $s' \in \mathcal{S}^i$ with the probability $\mathcal{Q}_{s'}^i(s, a) \geq 0$ ($\sum_{s' \in \mathcal{S}^i} \mathcal{Q}_{s'}^i(s, a) = 1$).

A policy $u^i = \{u_0^i, u_1^i, \dots, u_j^i, \dots, u_{\lfloor \epsilon^{-1} \rfloor}^i\}$ of the player i is a sequence of probability measures on A^i chosen as functions of the present state of the player and of its states/actions history. That is, $u_0^i = u_0^i(s_0^i)$ and $u_{j+1}^i = u_{j+1}^i(h_0^{j,i}, s_{j+1}^i)$, where $h_0^{j,i}$ is the states/actions history of the i^{th} system/player from time 0 to time t :

$$h_0^{j,i} := (s_l^i, a_l^i, l = 0, \dots, j), \quad j = 1, \dots, \lfloor \epsilon^{-1} \rfloor, \quad i = 1, 2. \quad (7.2)$$

Let \mathcal{U}_i stand for the set of all such policies and let \mathcal{F}^i be the discrete σ -algebra of all subsets of \mathbf{H}_i , which is the set of all possible states-actions histories of player i that can be observed until time ϵ^{-1} . Each initial distribution β of the initial states (s_0^1, s_0^2) and a policies pair $(u^1, u^2) \in \mathcal{U}_1 \times \mathcal{U}_2$ uniquely define a probability measure $\mathbb{P}_\beta^{(u^1, u^2)}$ over the space of samples $\mathbf{H} := \mathbf{H}_1 \times \mathbf{H}_2$ equipped with the discrete σ -algebra $\mathcal{F} := \mathcal{F}^1 \otimes \mathcal{F}^2$. Denote by $\mathbb{E}_\beta^{(u^1, u^2)}$ the corresponding expectation operator. When the distribution of the initial states is concentrated on a single states pair $\zeta = (\zeta^1, \zeta^2)$, we shall denote the corresponding probability measure and the mathematical expectation operator as $\mathbb{P}_\zeta^{(u^1, u^2)}$ and $\mathbb{E}_\zeta^{(u^1, u^2)}$.

Let $g^i : \mathcal{S}^i \times A^i \rightarrow \mathbb{R}^{N_i}$ be a given vector function and let S_j^i and A_j^i , $j = 0, 1, \dots, \lfloor \epsilon^{-1} \rfloor$, be the state-action processes of the system/player i . Then $Y^i(t)$ in (7.1) are defined by the equations

$$Y^i(t) = g^i(S_{\lfloor t/\epsilon \rfloor}^i, A_{\lfloor t/\epsilon \rfloor}^i), \quad i = 1, 2. \quad (7.3)$$

Note that the dynamics of the state vector $Z(t)$ is fully determined by the states/actions realizations $\{S_j^i, A_j^i\}$, $i = 1, 2$. For convenience, $Z(t)$ will be referred to as a “**macro state**” vector of the system and $\{S_j^i\}$ will be referred to as “**micro states**” of the players $i = 1, 2$.

Along with the class of policies $\mathcal{U} := \mathcal{U}_1 \times \mathcal{U}_2$ described above, we will be dealing with two other classes of policies $\mathcal{U}_M := \mathcal{U}_{M_1} \times \mathcal{U}_{M_2}$ and $\mathcal{U}_S := \mathcal{U}_{S_1} \times \mathcal{U}_{S_2}$, where \mathcal{U}_{M_i} and \mathcal{U}_{S_i} ($i = 1, 2$) are defined as follows:

- \mathcal{U}_{M_i} is the set of Markov policies of the player i such that, at every moment $t_j = j\epsilon$ of decision making, the probability measure u_j^i on A_i is chosen as a function of two arguments, one being the current moment of time and the other being the current micro state of the player i .
- \mathcal{U}_{S_i} is the set of stationary policies of the player i . That is, it is the set of policies

such that at any moment $t_j = j\epsilon$, the probability measure u_j^i on A_i is chosen as a function of the current micro state of the player i .

Note that, as follows from the definitions above, $\mathcal{U}_S \subset \mathcal{U}_M \subset \mathcal{U}$.

Assumption 5. *Under any stationary policy, the state space of the stochastic processes $\{S_j^i, A_j^i\}$ ($i = 1, 2$) forms an aperiodic Markov chain such that all states communicate (regular Markov chain).*

As mentioned above, we make the following assumption about the functions f^1 and f^2 .

Assumption 6. *The functions $f_i(\cdot, \cdot)$, $i = 1, 2$ are continuous and satisfy Lipschitz conditions in the first argument. That is, there exist positive constants C_1, C_2 such that*

$$\|f^i(z, y^i) - f^i(\tilde{z}, y^i)\| \leq C_i \|z - \tilde{z}\|, \quad i = 1, 2 \quad (7.4)$$

for arbitrary z and \tilde{z} from a sufficiently large subset of \mathbb{R}^N .

Remark 8. *According to their definitions, the processes $Y^1(t)$ and $Y^2(t)$ take values in some finite subsets D_1 and D_2 of \mathbb{R}^{N_1} and \mathbb{R}^{N_2} , that is, $Y^i(t) \in D_i$, $i = 1, 2$. Due to (7.4), it implies that $Z(t) \in D \forall t \in [0, 1]$, where D is a compact subset of \mathbb{R}^N . Note that, since $f_i(\cdot, \cdot)$, $i = 1, 2$ are continuous, there exist positive constants M_i , $i = 1, 2$, such that*

$$\|f^i(z, y^i)\| \leq M_i, \quad \forall (z, y^i) \in D \times D_i, \quad i = 1, 2. \quad (7.5)$$

Assume that the player i wishes to minimize its cost that only depends on the final value of the macro state, $G^i(\cdot) : \mathbb{R}^N \rightarrow \mathbb{R}$, $i = 1, 2$. To simplify the presentation, we assume that the cost functions $G_i(\cdot)$ $i = 1, 2$, satisfy Lipschitz conditions on D (although the continuity of these functions would suffice our purposes). Thus, we make the following assumption.

Assumption 7. *There exist positive constants C_G^i , $i = 1, 2$, such that*

$$|G^i(z) - G^i(\tilde{z})| \leq C_G^i \|z - \tilde{z}\| \quad \forall Z, \tilde{Z} \in D, \quad i = 1, 2. \quad (7.6)$$

Definition 8. *A pair of policies $u_\epsilon^* = (u_\epsilon^{1*}, u_\epsilon^{2*}) \in \mathcal{U}$ is an **Asymptotic Nash Equilibrium (ANE)** of the hybrid game if*

$$\begin{cases} \lim_{\epsilon \rightarrow 0} \mathbb{E}_\zeta^{(u_\epsilon^{1*}, u_\epsilon^{2*})} [G^1(Z(1))] \leq \lim_{\epsilon \rightarrow 0} \mathbb{E}_\zeta^{(u_\epsilon^1, u_\epsilon^{2*})} \{G^1(Z(1))\}, & \forall u_\epsilon^1 \in \mathcal{U}_1 \\ \lim_{\epsilon \rightarrow 0} \mathbb{E}_\zeta^{(u_\epsilon^{1*}, u_\epsilon^{2*})} \{G^2(Z(1))\} \leq \lim_{\epsilon \rightarrow 0} \mathbb{E}_\zeta^{(u_\epsilon^{1*}, u_\epsilon^2)} \{G^2(Z(1))\}, & \forall u_\epsilon^2 \in \mathcal{U}_2 \end{cases} \quad (7.7)$$

where limits in the left-hand-sides are assumed to exist and the initial micro states $\zeta = (\zeta_1, \zeta_2)$ are fixed (and known to the players).

The fact that the changes of the micro states/actions occur frequently (every moment $t_j = j\epsilon$) means that the processes $Y_1(t)$ and $Y_2(t)$ change their values on a much faster scale than does the macro state $Z(t)$. This allows one to approximate the solutions of the hybrid system (7.1) with the solutions of the deterministic averaged control system introduced in the next section. Our main result is the construction of ANE of the stochastic hybrid game on the basis of the Nash equilibrium of the deterministic dynamic game considered on the trajectories of the averaged system (see Sections 3-5 below).

7.3 Averaged Dynamic Game

Let $\omega_i(u_s^i) = \omega_i(u_s^i; s, a)$ be the vector of **steady state probabilities** of the micro state-action pair (s, a) of the player i when it uses a stationary policy $u_s^i \in \mathcal{U}_{S_i}$. That is,

$$\omega^i(u_s^i; s, a) := \lim_{j \rightarrow \infty} \mathbb{P}_{\zeta}^{u_s^i}(S_j^i = s, A_j^i = a), \quad \omega^i(u_s^i) = \{\omega^i(u_s^i; s, a)\}. \quad (7.8)$$

Due to the Assumption 5, the limit value $\omega^i(u_s^i; s, a)$ is independent of the initial conditions. Define the sets W_i by the equations

$$W_i = \bigcup_{u_s^i \in \mathcal{U}_{S_i}} \{\omega^i(u_s^i)\}, \quad i = 1, 2. \quad (7.9)$$

Note that the sets W_i are polyhedrons (see, e.g., (Derman, 1970), pp. 93-95).

Consider a deterministic system, in which the dynamics of the state vector $z(t)$ is described by the equation

$$\dot{z}(t) = \hat{f}^1(z(t), \omega^1(t)) + \hat{f}^2(z(t), \omega^2(t)), \quad z(0) = z_0 \quad t \in [0, 1], \quad (7.10)$$

where

$$\hat{f}^i(z, \omega^i) := \sum_{s,a} f^i(z, g^i(s, a)) \omega^i(s, a) \quad \forall \omega^i \in W_i, \quad i = 1, 2. \quad (7.11)$$

The functions $\omega^i(\cdot)$, $i = 1, 2$, are controls chosen by the players. These are assumed to be measurable functions of t that satisfy the inclusions $\omega^i(t) \in W_i \quad \forall t \in [0, 1]$, $i = 1, 2$.

As one can readily see, from Assumption 6 it follows that

$$\|\hat{f}^i(z, \omega^i) - \hat{f}^i(\tilde{z}, \omega^i)\| \leq C_i \|z - \tilde{z}\|, \quad \forall z, \tilde{z} \in D, \quad \forall \omega^i \in W_i, \quad (7.12)$$

and from Remark 8 it follows that

$$\|\hat{f}^i(z, \omega^i)\| \leq M_i, \quad \forall (z, \omega^i) \in D \times W_i. \quad (7.13)$$

Note that from (7.12) it follows that the solution of (7.10) exists and is unique with any choice of controls $\omega_i(\cdot)$, $i = 1, 2$.

Assume that the player i wishes to minimize the terminal cost function $G_i(z(1))$, where $G_i(\cdot)$ is the same as in the previous section ($i = 1, 2$). Given a pair of controls $(\omega^1(t), \omega^2(t))$, let $J_i(\omega^1, \omega^2)$ stand for the cost function of the payer i obtained with the players adopting these controls. That is,

$$J_i(\omega^1, \omega^2) := G^i(z(1)), \quad (7.14)$$

where $z(t)$ is the solution of (7.10) obtained applying $(\omega^1(t), \omega^2(t))$.

Definition 9. A pair of controls $(\omega^{1*}(\cdot), \omega^{2*}(\cdot))$ is a *Nash equilibrium of the averaged game* if

$$\begin{cases} J_1(\omega^{1*}, \omega^{2*}) \leq J_1(\omega^1, \omega^{2*}), \\ J_2(\omega^{1*}, \omega^{2*}) \leq J_2(\omega^{1*}, \omega^2), \end{cases} \quad (7.15)$$

for any $\omega^1(\cdot)$ (resp. $\omega^2(\cdot)$).

Note that the Nash equilibrium of the averaged game is defined in the loop setting, which means that the players have no information about their own past state/action history, nor about the state of the system. In the next section, we will show how ANE policies of the hybrid game can be constructed on the basis of a Nash equilibrium pair of the averaged game.

7.4 Construction of ANE Policies - Main Results

Let $\omega^i(\cdot)$ be a control of the player i in the averaged game. Partition the time interval $[0, 1]$ by the points

$$\tau_l := l\Delta(\epsilon), \quad l = 0, 1, 2, \dots, \ell(\epsilon), \quad \ell(\epsilon) := \lfloor \Delta(\epsilon)^{-1} \rfloor, \quad \tau_{\ell(\epsilon)+1} = 1, \quad (7.16)$$

where $\Delta(\epsilon) > 0$ is a function of ϵ such that

$$\lim_{\epsilon \rightarrow \infty} \Delta(\epsilon) = 0, \quad \lim_{\epsilon \rightarrow 0} \frac{\Delta(\epsilon)}{\epsilon} = \infty. \quad (7.17)$$

On each interval $[\tau_l, \tau_{l+1}]$ ($l = 0, 1, \dots, \lfloor \Delta(\epsilon)^{-1} \rfloor - 1$), define the time averages ω_l^i ,

$$\omega_l^i := \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} \omega^i(t) dt \in W_i, \quad i = 1, 2, \quad (7.18)$$

(the validity of the last inclusions follows from the convexity of W_i , $i = 1, 2$). Note that, from the fact that $\omega_l^i \in W_i$, it follows that there exists a stationary policy $u_s^i \in \mathcal{U}_{S_i}$ of the player i such that

$$\omega_l^i = \omega_l^i(s_l^i)$$

(see (7.8) and (7.9)). Let us define a policy of the player i in the hybrid game that consists of:

- i. Applying, at each $j = \lfloor \tau_l/\epsilon \rfloor, \lfloor \tau_l/\epsilon \rfloor + 1, \dots, \lfloor \tau_{l+1}/\epsilon \rfloor - 1$, the policy s_l^i , for any $l = 0, 1, 2, \dots, \ell(\epsilon) - 1$;
- ii. Applying an arbitrary stationary policy for $\lfloor \tau_{\ell(\epsilon)}/\epsilon \rfloor, \lfloor \tau_{\ell(\epsilon)}/\epsilon \rfloor + 1, \dots, \lfloor \epsilon^{-1} \rfloor$.

Let us denote this policy as $u_\epsilon^i(\omega^i)$. Note that, by construction, $u_\epsilon^i(\omega^i) \in \mathcal{U}_{M_i}$, $i = 1, 2$. The main results of the paper are Propositions 15 and 16 stated below.

Proposition 15. *Let $\omega(t) = (\omega^1(t), \omega^2(t))$ be a pair of controls and let $z(t)$ be the corresponding solution of (7.10). Let also $u_\epsilon(\omega) = (u_\epsilon^1(\omega^1), u_\epsilon^2(\omega^2))$ be the pair of policies defined above and let $Z(t)$ be the random trajectory of system (7.1) obtained with the players using these policies. Then*

$$\max_{t \in [0, 1]} \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|Z(t) - z(t)\| \leq \gamma(\epsilon), \quad (7.19)$$

where $\lim_{\epsilon \rightarrow 0} \gamma(\epsilon) = 0$. Also,

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}_\zeta^{u_\epsilon(\omega)} G^i(Z(1)) = J^i(\omega^1, \omega^2), \quad i = 1, 2. \quad (7.20)$$

Proposition 16. *Let $\omega^*(t) := (\omega^{1*}(t), \omega^{2*}(t))$ be a Nash equilibrium of the averaged game. Let also $u_\epsilon^*(\omega^*) = (u_\epsilon^{1*}(\omega^{1*}), u_\epsilon^{2*}(\omega^{2*}))$ be defined as above (considering $\omega(t) = \omega^*(t)$). Then $u_\epsilon^*(\omega^*)$ is an asymptotic Nash equilibrium of the hybrid game.*

The proofs of Propositions 15 and 16 are given in the next section.

7.4.1 Proofs of Propositions 15 and 16

Let us first recall some results from the MDP theory that are needed for the proofs of the main results. Let h^i stand for the full micro states-actions history of the player i (that is, $h^i = h_0^{\lfloor \epsilon^{-1} \rfloor, i}$; see (7.2)) and let $h_m^{K, i}$ be a part of this history corresponding to the interval

$[t_m, t_{m+K}]$,

$$h_m^{K,i} := \{s_m^i, a_m^i, \dots, s_{m+K}^i, a_{m+K}^i\}.$$

Denote by $\phi_m^{K,i}(h^i; s, a)$ the frequency of appearance of the micro state-action pair $(s, a) \in \mathcal{S}^i \times A^i$ on the interval $[t_m, t_{m+K}]$,

$$\phi_m^{K,i}(h^i; s, a) := \frac{1}{K+1} \sum_{j=m}^{m+K} \mathbb{1}_{\{s_j^i=s, a_j^i=a\}}, \quad \phi_m^{K,i}(h^i) = \{\phi_m^{K,i}(h^i, s, a)\}, \quad (7.21)$$

where $\mathbb{1}_{\{s_n^i=s, a_n^i=a\}}$ is the indicator function. If H^i is a random realization of h^i , we denote:

$$\phi_m^{K,i}(s, a) = \phi_m^{K,i}(H^i, s, a), \quad \phi_m^{K,i} = \{\phi_m^{K,i}(s, a)\}. \quad (7.22)$$

Lemma 4. *The following relationships are valid:*

$$\lim_{K \rightarrow \infty} \sup_{\zeta} \sup_{u_s^i \in \mathcal{U}_{S_i}} \mathbb{E}_{\zeta}^{u_s^i} \|\phi_0^{K,i} - \omega^i(u_s^i)\| = 0, \quad (7.23)$$

and also

$$\lim_{K \rightarrow \infty} \sup_{\zeta} \sup_{u^i \in \mathcal{U}_i} \mathbb{E}_{\zeta}^{u^i} d_0^{K,i} = 0, \quad (7.24)$$

where

$$d_0^{K,i} := \text{dist}(\phi_0^{K,i}, W_i) = \inf_{\omega^i \in W_i} \|\phi_0^{K,i} - \omega^i\|.$$

Proof. For a proof of the lemma, see Theorem 4.1 in (Altman and Gaitsgory, 1997). \square

Proof of Proposition 15. Let $\omega(t) = (\omega^1(t), \omega^2(t))$ be as in the statement of Proposition 15 and let $z(t)$ be the corresponding solution of (7.10). Define the sequence of vectors ξ_l , $l = 0, \dots, \ell(\epsilon)$, as the solution of the following difference equation

$$\xi_{l+1} = \xi_l + \Delta(\epsilon) \left[\hat{f}^1(z(\tau_l), \omega_l^1) + \hat{f}^2(z(\tau_l), \omega_l^2) \right], \quad \xi_0 := z_0. \quad (7.25)$$

By definition

$$z(\tau_{l+1}) = z(\tau_l) + \int_{\tau_l}^{\tau_{l+1}} \hat{f}^1(z(t), \omega^1(t)) dt + \int_{\tau_l}^{\tau_{l+1}} \hat{f}^2(z(t), \omega^2(t)) dt. \quad (7.26)$$

Hence,

$$\|z(\tau_{l+1}) - \xi_{l+1}\| \leq \|z(\tau_l) - \xi_l\| + \Delta(\epsilon) \sum_{i=1}^2 \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} \hat{f}^i(z(t), \omega^i(t)) dt - \hat{f}^i(z(\tau_l), \omega_l^i) \right\|.$$

The function $\hat{f}^i(z, \omega^i)$ is linear in ω^i and Lipschitz continuous in z . Consequently, there exists an appropriately chosen positive constant \tilde{M} such that

$$\|z(\tau_{l+1}) - \xi_{l+1}\| \leq \|z(\tau_l) - \xi_l\| + \Delta(\epsilon)^2 \tilde{M},$$

which implies that, for any $l = 1, \dots, \ell(\epsilon)$,

$$\|z(\tau_l) - \xi_l\| \leq \ell(\epsilon) \Delta(\epsilon)^2 \tilde{M} \leq \Delta(\epsilon) \tilde{M}. \quad (7.27)$$

Let Z_l , $l = 0, 1, \dots, \ell(\epsilon)$, be the sequence of random vectors defined by the equation

$$Z_{l+1} = Z_l + \int_{\tau_l}^{\tau_{l+1}} f^1(Z_l, Y^1(t)) dt + \int_{\tau_l}^{\tau_{l+1}} f^2(Z_l, Y^2(t)) dt, \quad l = 0, 1, \dots, \ell(\epsilon) - 1. \quad (7.28)$$

By subtracting the latter from (7.25) and taking the expectation over the probability measure corresponding to the policies $u_\epsilon(\omega) = (u_\epsilon^1(\omega^1), u_\epsilon^2(\omega^2))$, we obtain

$$\begin{aligned} \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|\xi_{l+1} - Z_{l+1}\| &\leq \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|\xi_l - Z_l\| + \Delta(\epsilon) \sum_{i=1}^2 \mathbb{E}_\zeta^{u_\epsilon(\omega)} \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(Z_l, Y^i(t)) dt - \hat{f}^i(z(\tau_l), \omega_l^i) \right\| \\ &\leq \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|\xi_l - Z_l\| + \Delta(\epsilon) \sum_{i=1}^2 \mathbb{E}_\zeta^{u_\epsilon(\omega)} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(Z_l, Y^i(t)) dt - \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt \right\| \right. \\ &\quad \left. + \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt - \hat{f}^i(z(\tau_l), \omega_l^i) \right\| \right\}. \end{aligned} \quad (7.29)$$

Using the Lipschitz continuity of $f^i(z, \omega^i)$ in z and the estimate (7.27), one can obtain that

$$\begin{aligned} \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(Z_l, Y^i(t)) dt - \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt \right\| &\leq \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} C_i \|Z_l - z(\tau_l)\| dt \\ &\leq C_i (\|Z_l - \xi_l\| + \|\xi_l - z(\tau_l)\|) \leq C_i (\|Z_l - \xi_l\| + \tilde{M} \Delta(\epsilon)). \end{aligned} \quad (7.30)$$

By substituting the latter inequality in (7.29), one obtains

$$\begin{aligned} \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|\xi_{l+1} - Z_{l+1}\| &\leq \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|\xi_l - Z_l\| + (C_1 + C_2) \tilde{M} \Delta(\epsilon)^2 + \Delta(\epsilon) \mathbb{E}_\zeta^{u_\epsilon(\omega)} \left\{ \sum_{i=1}^2 C_i \|Z_l - \xi_l\| \right. \\ &\quad \left. + \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt - \hat{f}^i(z(\tau_l), \omega_l^i) \right\| \right\}. \end{aligned} \quad (7.31)$$

Let $K(\epsilon) = \min_{l=0,1,\dots,\ell(\epsilon)-1} (\lfloor \tau_{l+1/\epsilon} \rfloor - \lfloor \tau_{l/\epsilon} \rfloor)$. Note that the following estimates are valid

(see (Shi et al., 1998))

$$\begin{aligned}
 2 &\geq \lfloor \tau_{l+1/\epsilon} \rfloor - \lfloor \tau_{l/\epsilon} \rfloor - K(\epsilon), \quad \left| K(\epsilon) - \frac{\Delta(\epsilon)}{\epsilon} \right| \leq 1 \\
 &\Rightarrow \left| K(\epsilon)^{-1} - \frac{\epsilon}{\Delta(\epsilon)} \right| \leq \frac{\epsilon^2}{\Delta(\epsilon)^2} \left(\frac{1}{1 - \epsilon/\Delta(\epsilon)} \right). \tag{7.32}
 \end{aligned}$$

From (7.32) it follows that there exist positive constants L_1^i, L_2^i such that, for $i = 1, 2$, and $l = 1, \dots, \ell(\epsilon)$,

$$\left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt - \frac{\epsilon}{\Delta(\epsilon)} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) \right\| \leq L_1^i \frac{\epsilon}{\Delta(\epsilon)}, \tag{7.33}$$

$$\left\| \frac{\epsilon}{\Delta(\epsilon)} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) - \frac{1}{K(\epsilon) + 1} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) \right\| \leq L_2^i \frac{\epsilon}{\Delta(\epsilon)}. \tag{7.34}$$

Let (as above) $\phi_m^{K,i} = \{\phi_m^{K,i}(s, a)\}$ stand for the state-action frequencies of the controller i that corresponds to a random realization of the history H^i (see (7.22)). Then

$$\frac{1}{K(\epsilon) + 1} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) = \sum_{(s,a) \in \mathcal{S}^i \times \mathcal{A}^i} \phi_{\lfloor \tau_{l/\epsilon} \rfloor}^{K(\epsilon),i}(s, a) f^i(z(\tau_l), g^i(s, a)).$$

Note that from (7.23) (see Lemma 4) it follows that there exists a function $\mu_i : \mathbb{N} \rightarrow \mathbb{R}$, with $\lim_{K \rightarrow \infty} \mu_i(K) = 0$ (implying $\lim_{\epsilon \rightarrow 0} \mu_i(K(\epsilon)) = 0$) such that

$$\mathbb{E}_{S_{\lfloor \tau_{l/\epsilon} \rfloor}^i}^{u_s^i} \left\{ \max_{(s,a)} |\phi_{\lfloor \tau_{l/\epsilon} \rfloor}^{K(\epsilon),i}(s, a) - \omega_l^i(s, a)| \right\} \leq \mu_i(K(\epsilon)). \tag{7.35}$$

Since

$$\begin{aligned}
 &\sum_{i=1}^2 \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt - \hat{f}^i(z(\tau_l), \omega_l^i) \right\| \right\} \\
 &\leq \sum_{i=1}^2 \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt - \frac{\epsilon}{\Delta(\epsilon)} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) \right\| \right. \\
 &\quad + \left\| \frac{\epsilon}{\Delta(\epsilon)} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) - \frac{1}{K(\epsilon) + 1} \sum_{n=\lfloor \tau_{l/\epsilon} \rfloor}^{\lfloor \tau_{l+1/\epsilon} \rfloor + K(\epsilon)} f^i(z(\tau_l), g^i(S_n^i, A_n^i)) \right\| \\
 &\quad \left. + \left\| \sum_{s,a} \phi_{\lfloor \tau_{l/\epsilon} \rfloor}^{K(\epsilon),i}(H^i; s, a) f^i(z(\tau_l), g^i(s, a)) - \hat{f}^i(z(\tau_l), \omega_l^i) \right\| \right\}
 \end{aligned}$$

and since, by definition, $\hat{f}^i(z(\tau_l), \omega_l^i) := \sum_{s,a} \omega_l^i(s,a) f^i(z(\tau_l), g^i(s,a))$, one can obtain, using (7.35), (7.33), (7.34),

$$\begin{aligned}
 & \sum_{i=1}^2 \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^i(z(\tau_l), Y^i(t)) dt - \hat{f}^i(z(\tau_l), \omega_l^i) \right\| \right\} \leq (\tilde{L}_1 + \tilde{L}_2) \frac{\epsilon}{\Delta(\epsilon)} \\
 & + \sum_{i=1}^2 \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} \left\{ \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], i}^{K(\epsilon)}(H^i; s, a) - \omega_l^i(s, a) \right| \|f^i(z(\tau_l), g^i(s, a))\| \right\} \\
 & \leq (\tilde{L}_1 + \tilde{L}_2) \frac{\epsilon}{\Delta(\epsilon)} + \sum_{i=1}^2 M_i \mathbb{E}_{\zeta}^{u_{\epsilon}^i(\omega)} \left\{ \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], i}^{K(\epsilon)}(H^i; s, a) - \omega_l^i(s, a) \right| \right\} \\
 & \leq (\tilde{L}_1 + \tilde{L}_2) \frac{\epsilon}{\Delta(\epsilon)} + \sum_{i=1}^2 M_i \mathbb{E}_{\zeta}^{u_{\epsilon}^i(\omega)} \left\{ \mathbb{E}_{S_{[\tau_l/\epsilon]}}^{s_l^i} \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], i}^{K(\epsilon)}(H^i; s, a) - \omega_l^i(s, a) \right| \right\} \\
 & \leq (\tilde{L}_1 + \tilde{L}_2) \frac{\epsilon}{\Delta(\epsilon)} + \sum_{i=1}^2 M_i \mu_i(K(\epsilon))
 \end{aligned} \tag{7.36}$$

where $\tilde{L}_i = L_1^i + L_2^i$ (L_j^i , $j = 1, 2$, being the constants from (7.33) and (7.34)) and M_i , $i = 1, 2$, are the constants from (7.5). The substitution of the latter into (7.31) leads to

$$\mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} [\|\xi_{l+1} - Z_{l+1}\|] \leq \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} [\|\xi_l - Z_l\|] + \tilde{M}\Delta(\epsilon) \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} [\|\xi_l - Z_l\|] + \Delta(\epsilon)\kappa(\epsilon), \tag{7.37}$$

where $\lim_{\epsilon \rightarrow 0} \kappa(\epsilon) = 0$ and \tilde{M} is an appropriately chosen constant. This implies (see Proposition 5.1 in (Gaitsgory, 1992)) that

$$\mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} [\|\xi_l - Z_l\|] \leq \nu(\epsilon), \quad l = 1, \dots, \ell(\epsilon), \tag{7.38}$$

with $\lim_{\epsilon \rightarrow 0} \nu(\epsilon) = 0$. By definition, for $l = 1, \dots, \ell(\epsilon) - 1$:

$$Z(\tau_{l+1}) = Z(\tau_l) + \int_{\tau_l}^{\tau_{l+1}} f^1(Z(t), Y^1(t)) dt + \int_{\tau_l}^{\tau_{l+1}} f^2(Z(t), Y^2(t)) dt \tag{7.39}$$

By subtracting (7.28) from (7.39), one obtains

$$\mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} \|Z(\tau_{l+1}) - Z_{l+1}\| \leq \mathbb{E}_{\zeta}^{u_{\epsilon}(\omega)} \left[\|Z(\tau_l) - Z_l\| + (C_1 + C_2) \int_{\tau_l}^{\tau_{l+1}} \|Z(t) - Z_l\| dt \right]. \tag{7.40}$$

Note that, due to (7.5),

$$\|Z(t) - Z(\tau_l)\| \leq (M_1 + M_2)\Delta(\epsilon) \quad \forall t \in [\tau_l, \tau_{l+1}], \tag{7.41}$$

Hence,

$$\mathbb{E}_\zeta^{u_\epsilon(\omega)} \|Z(\tau_{l+1}) - Z_{l+1}\| \leq \mathbb{E}_\zeta^{u_\epsilon(\omega)} [\|Z(\tau_l) - Z_l\| + C\Delta(\epsilon)\|Z(\tau_l) - Z_l\|] + C(M_1 + M_2)\Delta(\epsilon)^2. \quad (7.42)$$

Due to Proposition 5.1 from (Gaitsgory, 1992), the latter implies that there exists an appropriately chosen positive constant \tilde{M} such that

$$\mathbb{E}_\zeta^{u_\epsilon(\omega)} \|Z(\tau_l) - Z_l\| \leq \tilde{M}\Delta(\epsilon), \quad l = 0, 1, \dots, \ell(\epsilon). \quad (7.43)$$

By combining (7.27), (7.38), (7.43), we may conclude that

$$\begin{aligned} \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|z(\tau_l) - Z(\tau_l)\| &\leq \mathbb{E}_\zeta^{u_\epsilon(\omega)} \{\|z(\tau_l) - \xi_l\| + \|\xi_l - Z_l\| + \|Z_l - Z(\tau_l)\|\} \\ &\leq \nu(\epsilon) + \Delta(\epsilon)\tilde{M}, \quad l = 0, 1, \dots, \ell(\epsilon). \end{aligned} \quad (7.44)$$

This and the fact that $\|z(t) - z(\tau_l)\| \leq M\Delta(\epsilon) \forall t \in [\tau_l, \tau_{l+1}]$ imply the validity of (7.19) with some $\gamma(\epsilon)$ such that $\lim_{\epsilon \rightarrow 0} \gamma(\epsilon) = 0$. The validity of (7.20) follows from (7.19) since (see (7.6))

$$|\mathbb{E}_\zeta^{u_\epsilon(\omega)} G^i(Z(1)) - G^i(z(1))| \leq C_G^i \mathbb{E}_\zeta^{u_\epsilon(\omega)} \|Z(1) - z(1)\| \leq C_G^i \gamma(\epsilon).$$

Thus, the proof of the proposition is completed. \square

Proof of Proposition 16. Let $h^1 = h_0^{\lfloor \epsilon^{-1} \rfloor, 1} = (s_0^1, a_0^1, \dots, s_{\lfloor \epsilon^{-1} \rfloor}^1, a_{\lfloor \epsilon^{-1} \rfloor}^1)$ be a realization of a state-action trajectory of player 1 and let

$$y^1(t, h^1) := g^1(s_{\lfloor t/\epsilon \rfloor}^1, a_{\lfloor t/\epsilon \rfloor}^1). \quad (7.45)$$

Define the projection of the vector of the state action frequencies $\phi_{\lfloor \tau_l/\epsilon \rfloor}^{K(\epsilon), 1}(h^1)$ (see (7.21)) onto W_1 by the equation:

$$\omega_l^1(h^1) := \arg \min_{\omega^1 \in W_1} \left\| \phi_{\lfloor \tau_l/\epsilon \rfloor}^{K(\epsilon), 1}(h^1) - \omega^1 \right\|. \quad (7.46)$$

where $K(\epsilon)$ is as in the proof of Proposition 15. Define the function

$$\omega^1(t, h^1) := \{\omega_l^1(h^1) \text{ for } t \in [\tau_l, \tau_{l+1}], \quad l = 1, \dots, \ell(\epsilon)\},$$

and denote by $z(t, h^1)$ the solution of the differential equation

$$\dot{z}(t, h^1) = \hat{f}_1(z(t, h^1), \omega^1(t, h^1)) + \hat{f}_2(z(t, h^1), \omega^{2*}(t)). \quad (7.47)$$

Define the piecewise constant function

$$\tilde{\omega}^{2*}(t) = \{\omega_l^{2*} \text{ for } t \in [\tau_l, \tau_{l+1}], l = 1, \dots, \ell(\epsilon)\},$$

where $\omega_l^{2*} := \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} \omega^{2*}(t) dt$. Let $\tilde{z}(t, h^1)$ be the solution of the following differential equation

$$\dot{\tilde{z}}(t, h^1) = \hat{f}_1(\tilde{z}(t, h^1), \omega^1(t, h^1)) + \hat{f}_2(\tilde{z}(t, h^1), \tilde{\omega}^{2*}(t)). \quad (7.48)$$

By subtracting (7.47) from (7.48) (and having in mind linearity of $\hat{f}^2(z, \omega^2)$ in ω^2), one obtains

$$\|\tilde{z}(\tau_{l+1}, h^1) - z(\tau_{l+1}, h^1)\| \leq \|\tilde{z}(\tau_l, h^1) - z(\tau_l, h^1)\| + \tilde{M}\Delta\|\tilde{z}(t, h^1) - z(t, h^1)\| + \tilde{M}\Delta^2(\epsilon),$$

where \tilde{M} is an appropriate positive constant. Due to Proposition 5.1 in (Gaitsgory, 1992), the latter implies that

$$\|\tilde{z}(\tau_l, h^1) - z(\tau_l, h^1)\| \leq \tilde{M}\Delta(\epsilon) \quad \forall l = 1 \dots, \ell(\epsilon), \quad (7.49)$$

which, in turn, implies that

$$\max_{t \in [0, 1]} \|\tilde{z}(t, h^1) - z(t, h^1)\| \leq \tilde{M}\Delta(\epsilon), \quad (7.50)$$

and, by (7.6),

$$\|G^1(\tilde{z}(1, h^1)) - G^1(z(1, h^1))\| \leq C_G^1 \|\tilde{z}(1, h^1) - z(1, h^1)\| \leq C_G^1 \tilde{M}\Delta(\epsilon). \quad (7.51)$$

Due to the Definition 9 of the Nash equilibrium ,

$$G^1(z(1, h^1)) \geq G^{1*}. \quad (7.52)$$

Consequently,

$$G^1(\tilde{z}(1, h^1)) \geq G^{1*} - C_G^1 \tilde{M}\Delta(\epsilon). \quad (7.53)$$

Let H^1 be a random realization of h^1 . Using the fact that $\hat{f}^i(z, \omega^i)$ ($i = 1, 2$) are Lipschitz

continuous in z and linear in ω^i , one can obtain

$$\begin{aligned}\tilde{z}(\tau_{l+1}, H^1) &= \tilde{z}(\tau_l, H^1) + \int_{\tau_l}^{\tau_{l+1}} \hat{f}^1(\tilde{z}(t, H^1), \omega_1(t, H^1)) dt + \int_{\tau_l}^{\tau_{l+1}} \hat{f}^2(\tilde{z}(t, H^1), \tilde{\omega}_2^*(t)) dt \\ &= \tilde{z}(\tau_l, H^1) + \Delta(\epsilon) \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) + \Delta(\epsilon) \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) + O(\Delta(\epsilon)^2).\end{aligned}\tag{7.54}$$

By subtracting (7.54) from (7.39), one obtains

$$\begin{aligned}\|Z(\tau_{l+1}) - \tilde{z}(\tau_{l+1}, H^1)\| &\leq \|Z(\tau_l) - \tilde{z}(\tau_l, H^1)\| \\ &+ \Delta(\epsilon) \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(Z(t), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \\ &+ \Delta(\epsilon) \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(Z(t), Y^2(t)) dt - \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) \right\| + O(\Delta(\epsilon)^2)\end{aligned}\tag{7.55}$$

By (7.4),

$$\begin{aligned}&\left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(Z(t), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \\ &\leq \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(Z(t), Y^1(t)) dt - \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(\tilde{z}(\tau_l, H^1), Y^1(t)) dt \right\| \\ &+ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(\tilde{z}(\tau_l, H^1), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \\ &\leq \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(\tilde{z}(\tau_l, H^1), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \\ &+ C_1 \Delta(\epsilon) \|Z(\tau_l) - \tilde{z}(\tau_l, H^1)\| + O(\Delta(\epsilon)^2)\end{aligned}\tag{7.56}$$

Analogously,

$$\begin{aligned}&\left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(Z(t), Y^2(t)) dt - \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) \right\| \\ &\leq \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(Z(t), Y^2(t)) dt - \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(\tilde{z}(\tau_l, H^1), Y^2(t)) dt \right\| \\ &+ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(\tilde{z}(\tau_l, H^1), Y^2(t)) dt - \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) \right\| \\ &\leq \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(\tilde{z}(\tau_l, H^1), Y^2(t)) dt - \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) \right\| \\ &+ C_2 \Delta(\epsilon) \|Z(\tau_l) - \tilde{z}(\tau_l, H^1)\| + O(\Delta(\epsilon)^2).\end{aligned}\tag{7.57}$$

For brevity, let us re-denote $u_\epsilon^{i*}(\omega^{i*})$ as u_ϵ^{i*} ($i = 1, 2$). From (7.55), (7.56) and (7.57) it

follows that, for any $u_\epsilon^1 \in \mathcal{U}_1$,

$$\begin{aligned} \mathbb{E}_\zeta^{u_\epsilon^1, u_\epsilon^{2*}} \|Z(\tau_{l+1}) - \tilde{z}(\tau_{l+1}, H^1)\| &\leq \mathbb{E}_\zeta^{u_\epsilon^1, u_\epsilon^{2*}} \left\{ \|Z(\tau_l) - \tilde{z}(\tau_l, H^1)\| (1 + (C_1 + C_2)\Delta(\epsilon)) \right\} \\ &+ \Delta(\epsilon) \mathbb{E}_\zeta^{u_\epsilon^1, u_\epsilon^{2*}} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(\tilde{z}(\tau_l, H^1), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \right. \\ &\left. + \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(\tilde{z}(\tau_l, H^1), Y^2(t)) dt - \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) \right\| \right\} + O(\Delta(\epsilon)^2). \end{aligned} \quad (7.58)$$

Similarly to (7.36), one can obtain

$$\begin{aligned} &\mathbb{E}_\zeta^{u_\epsilon^1, u_\epsilon^{2*}} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^2(\tilde{z}(\tau_l, H^1), Y^2(t)) dt - \hat{f}^2(\tilde{z}(\tau_l, H^1), \omega_l^{2*}) \right\| \right\} \\ &\leq \tilde{L}_2 \frac{\epsilon}{\Delta(\epsilon)} + M_2 \mathbb{E}_\zeta^{u_\epsilon^{2*}} \left\{ \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], 2}^{K(\epsilon)}(H^2; s, a) - \omega_l^{2*}(s, a) \right| \right\} \\ &\leq \tilde{L}_2 \frac{\epsilon}{\Delta(\epsilon)} + M_2 \mathbb{E}_\zeta^{u_\epsilon^{2*}} \left\{ \mathbb{E}_{S_{[\tau_l/\epsilon]}^2} \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], 2}^{K(\epsilon)}(H^2; s, a) - \omega_l^{2*}(s, a) \right| \right\} \\ &\leq \tilde{L}_2 \frac{\epsilon}{\Delta(\epsilon)} + M_2 \mu_2(K(\epsilon)), \end{aligned} \quad (7.59)$$

where \tilde{L}_2 and M_2 are the same constants as in (7.36). In an analogous way,

$$\begin{aligned} &\mathbb{E}_\zeta^{u_\epsilon^1, u_\epsilon^{2*}} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(\tilde{z}(\tau_l, H^1), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \right\} \\ &\leq \tilde{L}_1 \frac{\epsilon}{\Delta(\epsilon)} + M_1 \mathbb{E}_\zeta^{u_\epsilon^1} \left\{ \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], 1}^{K(\epsilon)}(H^1; s, a) - \omega_l^1(s, a; H^1) \right| \right\} \end{aligned} \quad (7.60)$$

where M_1 is the constant from (7.5). From the definition of $\omega_l^1(H^1) = \{\omega_l^1(s, a; H^1)\}$ as *argmin* in (7.46) and from (7.24) (see Lemma 4) it follows that there exists a function $\tilde{\mu}_1 : \mathbb{N} \rightarrow \mathbb{R}$, with $\lim_{K \rightarrow \infty} \tilde{\mu}_1(K) = 0$ (implying $\lim_{\epsilon \rightarrow 0} \tilde{\mu}_1(K(\epsilon)) = 0$) such that

$$\mathbb{E}_\zeta^{u_\epsilon^1} \left\{ \sum_{s,a} \left| \phi_{[\tau_l/\epsilon], 1}^{K(\epsilon)}(H^1; s, a) - \omega_l^1(s, a; H^1) \right| \right\} \leq \tilde{\mu}_1(K(\epsilon)). \quad (7.61)$$

Hence, by (7.60),

$$\mathbb{E}_\zeta^{u_\epsilon^1, u_\epsilon^{2*}} \left\{ \left\| \frac{1}{\Delta(\epsilon)} \int_{\tau_l}^{\tau_{l+1}} f^1(\tilde{z}(\tau_l, H^1), Y^1(t)) dt - \hat{f}^1(\tilde{z}(\tau_l, H^1), \omega_l^1(H^1)) \right\| \right\} \leq \tilde{L}_1 \frac{\epsilon}{\Delta(\epsilon)} + M_1 \tilde{\mu}_1(K(\epsilon)). \quad (7.62)$$

By substituting (7.59) and (7.62) into (7.58), one obtains

$$\begin{aligned} \mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \|Z(\tau_{l+1}) - \tilde{z}(\tau_{l+1}, H^1)\| &\leq (1 + C\Delta(\epsilon)) \mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \left\{ \|Z(\tau_l) - \tilde{z}(\tau_l, H^1)\| \right\} \\ &+ \Delta(\epsilon) \tilde{v}(\epsilon), \end{aligned} \quad (7.63)$$

with $\lim_{\epsilon \rightarrow 0} \tilde{v}(\epsilon) = 0$. By virtue of Proposition 5.1 in (Gaitsgory, 1992), the latter implies that there exists a function $\eta(\epsilon)$, with $\eta(\epsilon) \rightarrow 0$, such that $\forall l = 1, \dots, \ell(\epsilon)$,

$$\mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \|Z(\tau_l) - \tilde{z}(\tau_l, H^1)\| \leq \eta(\epsilon). \quad (7.64)$$

Thus,

$$\left| \mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \{G^1(Z(1))\} - \mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \{G^1(\tilde{z}(1, H^1))\} \right| \leq C_G^1 \eta(\epsilon) \quad \forall u_{\epsilon}^1 \in \mathcal{U}_1. \quad (7.65)$$

By (7.53),

$$\mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \{G^1(\tilde{z}(1, H^1))\} \geq G^{1*} + O(\Delta(\epsilon)),$$

and from (7.65) it follows that

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}_{\zeta}^{u_{\epsilon}^1, u_{\epsilon}^{2*}} \{G^1(Z(1))\} \geq G^{1*} \quad \forall u_{\epsilon}^1 \in \mathcal{U}_1.$$

A similar procedure can be applied to show that

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}_{\zeta}^{u_{\epsilon}^{1*}, u_{\epsilon}^2} \{G^2(Z(1))\} \geq G^{2*}.$$

Since, by Proposition (15),

$$G_i^* = \lim_{\epsilon \rightarrow 0} \mathbb{E}_s^{u_{\epsilon}^{1*}, u_{\epsilon}^{2*}} \{G^i(Z(1))\}, \quad i = 1, 2,$$

the proposition is proved. □

7.5 Conclusion

In this chapter we presented a non-zero sum dynamic game whose state is described by a hybrid dynamical system that evolves in continuous time and that is subjected to abrupt changes of the parameters. By supposing that the MDPs associated to the jumps of the parameters and the macro state of the system evolve with different velocities,

we proved that the state trajectory of the hybrid system can be approximated by the solutions of a certain deterministic averaged control system. We showed how to construct an asymptotic Nash equilibrium Markov policy of the game obtained solving the hybrid system, on the basis of a Nash equilibrium of the game obtained solving the deterministic averaged system.

Conclusions

Conclusions and Future Perspectives

8.1 Summary of Contributions

In this dissertation we investigated different theoretical aspects of EGT, with a particular focus on the notion and the characterization of the player. We followed two major directions: we first revisited the classical idea of individual-selfish players, defining a model for group-players, and we then focused on games combining standard EGT and MDP, which allows one to characterize players through an individual state evolving in time, beside the action used. We also suggested various scenarii which provide useful applications for our models, particularly in telecommunication networks.

More precisely, in Part I we proposed our new approach to evolutionary games to represent those situations in which individuals interact, maximizing the fitness of the group they belong to. Therefore, the actual players of a game are not the interacting individuals but the groups. We studied a simple but meaningful example with two pure actions, the Hawk-Dove game, in two cases, respectively a finite and an infinite population of individuals, divided into a finite number N of symmetric groups of the same size. As the number of players is finite, we considered the Nash equilibrium as solution concept (rather than the ESS). Moreover, since the Nash equilibrium is a multi-action such that no player can profit by deviating, in our group-players context, it is stable against deviations of a share of individuals, since the player consists in a whole group of individuals. We verified that the Hawk-Dove group-players game (in both the finite and the infinite population case) is convex, which allowed us to prove the existence and the uniqueness of the equilibrium through Rosen's conditions ([Rosen, 1965](#)). We then explicitly computed the symmetric mixed Nash equilibria, showing how the presence of

groups plays an important role for the equilibrium structure. In particular, we showed that, for the two group-players game, the level of aggressiveness at the equilibrium is lower than in the standard Hawk-Dove game. We finally defined a gradient-based dynamics, such that the equilibria obtained are asymptotically stable for this dynamics.

We then extended the group players model to a more general framework, where the N groups have different sizes and dispose of a finite set of pure actions. We presented a new solution concept, the GFESS, implying stability against local deviations within each group. For the particular case of two pure actions games, we provided a detailed characterization of the GFESSs and we illustrated our results studying three classical examples considering group-players, the Hawk-Dove game, the Stag Hunt game and the Prisoner's Dilemma, through which we showed that the presence of groups impacts the equilibrium output. We then further generalized our model by redefining the fitness of a group, assuming that the immediate payoff associated with the pairwise interactions among individuals in the same group differs from that of inter-group interactions. We applied this model to a MAC problem, where we studied the impact of groups on mobiles behavior.

In the second part of the dissertation, we considered a particular class of evolutionary games, the MDEG, where each (individual) player is characterized by a state. The action played by an individual not only determines its immediate fitness but it also impacts the transition probabilities to the individual's next state. Furthermore, the fitness of a player does not only depend on its action and on the distribution of actions in the population, but also on the distribution of individual states. We proposed a new dynamical approach to this type of games, focusing on the dynamics involved. We defined the interdependent dynamics of the policies distribution and of the individual states, where the first is assumed to follow the well-known replicator dynamics. For a particular game with two states and two pure actions, we proved the correspondence between stable rest points of the dynamics and the equilibrium profiles of the evolutionary game. Under the assumption that the two dynamics evolve with different timescales, we proposed two methods to obtain the rest points. We gave a complete characterization of these equilibrium profiles and we showed that these equilibria are equivalent in terms of average sojourn times and expected fitnesses. Finally, we illustrated our model through two application scenarios in network systems.

We then defined a particular Hawk-Dove game in a MDEG type of framework, where an individual may be in one of four possible states, representing its age and its strength. In contrast with standard MDEG, in this particular example, we supposed that the transition probabilities of a player do not depend on its action but on the action chosen by

its opponent. By considering stationary policies and supposing that the distribution of the individual states is stationary, we transformed the dynamic Hawk-Dove into a standard evolutionary game and we computed its Nash equilibria. We then further extended our dynamic Hawk-Dove game, by considering group-players as presented in Chapter 2. We showed that, for a two groups game with fixed values of the parameters, the presence of groups in the dynamic Hawk-Dove game brings important novel features with respect to the individual-player game, as the fact that the non-aggressive pure policy can be an equilibrium. We also compared the mixed equilibria obtained respectively for the individual-players and group-players games and we obtained that, also in this case, the probability of being aggressive in strong adult state is lower for group-players.

Finally, we presented a different problem which is not directly related to evolutionary games but which is connected to the dynamic approach adopted for MDEG problems. In particular, we studied a non-zero sum dynamic game whose state is described by a hybrid dynamical system that evolves in continuous time and that is subjected to abrupt changes of the parameters. By supposing that the MDPs associated with the jumps of the parameters and the macro state of the system evolve with different velocities, we proved that the state trajectory of the hybrid system can be approximated by the solutions of a certain deterministic averaged control system. In particular, we show that an asymptotic Nash equilibrium of such hybrid game can be constructed on the basis of a Nash equilibrium of the deterministic averaged dynamic game.

8.2 Perspectives

Group-players evolutionary games leave many issues open for future studies. We have introduced a gradient based dynamics for the Hawk-Dove group-players game. We may define a dynamics in the general group-players framework, comparing different possible dynamics (gradient based, replicator, etc.). Following the folk theorem approach for standard evolutionary games, we could then establish the relation between rest points and equilibria. We may also unify the two lines of research presented here by considering group-players in a MDEG framework, as we did in Chapter 6 for a particular Hawk-Dove dynamic game. The notion of the player can be further investigated. We may consider the case of a population of individuals divided into groups, where the interacting individuals maximize a fitness function which considers both an individual and a group component. It would be interesting also to consider a population which is composed, at the same time, by group-players and individual players, i.e. where some

individuals maximize the fitness of the group they belong to, while others maximize their own fitness, and then to define a suitable notion of equilibrium in this context.

Also the SPcD and, more generally, MDEG problems offer various possible direction for forthcoming developments. In standard MDEG and in our dynamical model we supposed that, respectively, transition probabilities and transition rates only depend on the action of the player. In future works we propose to relax this assumption, letting these functions depend on the actions of both individuals involved in an interaction. Moreover, we may extend our results concerning SPcD to a more general framework, characterizing the states and policies interdependent dynamics for games with more than two states and two pure actions. A major effort should be done in unifying the different classes of stochastic games into a comprehensive theory. As a matter of fact, as we have already mentioned, different classes of games dealing with stochastic processes have been defined in the literature, as stochastic (evolutionary) games (Shapley, 1953) (Flesch et al., 2013), anonymous sequential games (Jovanovic and Rosenthal, 1988), mean-field stochastic games (Weintraub et al., 2005) (Tembine, 2011), mean-field stochastic differential games (Lasry and Lions, 2007) (Guéant et al., 2010). It would be interesting and useful to deeper explore the relation between these types of games and see whether they can all be thought as different subclasses of a more general class of games.

As regards the hybrid stochastic system, in this dissertation we analyzed the *open loop* information structure case, which means that the player has no information about the state of the system, but only on its state and (eventually) its history. We are currently working to extend this model to the *feedback* information structure case, where, at each instant of time both players know exactly to which (macro) state the game has evolved and policies are determined as functions of this information. In order to generalize our results to the closed-loop case, we need to prove that the results in Lemma 4 still hold. Another direction that we would like to investigate in future works is the study of a hybrid stochastic system where the parameters evolve according to CT-MDPs, instead of discrete time processes. This would lead to a framework very close in nature to the one presented in Chapter 5, where we study the interdependent dynamics of individual states and policies distributions and we assume a two timescale behavior of the system. We may thus be able to combine hybrid stochastic dynamical system and EGT.

List of Publications

Journal Papers

- I. Brunetti, V. Gaitsgory, E. Altman, On Non-Zero Sum Game Considered on Solutions of a Hybrid System with Frequent Random Jumps, submitted to Dynamic Games and Applications Journal.
- I. Brunetti, Y. Hayel, E. Altman, State Policy Dynamics in Continuous Time Markov Decision Evolutionary Games, submitted to IEEE Transactions on Automatic Control journal, second round.

Conference Papers

- I. Brunetti, Y. Hayel, E. Altman, State Policy Dynamics in Evolutionary Games, American Control Conference 2015, Chicago.
- I. Brunetti, R. Elazouzi, E. Altman, Altruism in Groups: an Evolutionary Games Approach, NetGCoop NETwork Games COntrol and OPTimization, 2014, Trento. This work has been also presented at the International Conference on Game Theory 2014, Stony Brook.
- I. Brunetti, E. Altman, Revisiting Evolutionary Game Theory, IEEE Conference on Decision and Control 2013, Florence
- E. Altman, I. Brunetti, A Markov Decision Evolutionary Game for the Study of a Dynamic Hawk and Dove Problem, ISDG Workshop 2013, Barcelona.
- I. Brunetti, E. Altman, M. Haddad, Pricing Access to the Internet with Partial

Information, 6th International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS), Cargèse, 2012

List of Figures

2.1	Individuals are divided into a finite number of symmetric groups. Each individual can interact with a member of its own group or of a different one.	14
2.2	The value of p^* as a function of the number of groups $2 \leq N \leq 20$ for three different values of δ . The upper dashed curve is obtained with $\delta = 0.6$, the middle curve with $\delta = 1$ and the lower dotted one with $\delta = 4$	23
2.3	The value of p^* as a function of the number of groups $0 \leq N \leq 20$ for three different values of K and $\delta = 2$. Upper dashed line is obtained with $K = 2$, middle line with $K = 10$ and lower dotted line with $K = 70$	25
2.4	The value of p^* as a function of the size of groups $1 \leq K \leq 50$ for three different values of N . Upper dashed line is obtained with $N = 2$, middle line with $N = 10$, lower dotted line with $N = 70$	25
3.1	The population of individuals is divided into a finite number of groups $G_i, i = 1, \dots, N$. Each individual can interact with a member of its group or of a different one. All individuals in G_i adopts the same strategy q_i (no matter which individual they encounter).	32
3.2	The fully mixed equilibrium actions q^{1*} (upper line) and q^{2*} (lower line), respectively of the first and the second group for $0.3 < \alpha < 0.5$ for the two group-players Hawk-Dove game.	41
3.3	When a mobile k transmits, all mobiles within a circle of radius R centered at node $R(k)$ cause interference to k for its transmission to $R(k)$, so that more than one transmission within a distance R of the receiver in the same slot causes a collision and the loss of mobile's k packet at $R(k)$. A transmitter in group G_i transmits with probability $p^i \in [0, 1], i = 1, \dots, N$	44

3.4	The value of the equilibrium strategy q^* in a two symmetric groups MAC game as a function of γ for $\alpha = 0.4$, compared to the ESS q_{std}^* of the standard game.	48
3.5	The probability $p_s(\mathbf{q}^*)$ of a successful transmission at the equilibrium for the fully mixed GFESS in a two symmetric groups MAC game and in the standard game, as a function of γ	48
3.6	The value of the equilibria of the two groups in a two groups MAC game as a function of γ for $\alpha = 0.4$	49
3.7	The value of the equilibrium strategy of the second group in the fully mixed equilibrium (q^{1*}, q^{2*}) and in the mixed equilibrium (T, q^{T*}) as a function of γ for $\alpha = 0.4$	49
4.1	Each mobile can be in full or in almost empty battery state. In state full, it can choose to transmit at high power or at low power, while, if the battery is almost empty, it can only transmit at low power. If a destination receives simultaneously a transmission from two terminals, packets collide, but one of the packets can still be received correctly if transmitted at a higher power.	55
5.1	A player can choose between policy u_x and u_y . Accordingly with this choice, in state 1 the individual plays, respectively x or y and move to state 0, at a rate μ_x or μ_y . In state 0, under both policies, the only available action is y and the rate at which the player returns to state 1 is μ	72
5.2	Trajectories of the system (S_ϵ) from different starting points and the slow manifold M_0 for $\epsilon = 0.01$	80
5.3	Convergence of the replicator dynamics to the equilibrium $\delta_x^* = 0.71875$, obtained by setting $\mu = 10$, $\mu_x = 1.5$ and $\mu_y = 1$, and $a = -0.3$, $c = 0$, $b = 1$ and $d = 0.5$, starting from $q_x(0) = 0.2$ and $q_x(0) = 0.9$	83
5.4	Energy control in wireless networks. Each mobile continuously moves from full battery state (F) to almost empty battery state (A) and then back to F. If the mobile chooses policy u_h (left side), in state F it transmits at high power (h), otherwise, if it chooses policy u_l it transmits at low power (l) (right side). In state A, it always transmits at low power. Since high power transmission is costly in terms of energy consumption, when transmitting at high power, the transition rate to state A is faster, and thus the time spent in state F is shorter.	88

6.1	The equilibrium q_{v*} plotted as a function of the proportion $P(A)$ of adults in the population for different values of δ . The higher line is obtained with $\delta = 4$, the middle line with $\delta = 6$, the lower one with $\delta = 10$	98
6.2	The equilibrium q_{v*} as a function of the value of δ for different values of $P(A)$. The higher line is obtained with $P(A) = 0.15$, the middle line with $P(A) = 0.5$, the lower one with $P(A) = 0.88$	98
6.3	The probability of being aggressive in state A_S at the equilibria, obtained with $N = 2$, $P(A) = 0.5$ and $\Delta = 1$, as a function of the cost δ	103
6.4	The equilibrium value q_{u*} as a function of the value of δ	103

Bibliography

- Abbad, M. and Filar, J. (1992). Perturbation and stability theory for Markov control problems. *IEEE Trans. Automat. Control*, 37:1415–1420. [107](#), [108](#)
- Altman, E., El-Azouzi, R., Hayel, Y., and Tembine, H. (2008a). An evolutionary game approach for the design of congestion control protocols in wireless networks. In *Physicomnet workshop, Berlin, April 4*. [11](#)
- Altman, E., El-Azouzi, R., Hayel, Y., and Tembine, H. (2008b). Evolutionary power control games in wireless networks. In *7th IFIP Networking*. [xiii](#)
- Altman, E., El-Azouzi, R., Hayel, Y., and Tembine, H. (2009). The evolution of transport protocols: An evolutionary game perspective. *Computer Networks*, 53(10):1751–1759. [xiii](#), [11](#)
- Altman, E. and Gaitsgory, V. (1995). A hybrid differential stochastic zero sum game with fast stochastic part. *Annals of the International Society of Dynamic Games*, 3:47–59. [107](#)
- Altman, E. and Gaitsgory, V. (1997). Asymptotic optimization of a nonlinear hybrid system controlled by a Markov decision process. *SIAM J. Control Optim.*, 35(6):2070–2085. [107](#), [108](#), [114](#)
- Altman, E. and Gaitsgory, V. A. (1993). Control of a hybrid stochastic system. *Systems Control Lett.*, 20:307–314. [107](#)
- Altman, E. and Hayel, Y. (2008). Stochastic evolutionary game approach to energy management in a distributed aloha network. In *IEEE INFOCOM*. [xiii](#), [15](#), [55](#), [63](#), [86](#), [91](#)

- Altman, E. and Hayel, Y. (2010). Markov decision evolutionary games. *IEEE Transactions on Automatic Control*, 55(6). [v](#), [xiv](#), [xv](#), [53](#), [56](#), [59](#), [62](#), [70](#), [81](#)
- Beggs, A. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory*, 122(1). [62](#)
- Bensoussan, A. (1989). *Perturbation Methods in Optimal Control Problems*. John Wiley, New York. [107](#), [108](#)
- Bensoussan, A. and Blankenship, G. (1987). Singular perturbations in stochastic control. *Singular Perturbations and Asymptotic Analysis in Control Systems*, pages 171–260. Lecture Notes in Control and Inform. Sciences 90, Springer-Verlag, New York. [107](#), [108](#)
- Bielecki, D. and Filar, J. (1991). Singularly perturbed Markov control problem: Limiting average cost. *Ann. Oper. Res.*, 28:153–168. [107](#), [108](#)
- Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97:298–319. [4](#)
- Bonneau, N., Altman, E., Debbah, M., and Caire, G. (2005). An evolutionary game perspective to ALOHA with power control. In *19th International Teletraffic Congress, Beijing, China*. [xiii](#)
- Borgers, T. and Sarin, R. (1997). Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77. [62](#)
- Brown, G. and Neumann, J. V. (1950). *Solutions of Games by Differential Equations*. Princeton University Press. In H. Kuhn and A. Tucker, editors, Contributions to the Theory of Games I, Annals of Mathematics Studies 24. [8](#)
- Brunetti, I. and Altman, E. (2013). Revisiting evolutionary game theory. In *IEEE 52nd Annual Conference on Decision and Control, Florence*. [20](#)
- Cressman (1992). *The Stability Concept of Evolutionary Game Theory. A Dynamic Approach*. Springer-Verlag. [6](#), [11](#), [16](#), [30](#)
- Cressman, R. (2003). *Evolutionary dynamics and extensive form games*. Cambridge, Mass, MIT press. [10](#)
- Cressman, R., Garay, J., and Hofbauer, J. (2001). Evolutionary stability concepts for n-species frequency-dependent interactions. *J. theor. Biol.*, 211:1–10. [6](#), [30](#)

- Crowley, P. H. (2000). Hawks, doves, and mixed-symmetry games. *J. theor. Biol.*, 204:543–563. [11](#)
- Delebecque, F. and Quadrat, J. (1978). Contribution of stochastic control singular perturbation averaging and team theories to an example of large scale systems: Management and hydropower production. *IEEE Trans Automat. Control*, AC-23:209–222. [107](#), [108](#)
- Derman, C. (1970). *Finite State Markovian Decision Processes*. Academic Press, New York. [111](#)
- Enquist, M. and Ghirlanda, S. (2007). Evolution of social learning does not explain the origin of human cumulative culture. *Journal of theoretical biology*. [xiii](#)
- Filar, J. and Raghavan, T. (1984). A matrix game solution of the single-controller stochastic game. *Mathematics of Operations Research*, vol. 9(no. 3). [53](#)
- Filar, J. and Vrieze, K. (1997). *Competitive Markov Decision Processes*. Springer New York. [54](#)
- Fisher, R. A. (1930). *The Genetic Theory of Natural Selection*. Oxford, Clarendon Press. [3](#)
- Flesch, J., Parthasarathy, T., Thuijsman, F., and Uyttendaele, P. (2013). Evolutionary Stochastic Games. *Dynamic Games and Applications*, 3(2):207–219. [54](#), [128](#)
- Friedman, D. (1991). Evolutionary games in economics. *Econometrica*, 59:637–666. [xiii](#), [4](#)
- Fudenberg, D. and Levine, D. (1998). *Theory of Learning in Games*. MIT press. [8](#), [63](#)
- Fudenberg, D. and Tirole, J. (1991). *Game theory*. MIT Press, Cambridge, MA. [8](#)
- Gaitsgory, V. (1992). Suboptimization of singularly perturbed control systems. *SIAM J. Control Optim.*, 30:1228–1249. [117](#), [118](#), [119](#), [122](#)
- Garay, J. and Varga, Z. (2000). Strict ESS for n-species systems. *BioSystems*, 56:131–137. [6](#), [29](#), [30](#)
- Gilboa, I. and Matsui, A. (1991). Social stability and equilibrium. *Econometrica*, 3:859–67. [8](#)
- Guéant, O., Lasry, J., and Lions, P. (2010). Mean field games and applications. *Springer: Paris-Princeton Lectures on Mathematical Finance*, Eds. René Carmona, Nizar Touzi. [128](#)

- Guo, X. and Hernandez-Lerma, O. (2009). *Continuous Time Markov Decision Process*. Springer. [64](#)
- Hamilton, W. (1967). Extraordinary sex ratio. *Science*, 156:477–88. [3](#)
- Hanauske, M., Kunz, J., Bernius, S., and Konig, W. (2010). Doves and hawks in economics revisited: An evolutionary quantum game theory based analysis of financial crises. *Physica A: Statistical Mechanics and its Applications*, 389(21):5084–5102. [11](#)
- Harmsand, W. and Skyrms, B. (2008). *Evolution of Moral Norms*. in The Oxford Handbook of Philosophy of Biology, Oxford University Press. [xiii](#)
- Haurie, A. and Marcotte, P. (1985). On the relationship between Nash-Cournot and Wardrop equilibria. *Networks*, 15:295–308. [22](#)
- Hofbauer, J. and Sandholm, W. (2008). *Stable population games and integrability for evolutionary dynamics*. Unpublished manuscript, University of Vienna and University of Wisconsin. [19](#)
- Hofbauer, J., Schuster, P., and Sigmund, K. (1979). A note on evolutionary stable strategies and game dynamics. *J. Theor. Biology*, 81:609–612. [10](#)
- Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK. [6](#), [7](#), [62](#)
- Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *American Mathematical Society*, 40(4):479–519. [6](#), [10](#)
- Hoppensteadt, F. (1967). Stability in systems with parameters. *J. Math. Anal. Appl.*, 18:129–134. [67](#), [69](#)
- Houston, A. and McNamara, J. (1988). Fighting for food: a dynamic version of the Hawk-Dove game. *Evolutionary Ecology*, 2:51–64. [11](#), [91](#)
- Houston, A. and McNamara, J. (1991). Evolutionary stable strategies in the repeated Hawk-Dove game. *Behavioral Ecology*, 2(3):219–227. [91](#)
- Ibe, O. (2008). *Markov Processes for Stochastic Learning*. Academic Press. [70](#)
- Jackson, M. (2005). *A Survey of Network Formation Models: Stability and Efficiency*. Group Formation in Economics: Networks, Clubs, and Coalitions book, Cambridge University Press. [88](#)
- J.Hofbauer and K.Sigmund (1988). *The Theory of Evolution and Dynamical Systems*. Cambridge Univ. Press. [30](#)

- Jovanovic, B. and Rosenthal, R. (1988). Anonymous sequential games. *Journal of Mathematical Economics*, 17:77–87. [53](#), [54](#), [128](#)
- Kallenberg, L. C. M. (1983). *Linear Programming and Finite Markovian Control Problems*, volume 148. Mathematical Centre Tracts, Amsterdam. [57](#)
- Kokotovic, P., Khalil, H., and O'Reilly, J. (1986). Singular perturbation methods in control. *SIAM*. [67](#), [69](#), [76](#), [77](#), [107](#), [108](#)
- Kokotovic, P. V. (1984). Applications of singular perturbations techniques to control problems. *SIAM Review*, 26:501–550. [108](#)
- Kuran, T. and Sandholm, W. H. (2008). Cultural integration and its discontents. *Review of Economic Studies*, 75:201–228. [4](#)
- Kushner, H. (1990). *Weak Convergence and Singularly Perturbed Stochastic Control and Filtering Problems*. Birkhauser, Boston. [107](#)
- Lasry, J. and Lions, P. (2007). Mean field games. *J. Math.*, 2:229–260. [128](#)
- Levinson, N. (1950). Perturbations of discontinuous solutions of non-linear systems of differential equations. *Acta Math.*, 82:71–106. [67](#), [69](#)
- Lewontin, R. C. (1961). Evolution and the theory of games. *J. Theor. Biol.*, 1(3):382–403. [3](#)
- Mailath, G. (1992). Introduction: Symposium on Evolutionary Game Theory. *Journal of Economic Theory*, 57:259–277. [4](#)
- Maynard Smith, J., , and Parker, G. (1976). The logic of asymmetric contests. *Anim. Behav.*, 24:159–75. [6](#)
- Maynard Smith, J. (1972). Game theory and the evolution of fighting. *J. Maynard Smith, On Evolution* (Edinburgh: Edinburgh University Press), pages 8–28. [iv](#), [xii](#), [10](#)
- Maynard Smith, J. (1982). *Evolution and the theory of Games*. Cambridge University Press, UK. [3](#), [5](#), [63](#)
- Maynard Smith, J. and Price, G. R. (1973). The logic of animal conflict. *Nature*, 246:15 – 18. [v](#), [xiii](#), [3](#)
- McNamara, J., Merad, S., and Collins, E. J. (1991). The Hawk-Dove game as an average cost problem. *Advances of Applied Probability*, 23:667–682. [91](#)

- Menasche, D., Figueiredo, D., and de Souza e Silva, E. (2005). An evolutionary game-theoretic approach to congestion control. *Performance Evaluation*, 62:295–312. [xiii](#)
- Nachbar, J. H. (1990). Evolutionary selection dynamics in games: Convergence and limit properties. *International Journal of Game Theory*, 19:59–89. [4](#)
- Nash, J. (1950). *Non-cooperative Games*. PhD thesis, Princeton University. [iv](#), [xiii](#)
- Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behaviour*. Princeton University Press. [8](#)
- Nguyen, M. T., Altman, E., and Gaitsgory, V. (2001). On stochastic hybrid zero-sum games with nonlinear slow dynamics. *Annals of the International Society of Dynamic Games*, 6:129–145. [107](#), [108](#)
- Nowak, M., Plotkin, J., and Krakauer, D. (1999). The evolutionary language game. *Journal of Theoretical Biology*, (200):147–162. [xiii](#)
- O'Malley, R. (1971). Boundary layer methods for nonlinear initial value problems. *SIAM Review*, 26(4):425–434. [67](#)
- O'Malley, R. E. (1974). *Introduction to singular perturbations*. Academic Press, New-York. [108](#)
- Pervozvansky, A. and Gaitsgory, V. (1988). *Theory of Suboptimal Decisions*. Kluwer. [107](#), [108](#)
- Philips, R. G. and Kokotovic, P. (1981). A singular perturbation approach to modeling and control of Markov chains. *IEEE Trans. Automat. Control*, pages 1087–1094. [107](#), [108](#)
- Pohley, H. and Thomas, B. (1983). Non-linear ESS-models and frequency dependent selection. *Biosystems*, 16(2):87–100. [6](#)
- Puterman, M. (1994). *Markov Decision Processes*. Wiley. [70](#)
- Rosen, J. (1965). Existence and uniqueness of equilibrium points for concave n -person games. *Econometrica*, 33:520–534. [xv](#), [8](#), [14](#), [19](#), [20](#), [21](#), [23](#), [26](#), [27](#), [125](#)
- Sandholm, W. (2005). Negative externalities and evolutionary implementation. *Review of Economic Studies*, 72:885–915. [4](#)
- Sandholm, W. (2009). Evolutionary Game Theory. *Encyclopedia of Complexity and Systems Science - Springer*. [7](#)

- Sandholm, W. H. (2001). Potential games with continuous player sets. *Journal of Economic Theory*, 97:81–108. [4](#)
- Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. MIT Press. [xiv](#), [8](#), [16](#), [30](#), [63](#)
- Schuster, P. and Sigmund, K. (1983). Replicator dynamics. *Journal of Theoretical Biology*, 100:533–538. [3](#), [8](#)
- Selten, R. (1980). A note on evolutionarily stable strategies in asymmetric animal conflicts. *J. Theor. Biol.*, 84:93–101. [6](#)
- Selten, R. (1991). Evolution, learning, and economic behavior. *Games and Economic Behavior*, 3:3–24. [4](#)
- Shakkottai, S., Altman, E., and Kumar, A. (2006). The case for non-cooperative multi-homing of users to access points in ieee 802.11 wlans. *IEEE Infocom, Barcelona, Spain*. [xiii](#), [4](#)
- Shapley, L. (1953). Stochastic games. *Proceedings National Academy of Science USA*, 39:1095–1100. [53](#), [128](#)
- Shi, P., Altman, E., and Gaitsgory, V. (1998). On asymptotic optimization of a class of nonlinear stochastic hybrid systems. *Mathematical Methods of Operations Research*, 47:289–315. [108](#), [116](#)
- Slobodkin, L. and Rapoport, A. (1974). An optimal strategy of evolution. *Q. Rev. Biol.*, 49:181–200. [3](#)
- Taylor, P. (1979). Evolutionarily stable strategies with two types of player. *Journal of Applied Probability*, 16(1):76–83. [6](#), [15](#)
- Taylor, P. and Jonker, L. B. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156. [v](#), [xiii](#), [3](#), [6](#), [7](#)
- Tembine, H. (2011). Mean field stochastic games: Convergence, q/h- learning and optimality. In *American Control Conference, San Francisco, California, USA*. [128](#)
- Tembine, H., Boudec, J. Y. L., ElAzouzi, R., and Altman, E. (2009). Mean field asymptotic of Markov Decision Evolutionary Games. *International IEEE Conference on Game Theory for Networks, Gamenets*. [54](#)
- Thomas, B. (1984). Evolutionary stability: States and strategies. *Theor. Pop. Biology*, 26(1):49–67. [6](#)

- Thomas, B. (1985). On evolutionarily stable set. *J. Math. Biology*, 22:105–115. [6](#)
- Tomassini, M., Pestelacci, E., and Luthi, L. (2010). Mutual trust and cooperation in the evolutionary hawks-doves game. *BioSystems*, 99(1):50–59. [11](#)
- Weibull, J. (1995). *Evolutionary Game Theory*. MIT Press. [5](#), [6](#)
- Weibull, J. (1998). What have we learned from evolutionary games so far? *Working paper 487. Research Institute of Industrial Economics, Stockholm*. [iv](#), [xiii](#), [xvi](#)
- Weintraub, G., Benkard, L., and Roy, B. V. (2005). Oblivious equilibrium: A mean field approximation for large-scale dynamic games. *Advances in Neural Information Processing Systems*, 18. [128](#)
- Wiecek, P. and Altman, E. (2014). Stationary anonymous sequential games with undiscounted rewards. *Journal of Optimization Theory and Applications*. [89](#)
- Zhang, J. (2004). A dynamic model of residential segregation. *Journal of Mathematical Sociology*, 28:147–170. [4](#)
- Zheng, Y. and Feng, Z. (2001a). Evolutionary game and resources competition in the internet. In *Modern Communication Technologies, SIBCOM, The IEEE-Siberian Workshop of Students and Young Researchers*, pages 51–54. [xiii](#), [4](#)
- Zheng, Y. and Feng, Z. (2001b). Evolutionary game theoretic model and end-to-end flow control in the internet. In *ICII, Beijing*. [xiii](#)

Index

- Asymptotic Nash Equilibrium (ANE), [110](#)
- Asymptotic stability, [10](#)
- Average sojourn time, [65](#)
- Deterministic policies, [56](#)
- Diagonally strictly concave function, [19](#)
- Evolutionarily Stable State, [6](#)
- Evolutionarily Stable Strategy, [4](#)
- Folk theorem of evolutionary game theory, [9](#)
- Group Fitness Evolutionarily Stable Strategy (GFESS), [31](#)
- Hawk-Dove game, [10](#)
- History, [56](#)
- Hybrid Stochastic System, [108](#)
- Interior orbit, [10](#)
- Liapunov stability, [10](#)
- Macro state, [109](#)
- Markov Decision Evolutionary Games (MDEG), [58](#)
- Markov Decision Process (MDP), [56](#)
- Markov policies, [56](#)
- Micro state, [109](#)
- Multiaction, [17](#)
- Multiple Access Control (MAC), [42](#)
- Normal Form Symmetric Games, [4](#)
- Occupation measures, [57](#)
- Policy, [56](#)
- Population games, [6](#)
- Pseudogradient, [19](#)
- Replicator Dynamics, [7](#)
- rest point, [9](#)
- Revision protocol, [8](#)
- Rosen's Conditions, [19](#)
- Singular perturbation model, [67](#)
- Stationary policies, [56](#)
- Stationary system state, [58](#)
- Stochastic games, [54](#)
- Transition probabilities, [56](#)