Network games and strategic play

Social influence, cooperation and exerting control

Alain Govaert

The research described in this dissertation has been carried out at the Faculty of Science and Engineering, University of Groningen, the Netherlands.

The research reported in this dissertation is part of the research program of the Dutch Institute of Systems and Control (DISC). The author has successfully completed the educational program of DISC.

university of groningen

Network games and strategic play

Social influence, cooperation and exerting control

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus, Prof. E. Sterken,
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Friday 14 February 2020 at 12.45 hours

by

Alain Govaert

born on 14 April 1989
in Weert, the Netherlands

Supervisors
Prof. M. Cao
Prof. J.M.A. Scherpen

Assessment committee
Prof. T. Başar
Prof. D. Bauso
Prof. A. Flache

To my late father, an inspiring examplar

Acknowledgments

Let me begin by thanking my supervisor Ming Cao for allowing me to freely explore my curiosity during the last four years while also pushing me to expand my capabilities. If it was not for this freedom, your broad research perspectives and "subtle" management style, I doubt that I would have found the passion that I feel today for doing research. As a trained "industrial manager", I had a lot to learn about doing research *your way*, that, the way I see it, requires a strict academic attitude, open mind and technical skill (not necessarily in that order). Because I believe we are never done learning, I will continue to develop these aspects both in life and work, and add a personal touch to it, to make it my way. You once told me that I can be quite stubborn, perhaps this was also part of the reason why I could explore topics freely. Nevertheless, I truly hope you feel proud of some of the work that we have done together.

To my second supervisor Jacquelien Scherpen. Even though our research topics were quite different, you have always shown an interest in my progress and research topics. I am happy that you have added human behavior to your research portfolio as well. In the future, I will closely follow how you will approach challenging engineering problems with humans in the loop. Perhaps most of all, I would like to thank you for your positive spirit throughout the last four years and how this reflected in a very pleasant and social research group with you as the head.

I would also like to thank Michael Mäs. In our meetings, discussions and lab work, you have inspired me to look at decision-making processes from different perspectives. Going from equations to the real-world and back is a very challenging but inspiring task. Things that I used to take for granted, I now question. Naturally, this can be frustrating, but I am sure that this attitude is critical for obtaining new interesting insights. I hope in the future we will be able to finish the work that we have started.

I would also like to address some words to my co-authors. Many thanks to Pouria Ramazi, for the guidance at the beginning of my Ph.D. project and introducing me to game theory, evolutionary game theory, and network games. Even though I knew very little, your patience and enthusiasm always made our discussions enjoyable. I have especially put your patience to the test, promising to finish our paper time after time, and failing to do so, time after time. I am glad that afterwards you felt it was worth the wait. To Yuzhen Qin, I can still remember the first day that we met at the welcome day for new Ph.D. candidates. It has been great to share the typical Ph.D. “burdens” with you. To Carlo Cenedese and Sergio Grammatico, thank you for being open to new ideas and being enthusiastic about our joint work. There are many challenging open problems ahead.

Alain Govaert
Groningen
September, 2019

Contents

List of symbols and acronyms

$\mathbb{R}$	set of real numbers
$\mathbb{R}_{>}$	set of real positive numbers
$\mathbb{R}_{\geq}$	set of real nonnegative numbers
$\mathbb{Z}$	set of integers
$\mathbb{1}_n$	n-dimensional vector of all ones
$\mathcal{G}$	graph
$\mathcal{V}$	vertex set
$\mathcal{E}$	edge set
$\mathcal{N}_i$	the set of neighbors of i excluding i
$\bar{\mathcal{N}}_i$	the set of neighbors of i including i
N	Total number of players
n	groupsize of multiplayer game
Γ_{f}	finite game
Γ_{c}	convex game
$\boldsymbol{\mathcal{A}}$	action space
$\mathcal{A}_i$	Action set of player i
$\mathcal{F}_i$	Feasible action set of player i
$\boldsymbol{\sigma}$	action profile
$\boldsymbol{\mathcal{S}}$	action space of a strategically differentiated game
$\boldsymbol{s}$	action profile in a strategically differentiated game
π	combined payoff function
$\mathbf{p}$	memory-one strategy specifying cooperation probabilities
δ	continuation probability or discount factor
p_0	initial probability to cooperate
ϕ	Scaling parameter of zero-determinant strategy
s	slope of the zero-determinant strategy
l	Baseline of the zero-determinant strategy

AFIP	Approximate Finite Improvement Property
BR	Rest Response
ϵ-NE	Approximate Nash Equilibrium
ϵ-GNE	Approximate Generalized Nash Equilibrium
FIP	Finite Improvement Property
GNE	Generalized Nash Equilibrium
NE	Nash Equilibrium
PD	Prisoner's Dilemma
pTFT	Proportional Tit-for-Tat
PGG	Public Goods Game
RBR	Relative Best Response
RSP	Rock-Scissors-Paper
TFT	Tit-for-Tat
ZD	Zero-Determinant

CHAPTER 1

Introduction

Reciprocity is certainly not a good basis for a morality of aspiration. Yet it is more than just the morality of egoism.

Robert Axelrod

1.1 Background

1.1.1 Social dilemmas

In [1], social dilemmas are broadly defined to be situations that involve conflicts between immediate self-interest and longer-term collective interests. These situations are complex psychological, social and economic behaviors because the immediate self-interests make it tempting for individuals to choose selfish decisions that in the longer term become detrimental to the collective and possibly to themselves. A classical example is known as *the tragedy of the commons*, in which individual users in a shared resource system, by acting in their self-interests, deplete a resource through their collective actions [2,3]. Although the theory originated almost 200 years ago [4], the tragedy of the commons, and in a broader sense, social dilemmas, remain relevant for today's societal concerns. From over-fishing, global warming, smoking in public

Figure 1.1: The social dilemma and tragedy of the commons in over-fishing. By Cardow, The Ottawa Citizen.

places to the more recent social dilemma of autonomous vehicles [5]. All of these situations, to some extent, affect our day-to-day lives.

While some common resource systems have indeed collapsed due to overuse, for others the tragedy of the commons was averted through cooperation, regulation or some other mechanism that enables to *govern the commons* [6]. Knowledge of social dilemmas can thus help in understanding when personal interests are set aside for selfless cooperative behavior and under which conditions cooperation in large groups and organizations can be maintained or even promoted [1].

Because social dilemmas come in all sorts and sizes, and obtaining a uniform understanding of the consequences of individual choice and collective behavior is desirable, it is necessary to apply a unifying framework in which the large variety of social dilemmas can be studied formally. A defining feature of *game theory* [7] is that outcomes of decision-making processes or *games* do not only depend on one's own decision, but also on that the decisions of others. It is precisely this characteristic feature that makes game theory a suitable modeling framework for social dilemmas. The *prisoner's dilemma* is the most simple and widely studied game that captures a social dilemma between two individuals that *simultaneously* choose between two actions: to cooperate or defect. The *payoffs*, in this game are

$$\begin{bmatrix} R & S \\ T & P \end{bmatrix}, \quad T > R > P > S.$$

In the case of the prisoner's dilemma defection refers to betraying the other prisoner, while cooperation refers to staying silent. When both players cooperate, they receive the reward for mutual cooperation (R). When both defect, they receive the punishment for mutual defection P. When one player defects, while the other cooperates, the cooperator who kept silent is betrayed by his/her assailant and receives the sucker's payoff (S), while the defector obtains the temptation to defect (T). This classic game has a single *Nash equilibrium* [9, 10] at which both players make the rational decision to defect because this action receives a higher payoff independent of what the other player chooses (i.e. $T > R$ and $P > S$). To see that the prisoner's dilemma is indeed a social dilemma, notice that if the two prisoners neglected their self-interests and would choose to cooperate, they would receive R, that is higher than the payoff received when both players are selfish and choose to defect, i.e. $R > P$.

Social dilemmas do not always have a dominant strategy (like defection in the prisoner's dilemma), and there can exist more than just one equilibrium. A simple example is the game of Chicken, the Hawk-Dove game or Snowdrift game in which the payoffs satisfy $S > P > T > R$. There may also be more than two players, in this case, the game is called a *multiplayer* or n-player game. A famous example is the public goods game in which players need to decide to contribute to a publicly available good. Multiplayer games are interesting because they can capture the collective behavior of a large group of decision-makers.

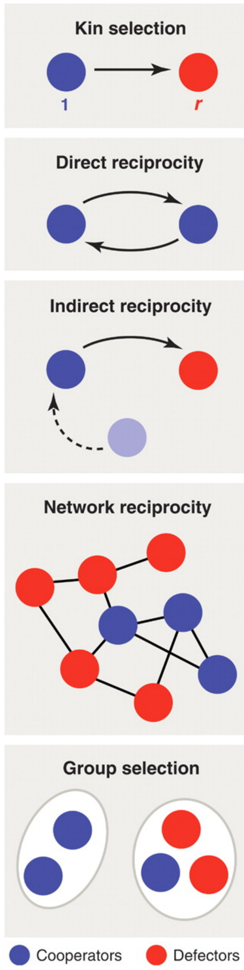


Figure 1.2: Mechanisms for the evolution of cooperation in social dilemmas from [8], reprinted with permission from AAAS.

In the simple static models described above, the emergence of this selfless cooperative behavior is impossible to achieve. However, when decisions are repeated, social structure or individual sanctioning is added, several solutions to social dilemmas present themselves. Through individual sanctioning, cooperators can be rewarded and defectors can be punished by the players themselves or an overarching institution. Punishment and reward have proven to be effective in promoting cooperation both experimentally and theoretically [11–14]. Punishment and rewards are related to *indirect reciprocity* [8, 15–18], through which cooperators enjoy good reputations, while defectors have bad reputations. Indirect reciprocity relies on the assumption that players are inclined to cooperate against players who have cooperated before, and thus have a good reputation. Provided that this reputation information is available to the players, indirect reciprocity, as the name suggests, enables cooperative actions to be played against "strangers", i.e. players that an individual has not interacted with in the past (Fig. 1.2). Empirical evidence of indirect reciprocity can be found in [19, 20]. From a more evolutionary point of view the mechanisms known as *kin selection* [21–25] and *group selection* [26–28] have been proposed as a means for promoting the evolution of selfless cooperation. Under kin selection, the relatedness between individuals, defined by the probability of sharing a gene, affects the behavior of the individual against their kin: cooperative actions are more likely if relatedness between individuals is higher. This idea supports the concept of *inclusive fitness*, in which payoffs, or the more biological term *fitness*, are evaluated by including the effect actions may have on closely related individuals or *kin* [8]. Inclusive fitness or kin selection is where the concept of *selfish genes* [24] comes from: cooperation against kin increases their fitness and hence increases the reproductive rate and spread of closely related genes. Under group selection, the natural selection forces act not only on individuals but also on groups: groups of cooperators can obtain a higher payoff than groups of defectors and can therefore grow and split into multiple groups faster than groups of defectors [8].

This thesis will not cover all of these mechanisms for the evolution of cooperation. Rather we will focus on structural and strategic solutions to social dilemmas that can allow for cooperative actions to evolve through *network reciprocity* and *direct reciprocity*, respectively (Fig. 1.2). These mechanisms are introduced in the following sections.

Structural solutions: network reciprocity

In its original application to an evolving biological population, *evolutionary game theory* [29] describes how competing strategies propagate through a well-mixed *population* via natural selection. In such a well-mixed population, all players interact equally likely with all other players [29–33]. In real populations, individuals often interact with each other via spatial or social structures that tend to be very different between individuals. These effects can be captured by *evolutionary graph theory* [34], that allows to study how spatial and social structures affect evolutionary dynamics [35, 36, 36–38]. The majority of these works focus on formulating conditions for *evolutionary success* of structured populations in which the micro-dynamics describe *birth-death* and *imitation* processes of the players occupying the nodes of the graph. In social dilemmas this evolutionary success depends on the emergence and maintenance of cooperation in the population [39]. The mechanism that allows cooperation to exist in evolutionary games on graphs is known as *network reciprocity* [8]: clusters of cooperators can form in the network in which the mutual cooperative actions help each other (Fig. 1.2). Unfortunately, evolutionary games on graphs are difficult to analyze mathematically because of the large number of configurations that are possible. When individuals interact in pairs and they have no more than two strategies, the conditions for evolutionary success can be characterized analytically by benefit-to-cost ratios and the average degree of the network [38, 40].

Evolutionary games on graphs stay true to their original application to evolving biological populations via natural selection and hence mainly study dynamical processes based on *replication*. In contrast, *network games* take a more economic perspective and typically describe how individual decision-makers change their actions over time under the *bounded rationality* principles [41]: even when individuals intend to make *rational* decisions, limitations on cognitive capacity or available information might limit their ability to make optimal decisions in complex situations. In this economic context, "evolutionary" dynamics driven by simple rational thinking, (e.g. *myopic best response*) have been studied extensively for games on networks using potential functions [42–44] and Markov chain theory [45, 46], and brought forth a number of algorithms that ensure convergence to an equilibrium [47–49]. However, as we have seen in the tragedy of the commons, myopic optimizations tend to generate outcomes with payoffs that are far from the system optimum [50]. Hence, under these rationality principles, network reciprocity is less effective. We will return to this problem in part I of the thesis.

Strategic solutions: direct reciprocity

We have seen that the only rational decision in the prisoner's dilemma game is to defect. However, when the prisoner's dilemma game is repeated, decisions become more cooperative [51]. *Repeated games* allows us to formalize how reciprocity [52] can influence the behavior of the players. In repeated games the reciprocity effects occur between the same set of players and hence, the mechanism that allows cooperative decisions to emerge is called *direct reciprocity*. Repeated games can capture a variety of complicated trade-offs in decision-making processes. For instance, players can learn from past decisions and adjust their behavior accordingly. Indeed, a *strategic* player would base his or her decision on what to do *now*, by taking into account what happened before. This allows for a variety of strategies that differ in memory, rewards, punishments, fairness, etc.

Another interesting process that is captured by repeated games is how one's current actions can affect *future* interactions and their associated payoffs. If one would consider to defect at some point in time, how large will the consequences of retaliation be? Is the fear of retaliation enough to remain cooperative? These strategic trade-offs are sometimes referred to as "*the shadow of the future*" and can be studied using discounting techniques. Direct reciprocity is only effective when the shadow of the future is uncertain. To see this, let us assume the players know they will interact in $0 < k < \infty$ rounds and payoffs are not discounted. Regardless of what happened in the $k-1$ rounds before, at round k the only rational choice is to defect because there will be no future play and hence no opportunities for retaliation. Under the rationality principle, both players will thus choose to defect at round k. Knowing this, the action made at the penultimate round $k-1$ cannot affect the actions at round k and defection strongly dominates cooperation. Hence, the players will choose to defect at $k-1$ as well. An induction argument shows that defection in all rounds is the only equilibrium. [53, 54]. The repeated prisoner's dilemma with an undetermined number of rounds (possibly finite) has many different equilibria. The famous *folk theorem* guarantees that any feasible average payoff can be obtained at an equilibrium, as long as the players obtain at least the mutual defection payoff [55]. However, in evolving populations these equilibria are not *evolutionarily stable* [29] i.e., the equilibrium strategies can be invaded by a mutant strategy that performs better [56–58]. This motivated researchers to identify strategies that perform well under a variety circumstances [59–62]. Perhaps the most famous of these strategies is known as *Tit-for-Tat* (TFT), in which players simply repeat the action that their co-player chose in the previous round. Next to TFT's ability to let cooperation evolve, in [63] it was shown that TFT is "*unbeatable*" in the class of exact potential games (See Preliminaries chapter), that includes all symmetric games with two players and two actions. This means that no other strategy can get strictly more than a player applying the simple imitations of the TFT rule. This rather surprising result can be placed into the broader context of *Zero-determinant strategies* (ZD) [64]. ZD strategies can enforce a linear payoff relation between the ZD strategist and their co-players. The n-player version of TFT known as proportional-TFT (pTFT) is a *fair* ZD strategy. That is, it enforces that the average payoffs of all players are equal. If pTFT is applied to a 2-player game, it naturally recovers the classic TFT strategy, implying that TFT is unbeatable.

In part II, we will investigate the existence, efficiency and evolutionary stability of ZD strategies under a variety of circumstances.

1.2 Contributions and thesis outline

Part I: rationality and social influence in network games

The contributions in Part I are mainly concerned with how network reciprocity can result in *rational cooperation* in social dilemmas on networks. New decision-making rules are introduced that combine rational economic behavior with social learning by imitation and a mechanism called *strategic differentiation* is introduced.

Chapter 3

The role of human decision making is becoming increasingly important for complex engineering systems. More often than not, this social behavior of large groups of humans is modeled based on *rationality*. However, behavioral and experimental economics suggest that humans are not always rational and our decisions are likely to be influenced by a form of *social learning* in which new behaviors result from imitation. In this chapter, novel evolutionary dynamics for network games are proposed, called the *h-Relative Best Response* (h-RBR) dynamics, that result from an intuitive mixture of rational Best Response (BR) and social learning by imitation. Under such a class of dynamics, the players optimize their payoffs over the set of actions employed by relatively successful neighbors. As such, the h-RBR dynamics share the defining non-innovative characteristic of imitation based dynamics that can lead to equilibria that differ from classic Nash equilibria. We study the asymptotic behavior of the h-RBR dynamics for both finite and convex games and provide preliminary sufficient conditions for finite-time convergence to an (approximate) generalized Nash equilibrium. We then couple the results to those obtained for classic best response dynamics and show how a mixture of rational best responding individuals and h-relative best responders can affect the equilibria of fundamental economic and behavioral problems that are more and more intertwined with today's engineering challenges.

Chapter 4

As mentioned before, in both economic and evolutionary theories of games two general classes of evolution can be identified: dynamics based on myopic optimization and dynamics based on imitations or replications. In network games, in which the players interact exclusively with a fixed set of neighbors, the dynamical features of these classes of dynamics vary significantly. In particular, myopic optimizations in social dilemmas tend to lead to Nash equilibrium payoffs that are well below the optimum (tragedy of the commons). Under imitation dynamics, the outcomes in terms of payoffs can be better, but convergence to an equilibrium is typically not guaranteed. In this chapter, we show that for a general class of public goods games, *rational imitation* dynamics converge to an imitation equilibrium in finite time independent of the spatial structure. For the more irrational 'imitate-the-best' dynamics, we identify network structures for which pure imitations lead to beneficial equilibrium profiles in which the players are satisfied with their decisions. Perhaps more importantly, we provide evidence that, in contrast to purely rational or purely imitation based decision rules, the combination of rationality and imitations in rational imitation dynamics guarantees both finite time convergence on arbitrarily connected graphs and high levels of cooperation in the imitation equilibrium profiles.

Chapter 5

In the existing models for finite non-cooperative network games, it is usually assumed that in each single round of play, regardless of the update rule driving the dynamics, each player selects the same action against all of its co-players. When a selfish player can distinguish the identities of his or her opponents, this assumption becomes highly restrictive. In this chapter, we will introduce the mechanism of *strategic differentiation* through which a subset of players in the network, called *differentiators*, can employ different actions against different opponents in their local game interactions. Within this new framework, we will study the existence of pure Nash equilibria and finite-time convergence of differentiated myopic best response dynamics by extending the theory of potential games to non-cooperative games with strategic differentiation. Finally, via simulation, we illustrate the effect of strategic differentiation on the equilibrium strategy profiles of a non-linear spatial public goods game. The simulation results show that depending on the position of differentiators in the network, the level of cooperation of the whole population at an equilibrium can be promoted or hindered. Moreover, if players imitate successful neighbors, a small number of differentiators placed on high degree nodes can result in large scale cooperation at very low benefit-to-cost ratios. Our findings indicate that strategic differentiation provides new ideas for solving the challenging free-rider problem on complex networks.

Part II: strategic play and control in repeated games

Part II is concerned with repeated games that are used to study the evolution of cooperation in social dilemmas through repeated interactions and the possibilities for future rewards and punishments. In particular, it is studied how individuals can exert control in n-player repeated games and in doing so can promote cooperation in repeated social dilemmas. New theory is developed for ZD strategies in a broad class of social dilemmas with discounting of future payoffs. Moreover, a novel discounting framework is proposed for repeated games that provides new insights into how individuals can exert control when the probability for future interactions is uncertain.

Chapter 6

The manipulative nature of ZD strategies attracted significant attention from researchers due to their close connection to controlling distributively the outcome of evolutionary games in large populations. In this chapter, we study the existence of ZD strategies in repeated n-player games with a finite but undetermined time horizon. Necessary and sufficient conditions are derived for a linear relation to be enforceable by a ZD strategist in n-player social dilemmas, in which the expected number of rounds is modeled by a fixed and common discount factor ($0 < \delta < 1$). For the first time in the studies of repeated games, ZD strategies are examined in the setting of finitely repeated n-player, two-action games. The results show that depending on the group size and the ZD-strategist's initial probability to cooperate, for *finitely* repeated n-player social dilemmas, it is possible for extortionate, generous and equalizer ZD-strategies to exist.

Chapter 7

In this chapter, we build upon the existence results in chapter 6 by developing a new theory that allows us to express threshold discount factors that determine how efficiently a strategic player can enforce a desired linear payoff relation. The efficiency is determined by a threshold discount factor that relies on the slope and baseline payoff of the desired linear relation and the variation in the "one-shot" payoffs of the n-player game. These general results apply to multiplayer and two-player repeated games and can be applied to a variety of complex social dilemma settings including the famous prisoner's dilemma, the public goods game, the volunteer's dilemma, the n-player snowdrift game and much more. The theory developed in this chapter can, for instance, be used to determine one's possibilities for exerting control given a constraint on the expected number of interactions or the general efficiency of generosity and extortion in n-player social dilemmas. To show the utility of these general results, we apply them to a variety of social dilemmas and show under which conditions mutual cooperation can be enforced by a single player in the group.

Chapter 8

In this chapter, we investigate the evolutionary stability of ZD strategies in a finite population. Necessary and sufficient conditions are provided for a resident ZD strategy to satisfy the equilibrium condition of evolutionarily stable strategies when they are invaded by a single ZD strategy. The derived conditions show that, for generous strategies that facilitate mutual cooperation to satisfy the stability condition with respect to one mutant strategy, the resident ZD strategists cannot be *too* generous. We provide an analytical expression for what exactly *too* generous is, and show that this depends on the one-shot payoff, the population size and the contest size of the n-player evolutionary game. Because in each contest, no other strategy can do better than an extortionate strategy, the evolutionary equilibrium conditions carry over to arbitrary mutant strategies in a finite population. Finally, a convenient method is proposed to check the evolutionary stability of resident ZD strategies with respect to any number of identical mutants.

Chapter 9

Evolutionary theories suggest that repeated interactions are necessary for direct reciprocity to be effective in promoting cooperative behavior in social dilemmas, and the discovery of zero-determinant strategies suggests that witty individuals can influence -for better or worse- the outcome of such repeated interactions. But what happens if the probability of repeating the mutual interactions is *uncertain*, and to what degree is it possible for a player to deal with this uncertainty in their efforts to influence the behavior of others? By incorporating the additional psychological complexity of an uncertain belief about the continuation probability into the framework of repeated games, in this chapter, we develop a general theory that can describe to what degree strategic players can influence the outcomes of multiplayer social dilemmas with uncertain future interactions. Our results suggest that this uncertainty can drastically alter one's opportunities to exert control and that some existing theories only hold in a more deterministic world. In particular, uncertainty may deny one's ability to ensure others do well, but the system remains vulnerable to extortion.

1.3 List of Publications

Journal articles

[1] Ye, M., Qin, Y., Govaert, A., Anderson, B. D., & Cao, M. (2019). An influence network model to study discrepancies in expressed and private opinions. Automatica, 107, 371-381.

[2] Govaert, A., & Cao, M. (2019). Zero-Determinant strategies in finitely repeated n-player games. Submitted. (Chapter 6 and 7)

[3] Govaert, A., & Cao, M. (2019). Uncertain discounting of future outcomes denies generosity in social dilemmas. Submitted. (Chapter 9)

[4] Govaert, A., Ramazi, P., & Cao, M. (2019). Imitation, rationality and cooperation in spatial public goods games. Submitted. (Chapter 4)

[5] Govaert, A., Cenedese, C., Grammatico, S., & Cao, M. (2019). Rationality and social influence in network games. In preparation. (Chapter 3)

Conference papers

[1] Govaert, A., Ramazi, P., & Cao, M. (2017, December). Convergence of imitation dynamics for public goods games on networks. In 2017 IEEE 56th Annual Conference on Decision and Control (CDC) (pp. 4982-4987). IEEE.

[2] Govaert, A., Qin, Y., & Cao, M. (2018, June). Necessary and sufficient conditions for the existence of cycles in evolutionary dynamics of two-strategy games on networks. In 2018 European Control Conference (ECC) (pp. 2182-2187). IEEE.

[3] Govaert, A., & Cao, M. (2019, June). Strategic Differentiation in Non-Cooperative Games on Networks. In 2019 18th European Control Conference (ECC) (pp. 3532-3537). IEEE. (Chapter 5)

[4] Govaert, A., & Cao, M. (2019, May). Zero-Determinant strategies in finitely repeated n-player games. In 2019 15th IFAC Symposium on Large Scale Complex Systems (LSS) (pp. 150-155). IFAC

[5] Govaert, A., Cenedese, C., Grammatico, S., & Cao, M. (2019) Relative Best Response Dynamics in finite and convex network games. In 2019 IEEE 58th Annual Conference on Decision and Control (CDC) (accepted). IEEE.

1.4 Notations

The set of real, positive, and non-negative numbers are denoted by $\mathbb{R}$, $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, respectively. The set of natural numbers is denoted by $\mathbb{N}$ and the set of integers is indicated by $\mathbb{Z}$. The cardinality of a set $\mathcal{A}$ is denoted by $|\mathcal{A}|$. For some vector $v \in \mathbb{R}^n$ we denote its i^{th} element by v_i. To emphasize a vector $v \in \mathbb{R}^n$ is obtained by stacking its elements v_i we write $v = (v_i) \in \mathbb{R}^n$. For a pair of vectors $w, u \in \mathbb{R}^n$, $w \cdot v = \sum_{i=1}^n w_i v_i$ is the dot product. Given a non-empty finite set $\mathcal{B}$ with cardinality m, the single valued function $\max^k(\mathcal{B})$, where $k \leq m$, evaluates the k^{th} highest value in the set $\mathcal{B}$. The power set of a non-empty set $\mathcal{B}$ is denoted by $2^{\mathcal{B}}$. We denote the n-ary Cartesian product over the sets $\mathcal{B}_1, \mathcal{B}_2, \ldots \mathcal{B}_n$ by $\prod_{i=1}^n \mathcal{B}_i$.

CHAPTER 2

Preliminaries

2.1 Network Games

Non-cooperative network games have three main ingredients: the network structure, the action space, and the combined payoff function. The action space is defined for both finite games, and convex games, in which the action sets are finite discrete sets and infinite compact and convex sets, respectively.

2.1.1 Network structure, action space and payoff functions

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph whose node set $\mathcal{V} = \{1, \ldots, N\}$ represents players. The edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, represents the player interaction topology. Let $\mathcal{A}_i$ denote the set of actions for player $i \in \mathcal{V}$ and let $\sigma_i \in \mathcal{A}_i$ denote the action of player i. The action space of the game is defined as the Cartesian product of the action sets of the players, i.e., $\boldsymbol{\mathcal{A}} = \prod_{i \in \mathcal{V}} \mathcal{A}_i$. An *action profile* of the game is an element of this set $\boldsymbol{\sigma} := (\sigma_i)_{i \in \mathcal{V}} \in \boldsymbol{\mathcal{A}}$, representing the actions chosen by all players in the network. To emphasize the i^{th} element of $\boldsymbol{\sigma} \in \mathbb{R}^N$, we write $\boldsymbol{\sigma} = (\sigma_i, \boldsymbol{\sigma_{-i}})$ where $\boldsymbol{\sigma_{-i}} = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_N)$. Let $\pi_i : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ indicate the payoff function of player i. The *combined* payoff function $\pi : \boldsymbol{\mathcal{A}} \to \mathbb{R}^N$ maps each action profile $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}$ to a payoff vector $\pi(\boldsymbol{\sigma}) = (\pi_i(\boldsymbol{\sigma}))_{i \in \mathcal{V}}$ whose elements correspond to the payoffs that the players receive for a single round of interaction. In network games, the spatial structure is incorporated into the payoff function π. Thus, the network structure determined by the graph $\mathcal{G}$, the action space $\boldsymbol{\mathcal{A}}$, and combined payoff function π defines the network game as the triplet $\Gamma = (\mathcal{G}, \boldsymbol{\mathcal{A}}, \pi)$.

2.1.2 Finite and convex games

We say Γ is a *finite* game if the action set of each player is a finite discrete set such that $\mathcal{A}_i \subset \mathbb{Z}$ and $\boldsymbol{\mathcal{A}} \subset \mathbb{Z}^N$. A finite game is denoted by Γ_f. We say Γ is a *convex* game if the action set of each player is a non-empty, convex subset of $\mathbb{R}^m$, i.e., $\mathcal{A} \subset \mathbb{R}^m$ and $\boldsymbol{\mathcal{A}} \subset \mathbb{R}^{Nm}$. A convex game is denoted by Γ_c. The convexity assumption over the action set for convex games is common in the literature of monotone games [47, 65].

2.2 Potential games

In Part I of the thesis, the theory of potential games is used. In [42], Monderer and Shapely identify several classes of games for which there exists a function that increases or decreases monotonically along the trajectory of rational decisions in a game. The most restrictive class is known as exact potential games that are defined as follows.

2.2.1 Finite games

Definition 1 (Exact potential game)**.** *Given a finite game Γ_f, if there exists a function $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ such that for every $i \in \mathcal{V}$, for every $\sigma_i, \sigma_i' \in \mathcal{A}_i$ and every $\boldsymbol{\sigma}_{-i} \in \prod_{j \in \mathcal{V} \setminus \{i\}} \mathcal{A}_j$, the following implication holds:*

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) = P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i}) \tag{2.1}$$

then Γ_f is an exact potential game.

Several generalizations of exact potential games exist. The following definitions provide an overview of increasingly general classes of games.

Definition 2 (Weighted potential game)**.** *Given a finite game Γ_f, if there exists a function $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ such that for every $i \in \mathcal{V}$, for every $\sigma_i, \sigma_i' \in \mathcal{A}_i$ and every $\boldsymbol{\sigma}_{-i} \in \prod_{j \in \mathcal{V} \setminus \{i\}} \mathcal{A}_j$, the following implication holds:*

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) = \alpha_i \left[P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i}) \right], \tag{2.2}$$

then Γ_f is a weighted potential game.

Definition 3 (Ordinal potential game)**.** *Given a finite game* Γ_{f}*, if there exists a function* $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ *such that for every* $i \in \mathcal{V}$*, for every* $\sigma_i, \sigma_i' \in \mathcal{A}_i$ *and every* $\boldsymbol{\sigma}_{-i} \in \prod_{j \in \mathcal{V} \setminus \{i\}} \mathcal{A}_j$*, the following implication holds:*

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0 \Leftrightarrow P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0, \tag{2.3}$$

then Γ_{f} *is an ordinal potential game.*

Definition 4 (Generalized ordinal potential game)**.** *Given a finite game* Γ_{f}*, if there exists a function* $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ *such that for every* $i \in \mathcal{V}$*, for every* $\sigma_i, \sigma_i' \in \mathcal{A}_i$ *and every* $\boldsymbol{\sigma}_{-i} \in \prod_{j \in \mathcal{V} \setminus \{i\}} \mathcal{A}_j$*, the following implication holds:*

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0 \Rightarrow P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0, \tag{2.4}$$

then Γ_{f} *is a generalized ordinal potential game.*

Potential games (and their generalizations) with finite action sets have an important property called the Finite Improvement Property (FIP) that is formalized as follows.

Definition 5 (Finite Improvement Property [42, Sec. 2])**.** *Let* $\gamma = (\boldsymbol{\sigma}(0), \boldsymbol{\sigma}(1), \dots)$ *denote a action profile sequence for* Γ*. If for every* $t \geq 1$ *there exists a unique player, say* $i_t \in \mathcal{V}$ *such that*

$$\boldsymbol{\sigma}(t) = (\sigma_{i_t}(t), \boldsymbol{\sigma}_{-i_t}(t-1)) \text{ for some } \sigma_{i_t}(t) \neq \sigma_{i_t}(t-1),$$

then γ *is called a path in the action profile. If additionally it holds that for each consecutive action profile in a path* γ *the payoff of the unique deviator* i_t *is strictly increasing, that is*

$$\forall t \geq 1 \; : \; \pi_{i_t}(\boldsymbol{\sigma}(t)) > \pi_{i_t}(\boldsymbol{\sigma}(t-1)),$$

then γ *is called an improvement path.* Γ *has the Finite Improvement Property (FIP) if every improvement path is finite.*

Lemma 1 (Finite Improvement paths in potential games [42, Sec. 2])**.** Γ_{f} *has the FIP if and only if* Γ_{f} *has a generalized ordinal potential function.*

2.2.2 Infinite games

In finite potential games, the action space is finite and in turn, the potential function is bounded. Naturally, these properties do not hold when the number of actions is infinite. In the following, we shortly introduce concepts from the theory of infinite potential games, i.e. potential games with an infinite number of actions. We will focus on convex games, that unless the action sets are singleton sets, can also be characterized as infinite games. We begin with the infinite game counterpart of improvement paths, commonly referred to as approximate improvement paths.

Definition 6 (ϵ-improvement paths)**.** *Let γ denote a sequence in the action profile of Γ_c and let $\epsilon > 0$ be an arbitrarily small positive real. When for every $t \geq 1$ there exists a unique player, say $i_t \in \mathcal{V}$, such that*

$$\boldsymbol{\sigma}(t) = (\sigma_{i_t}(t), \boldsymbol{\sigma}_{-i_t}(t-1)) \text{ for some } \sigma_{i_t}(t) \neq \sigma_{i_t}(t-1),$$

then γ is called a path in the action profile. When additionally it holds that for each consecutive action profile in a path γ the payoff of the unique deviator i_t is strictly increasing, i.e., $\forall t \geq 1$

$$\text{if } i = i_t : \ \pi_i(\boldsymbol{\sigma}(t)) > \pi_i(\boldsymbol{\sigma}(t-1)) + \epsilon,$$

then γ is called an ϵ–improvement path with respect to Γ_c.

The FIP for infinite games is known as the Approximate Finite Improvement Property (AFIP).

Definition 7 (Approximate Finite Improvement Property, [42])**.** *Γ_c has the AFIP if for every $\epsilon > 0$, every ϵ–improvement path is finite.*

Finite approximate improvement paths are naturally connected to the concept of an approximate Nash Equilibrium (NE), that is defined as follows.

Definition 8 (ϵ-Nash equilibrium)**.** *The action profile $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}$ is an ϵ–NE for Γ_c, if for all $i \in \mathcal{V}$, $\sigma_i \in \boldsymbol{\sigma}$ is such that*

$$\pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) > \pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) + \epsilon, \quad \forall \sigma_i' \in \mathcal{A}_i,$$

for some $\epsilon > 0$.

To characterize the class of infinite games that have the AFIP, it is necessary to introduce the concept of a bounded game.

Definition 9 (Bounded Game)**.** *Γ_c is a bounded game if for all $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}$ there exists $M \in \mathbb{R}$ such that $\forall i \in \mathcal{V}$ it holds that $|\pi_i(\boldsymbol{\sigma})| \leq M$.*

Bounded games thus have bounded payoff functions on the action space of the infinite game. For weighted and exact potential games, this implies that also the potential function is bounded. This leads to the following Lemma.

Lemma 2 (Lemma 4.2, [42])**.** *Every bounded w–potential game has the AFIP.*

Part I

RATIONALITY AND SOCIAL INFLUENCE IN NETWORK GAMES

CHAPTER 3

Relative Best Response dynamics in network games

> When people are free to do as they please, they usually imitate each other.
>
> *Eric Hoffer*

Game-theoretic scenarios in which players interact exclusively with a fixed group of neighbors traces back to the early 1990's when economists and biologists started to explore the effect of simple spatial structures in (probabilistic) decision-making processes driven by rational best response processes and more biologically inspired imitation processes [66–68]. Later, simple spatial structures were extended to arbitrary structures defined by graphs [34, 37, 45].

The long-run collective behavior of non-cooperative network games have been extensively studied for best response dynamics in which the players, given the history of plays of their neighbors, select a strategy that maximizes their payoffs. These extended research efforts have resulted in the identification of several classes of games that converge to a pure Nash equilibrium under a variety of such best response processes [42–44, 69] and brought forth a number of algorithms that ensure convergence to an equilibrium [47–49]. Best response dynamics are "*innovative*" in the sense that, to optimize their payoffs, players are always able to select new actions that are not played in the current strategy profile. They are in line with classic economic theories that support the idea that absolute optimization (or rational behavior) is a natural result of evolutionary forces [70]. Recently, the systems and control community has been interested in the analysis of dynamical systems driven by *imitation* [71–73]. Such dynamics are "*non–innovative*": players can only select actions that already exist in the networked population. Therefore, non–innovative dynamics can lead to equilibrium concepts that differ from traditional Nash equilibria. In [74, 75], the authors studied an evolutionary process where the players, most of the time, choose a best response from the set of actions that exist in the entire population strategy profile. In [75], this evolutionary process was simply referred to as imitation. Perhaps a more suitable name was proposed in [74], where such a revision was called a *Relative Best Response* (RBR). RBR combines the non-innovative nature of pure imitation with the rationality of best response. Such dynamics match classic economic studies that support the idea that rather than absolute performance, it is *relative performance*, that proves to be decisive in the long run [76]. Experimental evidences of such behavior are documented in [77, 78]. Another motivation for studying such dynamics is that they can take into account the effect of *word-of-mouth communication* and social learning in decision making processes [79]. For example, when reconsidering alternative technologies, an individual may ask friends or family about their current choice and benefits. This local spread of information, in turn, is likely to affect her decision, and may very well lead to a complete disregard of technology that is not used by her peers. Indeed, the adoption of new technologies is affected by social influence [80–82]. Traditional best response dynamics do not capture such a process of information exchange and social learning, rather they reflect situations in which an individual adopts some technology solely based on his/her *own* expectation, regardless of how others have perceived it. In many real-world decision-making processes, it is likely that both types of learning processes occur [83], but from a theoretical point of view the effects of social learning is often overlooked.

In this chapter, a novel game dynamics for finite and convex games on networks are proposed that result from an intuitive combination of rational behavior and social learning. We start on the basis of a spatial version of *Relative Best Response* (RBR) dynamics under which the players choose a best response from (a convex combination of) the current set of actions in their neighborhood. In this case, the players interact and relate their success exclusively with a *fixed* group of neighbors Even though this process contains an element of social learning, namely that the players prefer to *conform* themselves to observed actions, it does not take into account the relative performance of these actions. To this end, we generalize RBR dynamics to the h-RBR, where players relate their success to the subset of neighbors that obtain *at least* the h-highest payoffs within their neighborhood. This process relies on local information exchange of both decisions and benefits, that are fundamental to social learning by imitation. Even though under h-RBR dynamics the feasible action sets of the players are state-dependent and the overall problem is not-jointly convex, we show that for a general class of games such dynamics converge to an (approximate) generalized Nash equilibrium in finite-time, and relate the results to classes of games for which best response dynamics converge to a Nash equilibrium.

Throughout this chapter, it is assumed the action sets of the players are the same. This naturally allows players to imitate each other, and is in fact common in imitation dynamics [68, 71, 72].

Assumption 1 (Identical action sets)**.** *All players have the same action set, i.e., $\mathcal{A}_i = \mathcal{A}$ for all $i \in \mathcal{V}$.*

One can argue that there exist decision-making processes in which the action sets of the players are inherently different. For example, when individual A aims to go to destination Z, and individual B aims to go a different destination Y. In such cases, it does not make sense that individual A and B learn from each other how to arrive at their destinations. However, in many real-world decision-making processes, it is observed that, through *social learning*, new behaviors are acquired by imitating *others* [84]. For example, a company can decide to enter a market because they observed another company having success there. Assumption 1, in this sense, is a technical one that ensures all decision-makers can imitate each other's actions and affect one another in this process. We note that it is possible to relax this assumption, for instance by adding constraints on one's ability to imitate another player's action. However, the additional technicalities would defy the main purpose of this chapter, namely to illustrate *clearly* how rationality and social influence can be combined and studied in a *common* framework.

3.1 h-relative best response dynamics

Before defining the h–RBR dynamics, for the purpose of comparison, we give the definition of a best response.

Definition 10 (Best response)**.** *For player $i \in \mathcal{V}$, a best response is any action in the set*

$$\mathcal{B}_i(\boldsymbol{\sigma}_{-i}) := \operatorname*{argmax}_{y \in \mathcal{A}} \pi_i(y, \boldsymbol{\sigma}_{-i}).$$

The defining distinction of a *relative* best response is that, instead of optimizing over a fixed action set $\mathcal{A}$, player $i \in \mathcal{V}$ optimizes its payoffs over some *feasible* subset of $\mathcal{A}$ that depends on the actions of the neighbors of i and σ_i itself. For a game Γ

and an action profile $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}$, we denote the feasible action set for player $i \in \mathcal{V}$ by $\mathcal{F}_i(\boldsymbol{\sigma}) \subseteq \mathcal{A}$. For a finite game Γ_f, the feasible action set of player $i \in \mathcal{V}$ is simply determined as the local set of actions, i.e.,

$$\mathcal{F}_i^\mathrm{f}(\boldsymbol{\sigma}) := \{\sigma_j \in \boldsymbol{\sigma} \mid j \in \mathcal{N}_i\} \cup \{\sigma_i\} \subseteq \mathcal{A}. \tag{3.1}$$

Instead, for a convex game Γ_c, the action sets are convex and compact subsets of $\mathbb{R}^n$, hence the feasible action set for player $i \in \mathcal{V}$ is determined as

$$\mathcal{F}_i^\mathrm{c}(\boldsymbol{\sigma}) = \mathrm{conv}(\mathcal{F}_i^\mathrm{f}) \subseteq \mathcal{A}\,. \tag{3.2}$$

We are now ready to formalize the idea of RBR.

Definition 11 (Relative Best Response)**.** *Given a game* Γ*, a relative best response of player* $i \in \mathcal{V}$ *is any action in the set*

$$\mathcal{B}_i^\mathrm{r}(\boldsymbol{\sigma}_{-i}) := \underset{y \in \mathcal{F}_i(\boldsymbol{\sigma})}{\mathrm{argmax}}\ \pi_i(y, \boldsymbol{\sigma}_{-i}),$$

where the feasible action set $\mathcal{F}_i(\boldsymbol{\sigma}^*, h_i)$ *of a finite game and convex game are given by* Eq. (3.1) *and* Eq. (3.2)*, respectively.*

Imitations are often linked to *social learning*, in which new behaviors are acquired by observing and imitating others [84]. In the context of a game, to choose which neighbor's action to imitate, the players must thus have information about the actions and the current payoffs of their neighbors. It is this local exchange of information, that is absent in best response dynamics, that can lead to surprising "non-rational" behavior. As in BR, an RBR is based only on the local actions, and thus does not take into account the *payoffs* of others. An interesting and natural generalization of RBRs is a decision process in which the feasible action set of player $i \in \mathcal{V}$ depends on a subset of the neighbors that receive the h_i highest payoffs. Roughly speaking, only those actions that are taken by successful neighbors are considered in the action update. In this case, the *relative success* of the neighbors of i will have an influence on the future action of player i, and $h_i \in \mathbb{N}$ is a measure for how restricting this relative success is for player $i's$ feasible action set.

We dedicate the remainder of this section to formalize this novel revision process and illustrate its concepts with examples of interesting applications that are likely to be affected by relative performance considerations and social influence. Before defining the revision process formally, it is necessary to introduce some additional auxiliary sets. For some action profile $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}$, let us define the set of *distinct* payoffs

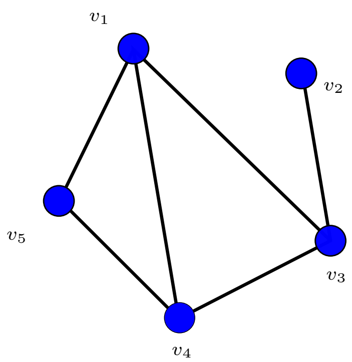


Figure 3.1

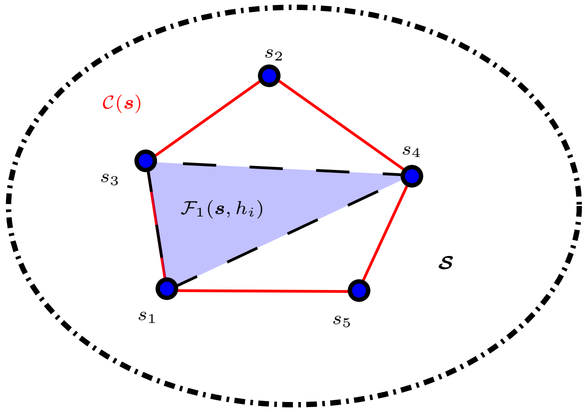


Figure 3.2

Figure 3.3: Suppose the network is as in (a) such that $n = 5$. The set of actions of the neighbors of 1 is $\mathcal{M}_1(\boldsymbol{\sigma}_{-i}) = \{s_3, s_4, s_5\}$. Moreover, suppose that $\pi_4(\boldsymbol{\sigma}) > \pi_3(\boldsymbol{\sigma}) > \pi_2(\boldsymbol{\sigma}) > \pi_5(\boldsymbol{\sigma})$ and $h_i = 2$. Then, $\mathcal{M}_1(\boldsymbol{\sigma}_{-i}, 2) = \{s_4, s_5\}$, $\mathcal{F}_1^{\mathrm{c}}(\boldsymbol{\sigma}, 2) = \{s_4, s_5, s_1\}$ and the shaded area with the dashed border in (b) illustrates $\mathcal{F}_1^{\mathrm{c}}(\boldsymbol{\sigma}, 2)$. Moreover, $\mathcal{C}(\boldsymbol{\sigma})$ is the convex hull of the entire action profile as is indicated by the region with the red border.

obtained by the neighbors of i as $\mathcal{R}_i(\boldsymbol{\sigma}) := \{\pi_j(\boldsymbol{\sigma}) \mid j \in \mathcal{N}_i\}$, and define the set of neighbors that receive at least the h_i highest payoff as

$$\mathcal{H}_i(\boldsymbol{\sigma}_{-i}, h_i) := \{j \in \mathcal{N}_i \mid \pi_j(\boldsymbol{\sigma}) \geq \max{}^{h_i}(\mathcal{R}_i(\boldsymbol{\sigma}))\},$$

Note that, it always holds that $|\mathcal{N}_i| \geq |\mathcal{H}_i(\boldsymbol{\sigma}_{-i}, h_i)| \geq h_i$. Then, the set of actions of these successful players is given by

$$\mathcal{M}_i(\boldsymbol{\sigma}_{-i}, h_i) := \{\sigma_j \in \boldsymbol{\sigma} \mid j \in \mathcal{H}_i(\boldsymbol{\sigma}_{-i}, h_i)\}. \tag{3.3}$$

In this case, for a finite game Γ_{f}, the feasible set of actions is determined by

$$\forall i \in \mathcal{V} : \mathcal{F}_i^{\mathrm{f}}(\boldsymbol{\sigma}, h_i) := \{\mathcal{M}_i(\boldsymbol{\sigma}_{-i}, h_i)\} \cup \{\sigma_i\} \subseteq \mathcal{A}, \tag{3.4}$$

while for a convex game Γ_{c}, it is

$$\forall i \in \mathcal{V} : \mathcal{F}_i^{\mathrm{c}}(\boldsymbol{\sigma}, h_i) := \mathrm{conv}\{\mathcal{F}_i^{\mathrm{f}}(\boldsymbol{\sigma}, h_i)\}. \tag{3.5}$$

Let $\boldsymbol{h} = (h_i)_{i \in \mathcal{V}} \in \mathbb{N}^N$. An h–RBR can now be formalized as follows.

Definition 12 (h-Relative Best Response)**.** *Given a game* Γ*, a* h*–relative best response of player* $i \in \mathcal{V}$ *is any action in the set*

$$\mathcal{B}_i^{\mathrm{r}}(\boldsymbol{\sigma}_{-i}, h_i) := \underset{y \in \mathcal{F}_i(\boldsymbol{\sigma}, h_i)}{\operatorname{argmax}} \ \pi_i(y, \boldsymbol{\sigma}_{-i}).$$

It is worth mentioning that, if $h_i = |\mathcal{N}_i|$ for every $i \in \mathcal{V}$, then Definition 12 recovers the definition of a relative best response. In contrast, for finite games, when $h_i = 1$, player i can only choose between his/her own action and the actions of the most successful neighbors. Therefore, if for all $i \in \mathcal{V}$, $h_i = 1$ the feasible actions of the h-RBR dynamics for finite games are exactly the feasible set of actions in an unconditional imitation process. We will explore this link to imitation dynamics in Chapter 4.

3.1.1 Examples of h-RBR applications

Example 1 (Adoption of competing products)**.** *Let us elaborate on the role of* h_i *in the context of the technology adoption example. Suppose an individual i is considering to adopt a new product and can choose between models* X*,* Y *and* Z*, to replace her current product* C*. In this case,* $\mathcal{A} = \{X, Y, Z, C\}$*. She values her current product with a* 3 *on a scale from zero to five. To make a decision about which product to adopt, she gathers information from three peers, labeled as* $\mathcal{N}_i = \{a, b, c\}$*, who she believes value the product in a similar manner as herself. Suppose model* X *is used by peer* a *and values the model with a full score of* 5 *out of five. In this case,* $\sigma_a = X$ *and* $\pi_a = 5$*. Model* Y *is used by peer* b *who values it with* 2 *(i.e.,* $\sigma_b = Y$*,* $\pi_b = 2$*)and model* Z *is used by peer* c *who values it with* 4 *(i.e.,* $s_c = Z, \pi_c = 4$*). In our notation, the distinct payoffs obtained by her neighbors is* $\mathcal{R}_i(\boldsymbol{\sigma}) = \{5, 2, 4\}$*. If* $h_i = 1$ *then, the individual would only consider to keep her current phone or buy model* X *because she believes model* Z *is worse than* X *and model* Y *is not worth the upgrade from her current product. In our notation, the set of action chosen by her most successful peer is* $\mathcal{M}_i^{\mathrm{f}}(\boldsymbol{\sigma}, 1) = \{\sigma_a\} = \{X\}$*, and the set of feasible actions is* $\mathcal{F}_i^{\mathrm{f}}(\boldsymbol{\sigma}, 1) = \{C, X\}$*. However, if* $h_i = 2$*, she would also consider buying model* Z *that due to individual differences in the perception of values may be a better choice for her. In this case,* $\mathcal{M}_i^{\mathrm{f}}(\boldsymbol{\sigma}, 1) = \{\sigma_a, \sigma_c\} = \{X, Z\}$ *and* $\mathcal{F}_i^{\mathrm{f}}(\boldsymbol{\sigma}, 2) = \{C, Z, X\}$*. In this example,* h_i *influences how the information from peers reflect her own valuation of a product. That is, if* $h_i = 3$ *then she would take into account every product because she could be uncertain if the low score of model* Y *reflects her own preferences accurately.*

Example 2 (Adoption of renewable energy)**.** *Suppose a fossil-fueled household is allowed to determine the fraction of energy obtained from renewable sources. In* *this case, $\mathcal{A} = [0, 1]$. To obtain an idea of how costly and sustainable the usage of renewable energy is compared to fossil fuel, they gather information from neighboring households with similar energy demands. If none of the neighbors are using renewable energy sources, due to inertia in the decision making the household may be inclined to refrain from using renewable energy simply because they lack information to make a reasonable decision about it and there are no forces of conforming to a green source of energy. In our notation, this would lead to $\mathcal{F}_i^c(\boldsymbol{\sigma}, h_i) = \{0\}$. However, if neighboring households are already using renewable energy and have informed the household that they are satisfied with the supply and costs, an appealing option is to choose some fraction of sustainable energy based on the fraction chosen by the neighbors. This decision is plausible because of two reasons: first, the information gathered from similar households suggests that renewable energy is a good alternative source of energy and second, conformity forces that result in peer pressures may lead the household to decide to try renewable energy sources [85].*

In some contexts it makes sense to apply a transformation to the action profile and payoffs before applying an h-relative best response.

Example 3 (Opinion dynamics)**.** *Take for example an opinion dynamics model in which $s_i \in \mathbb{R}$ represents an opinion that takes values on the unit interval. In these settings, it is well-established that social learning plays a crucial role in the evolution of opinions as individuals tend to adjust their opinion to a local weighted average [86, 87]. Such a process can be represented by a network game with best responses. Now, let us define a simple auxiliary "payoff function" that player i observes in neighbor j as*

$$\epsilon_{ij}(\boldsymbol{\sigma}) := 1 - |\sigma_i - \sigma_j|,$$

and let $\epsilon_i(\boldsymbol{\sigma}) \in \mathbb{R}^{|\mathcal{N}_i|+1}$ be the vector of these opinion errors. Now suppose the player applies the principle of selecting the h_i highest valued neighbors. Then the opinion dynamics would result in a bounded-confidence model in which the player only takes into account those neighbors that have an opinion similar to the player's own opinion.

Now that we have defined an h–RBR, let us introduce the *asynchronous*, or sequential, game dynamics that are associated with the h-RBR via an activation sequence: at each time step $t \in \mathbb{N}$ for which $\boldsymbol{\sigma}(t+1) \neq \boldsymbol{\sigma}(t)$, there exists a unique player $i_t \in \mathcal{V}$ such that the collective dynamics satisfy

$$\begin{aligned} \text{if } i = i_t: \quad \boldsymbol{\sigma}(t+1) &= (\sigma_i(t+1)), \boldsymbol{\sigma}_{-i}(t+1)) \\ &\in (\mathcal{B}_i^{\mathrm{r}}(\boldsymbol{\sigma}_{-i}(t), h_i), \boldsymbol{\sigma}_{-i}(t)). \end{aligned} \tag{3.6}$$

For the asynchronous dynamics in Eq. (3.6) we assume that the activation sequence ensures that at any time step, each player is guaranteed to be active at some finite future time.

Assumption 2 (persistent activation sequence)**.** *Every sequence of activated players* $(i_t)_{t\in\mathbb{N}}$ *driving the asynchronous dynamics* Eq. (4.8) *is persistent, i.e., if for every player* $j \in \mathcal{V}$ *and every time* $t \in \mathbb{N}$*, there exists some finite-time* $\bar{t} > t$ *at which player* j *is active again, i.e.,* $i_{\bar{t}} = j$*.*

3.1.2 Convergence problem statement

We are interested in characterizing the conditions under which the dynamics in Eq. (4.8) converge to an equilibrium action profile. In this case, all players in the network reach a decision with which they are satisfied. For the h–BRB dynamics, the local feasible action set for each player is constrained by the actions of the other players and hence the equilibrium action profiles of these dynamics correspond to a *Generalized Nash Equilibria* (GNE) [88].

Definition 13 (Generalized Nash Equilibrium)**.** *The action profile* $\boldsymbol{\sigma}^* \in \boldsymbol{\mathcal{A}}$ *is a GNE for* Γ*, if for all* $i \in \mathcal{V}$

$$\sigma_i^* \in \mathcal{B}_i^{\mathrm{r}}(\boldsymbol{\sigma}_{-i}^*, h_i), \tag{3.7}$$

where the feasible action set $\mathcal{F}_i(\boldsymbol{\sigma}^*, h_i)$ *of a finite game and convex game are given by* Eq. (3.4) *and* Eq. (3.5)*, respectively.*

It is worth mentioniong that, in the convex game case, our GNE problem is not jointly convex [89]. In Sections 3.2 and 3.3, we will study the convergence properties of Eq. (4.8) for finite and convex games under the following assumption which ensures that players only switch to another action if they have an incentive to deviate from their current action.

Assumption 3 (Incentive to deviate)**.** *For* Γ*,* $\sigma_i(t) \neq \sigma_i(t+1)$ *only if there exists* $y \in \mathcal{F}_i(\boldsymbol{\sigma}, h_i)$ *such that*

$$\pi_i(y, \boldsymbol{\sigma}_{-i}(t)) - \pi_i(\sigma_i(t), \boldsymbol{\sigma}_{-i}(t)) > 0.$$

3.2 Convergence in finite games

In this section, we study the convergence of the asynchronous h–RBR dynamics in Eq. (4.8) when all players choose h-relative best responses and they can have a finite set of actions that they can choose from. First, we define two sets that will prove useful in the analysis of the h-BRB dynamics in finite and convex games. For an initial action profile $\boldsymbol{\sigma}(0)$, let us denote the set that contains all actions that are employed by at least one player in the initial action profile by $\mathcal{A}_0 := \cup_{i\in\mathcal{V}}\{\sigma_i(0)\}$, and let $\boldsymbol{\mathcal{A}}_0 := \mathcal{A}_0^N$. The set $\mathcal{A}_0$ is called the *support* of $\boldsymbol{\sigma}(0)$ in [74]. The key property of $\boldsymbol{\mathcal{A}}_0$ is that it is positively invariant with respect to the h-RBR dynamics Eq. (4.8), due to their non-innovative nature. To study the convergence properties of finite games under the asynchronous h-RBR dynamics we use the theory of potential games [42]. Consider the following definition of a potential like function.

Definition 14 ($\mathcal{A}_0$–potential function)**.** *A function $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ is a $\mathcal{A}_0$-potential function for Γ_f and some $\boldsymbol{\sigma}(0) \in \boldsymbol{\mathcal{A}}$, if for every $i \in \mathcal{V}$, $\sigma_i, \sigma_i' \in \mathcal{A}_0$ and $\boldsymbol{\sigma}_{-i} \in \mathcal{A}_0^{N-1}$, it holds that if*

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0 \Rightarrow P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0. \tag{3.8}$$

If such a function exists, then we call Γ_f a relative potential game with respect to $\mathcal{A}_0$.

Remark 1. *When the initial action profile $\boldsymbol{\sigma}(0) \in \boldsymbol{\mathcal{A}}$ is such that $\boldsymbol{\mathcal{A}}_0 = \boldsymbol{\mathcal{A}}$, then Definition 14 is equivalent to the definition of a generalized ordinal potential function and a generalized ordinal potential game [42, Sec. 2]. In its classic definition, the implication in* Eq. (3.8) *needs to be satisfied on the entire action space $\boldsymbol{\mathcal{A}}$ to ensure convergence of the innovative best response dynamics to a pure Nash equilibrium.*

We are now ready to present the main result for finite games that relies on the existence of a $\mathcal{A}_0$-potential function.

***Theorem* 1.** *Suppose Assumption 3 is satisfied and that Γ_f is a relative potential game with respect to $\mathcal{A}_0$. Then, for all $\boldsymbol{\sigma}(0) \in \boldsymbol{\mathcal{A}}_0$ the asynchronous h–RBR dynamics in* Eq. (4.8) *converge to a GNE in finite-time.*

Proof. Suppose $\boldsymbol{\sigma}(0) \in \boldsymbol{\mathcal{A}}_0$. Because the h–RBR dynamics are non–innovative, it follows that $\boldsymbol{\sigma}(t) \in \boldsymbol{\mathcal{A}}_0$, for all $t \geq 0$. By Assumption, Γ is a relative potential game with respect to $\mathcal{A}_0$, hence there exists a function $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ such that for every $i \in \mathcal{V}$, for every $\sigma_i, \sigma_i' \in \mathcal{A}_0 \cap \mathcal{A}_i$ and every $\boldsymbol{\sigma}_{-i} \in \prod_{j\in\mathcal{V}\setminus\{i\}} \mathcal{A}_0$, the following implication holds:

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0 \Rightarrow P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i}) > 0. \tag{3.9}$$

By Definition 12, Eq. (3.4) and the asynchronous dynamics in Eq. (4.8) it follows that after a player switches, their payoff is at least as high as it was before. That is, for all $t \geq 1$:

$$\exists\ i_t \in \mathcal{V}:\ \pi_{i_t}(t) \geq \pi_{i_t}(t-1). \tag{3.10}$$

By Assumption 3, if a player switches, then inequality Eq. (3.10) holds *strictly* and hence the trajectory of relative best response dynamics generates an improvement path γ (see Definition 5). Since for all $t \geq 0$, we have $\mathcal{F}_i^c(\boldsymbol{\sigma}(t), h_i) \subseteq \mathcal{A}_0$. From the implication Eq. (3.9), it follows that the $\mathcal{A}_0$-potential function P is strictly increasing along γ. Since the action space is finite, P is a bounded function. This implies that the h-relative best response dynamics converge to a GNE in finite-time. □

It may happen that there exist $\mathcal{A}_0$-potential functions only for a subset of initial action profiles. To guarantee finite-time convergence for all initial condition, it is required there exists a generalized potential function, *not necessarily the same*, for every initial action profile. This is formalized in the following definition.

Definition 15 (Generalized relative potential game)**.** *If for* Γ_{f} *there exist generalized* $\mathcal{A}_0$*–potential functions for every* $\sigma(0) \in \boldsymbol{\mathcal{A}}$*, then* Γ_{f} *is called a generalized relative potential game.*

An example of a generalized relative potential game can be found in Example 4. An immediate consequence of Theorem 1 is stated in the following corollary.

Corollary 1. *For any finite generalized relative potential game, the asynchronous* h*–RBR dynamics converge globally to a GNE in finite-time.*

3.2.1 Relation to generalized ordinal potential games

From Definition 14, it can be easily seen that every generalized ordinal potential game is a generalized *relative* potential game. By means of the following counter-example we show that the converse is not always true, that is, not every generalized relative potential game is a generalized ordinal potential game.

Example 4. *Consider the symmetric Rock-Scissors-Paper (RSP) game with payoff matrix*

$$\boldsymbol{M} = \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}, \quad b > a \geq c. \tag{3.11}$$

Because each improvement path in the RSP game converges to the improvement cycle: $(R,S) \to (R,P) \to (P,S) \to (S,R) \to (P,R) \to (P,S) \to (R,S)$*, the RSP game is not a generalized ordinal potential game. However, for all initial action profile* $\boldsymbol{\sigma}(0) \in \boldsymbol{\mathcal{A}} := \{R,S,P\}^2$ *there exists a generalized* $\mathcal{A}_0$*–potential and thus the RSP game is a generalized relative potential game.*

Example 4 highlights that, especially for finite games in which the number of actions is larger than the number of players (i.e. $|\mathcal{A}| > N$), for the convergence of h-RBR dynamics Eq. (4.8) it is easier to rely on the existence of generalized $\mathcal{A}_0$–potential functions rather than generalized ordinal potential functions. In fact, it can be easily proven that every symmetric two-player $|\mathcal{A}| \times |\mathcal{A}|$ game converges to a GNE under Eq. (4.8) by using the fact that there always exist an exact potential function for 2×2 games. The RSP game also shows the relation to generalized ordinal potential games.

Proposition 1. *Let G,R denote the class of generalized ordinal potential games and generalized relative potential games, respectively. Then, $G \subset R$.*

Proof. The inclusion $G \subseteq R$ follows from Definitions 14 and 15. Strictness follows from Example 4. □

Corollary 2. *For any finite generalized ordinal potential game, the asynchronous h-RBR dynamics converge globally to a GNE in finite-time.*

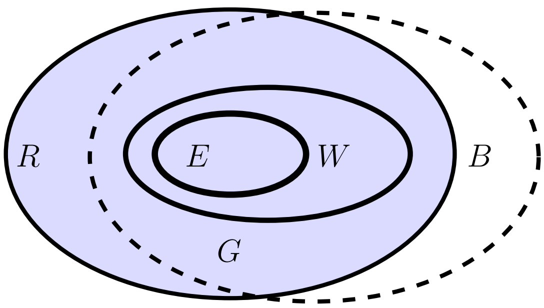


Figure 3.4: Let E, W, G, B, R represent the class of exact, weighted, generalized ordinal, best response, and generalized relative potential games, respectively. For finite games, the classic asynchronous best response dynamics are known to converge to a Nash equilibrium for E, W, G, B (Set indicated by dashed border) Corollary 2 and Proposition 1 shows that the asynchronous h–RBR dynamics will converge to a GNE for every game in the class $R \supset G \supset W \supset E$.

3.3 Convergence in convex games

In this section the concepts of *bounded* games and ϵ-improvement paths that are defined in the preliminaries chapter 2. For convex games, we are interested in the *finite* time convergence to an *approximate* GNE that is defined as follows.

Consider the following class of games inspired by weighted potential games [42].

Definition 16 (weighted $\mathcal{A}_0$-potential function)**.** *A function $P : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ is a weighted $\mathcal{A}_0$-potential function for Γ_{c} and some $\boldsymbol{\sigma}(0) \in \boldsymbol{\mathcal{A}}$, if for every $i \in \mathcal{V}$, $\sigma_i, \sigma_i' \in \mathcal{A}_0$ and $\boldsymbol{\sigma}_{-i} \in \mathcal{A}_0^{N-1}$, the following implication holds*

$$\pi_i(\sigma_i', \boldsymbol{\sigma}_{-i}) - \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) = w_i \left[P(\sigma_i', \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}_{-i})\right],$$

for some $w_i \in \mathbb{R}_+$. If such a function exists, then we call Γ_{c} a weighted relative potential game with respect to $\mathcal{A}_0$. Moreover, if $w_i = 1$ for all $i \in \mathcal{A}$, then Γ_{c} is called an exact relative potential game with respect to $\mathcal{A}_0$.

The following Lemma relates weighted $\mathcal{A}_0$-potential functions to exact $\mathcal{A}_0$-potential functions.

Lemma 3 (Equivalence weighted and exact $\mathcal{A}_0$–potential function)**.** *Γ_{c} is a weighted relative potential with respect to $\mathcal{A}_0$ if and only if Γ_{c}',with payoff functions $\frac{1}{w_i}\pi_i$, is an exact relative potential with respect to $\mathcal{A}_0$*

Proof. From the definition of a weighted potential game Γ we have $\pi_i(\sigma_i, \sigma_{-i}) - \pi_i(\sigma_i', \sigma_{-i}) = w_i \left(P(\sigma_i, \boldsymbol{\sigma}_{-i}) - P(\sigma_i', \boldsymbol{\sigma}_{-i})\right)$. On the other hand, from the definition of a potential game Γ' we have $\frac{1}{w_i}\pi_i(\sigma_i, \sigma_{-i}) - \frac{1}{w_i}\pi_i(\sigma_i', \sigma_{-i}) = P(\sigma_i, \boldsymbol{\sigma}_{-i}) - P(\sigma_i, \boldsymbol{\sigma}'_{-i})$. Clearly these are equivalent. □

The following result provides sufficient conditions for the convergence of h-relative best response dynamics in convex games.

Theorem 2. *Suppose Γ is a bounded game and a weighted relative potential game with respect to $\mathcal{A}_0$. Then for every $\epsilon > 0$, and initial action profile $\sigma(0) \in \boldsymbol{\mathcal{A}}_0$, every $\epsilon-$improvement path generated by* Eq. (4.8)*, converges to a ϵ–GNE in finite-time.*

Proof. Because of Lemma 3 it suffices to prove the statement if Γ is an exact relative potential game with respect to $\mathcal{A}_0$. By the definition of $\mathcal{F}_i^c(\boldsymbol{\sigma}_{-i}, h_i)$ in equation Eq. (3.5) it follows that the evolutionary dynamics Eq. (4.8) are positively invariant w.r.t $\boldsymbol{\mathcal{A}}_0$. That is,

$$\boldsymbol{\sigma}(t) \in \boldsymbol{\mathcal{A}}_0, \quad \forall t \geq 0. \tag{3.12}$$

Because Γ is a bounded game from Definition 16, it follows that P must be bounded as well. That is,

$$\exists M \in \mathbb{R}_+ : \quad |P(\boldsymbol{\sigma})| \leq M, \ \forall \boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}_0 \tag{3.13}$$

To prove that the game has the AFIP (see Preliminaries 2), a classic argument can be used based on a contradiction. Suppose γ is an infinite ϵ–improvement path. Denote the unique deviator at time t as i_t. By definition, if $i = i_t$ then

$$\pi_i(t+1) - \pi_i(t) > \epsilon,$$

if and only if

$$P(\sigma_i(t+1), \boldsymbol{\sigma}_{-i}(t+1)) - P(\sigma_i(t), \boldsymbol{\sigma}_{-i}(t)) > \epsilon. \tag{3.14}$$

This implies that

$$P(t) - P(0) > t\epsilon \Leftrightarrow P(t) > t\epsilon + P(0) \tag{3.15}$$

Then, for every $\epsilon > 0$

$$\lim_{t\to\infty} P(t) = \infty. \tag{3.16}$$

Because P is a bounded function this is a contradiction. Hence, every ϵ–improvement path terminates after a finite number of time steps T. At which it holds that

$$P(\boldsymbol{\sigma}(T)) \leq M < P(\boldsymbol{\sigma}(T)) + \epsilon \Rightarrow P(\boldsymbol{\sigma}(T)) > M - \epsilon.$$

This completes the proof. □

Remark 2. *The concept of generalized ordinal potential games also exists for convex games in which an increase in the payoff of the unique deviator implies an increase in the generalized ordinal potential function. However, for this class of convex games, in general, the bounded payoff functions do not imply the generalized ordinal potential function is bounded and hence one cannot guarantee convergence. If one, however, assumes this generalized potential function is bounded for every* $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{A}}_0$*, then the result in Theorem 2 carries over to this more general class of convex games.*

3.4 Networks of best and h-relative best responders

We have shown that the dynamics of network games in which all players choose h-relative best responses converge to a generalized Nash equilibrium. And that due to the their non-innovative nature, the relative best response dynamics converge for a more general class of games than best response dynamics. This also implies that any homogeneous action profile, in which all players choose the same action is a trivial generalized Nash equilibrium. Indeed, *payoff monotone* imitation dynamics share this property. In reality, *noise* in the decision-making process will destabilize most of these trivial equilibrium profiles. A characterization for the stochastically stable states of network games is beyond the scope of this chapter. Instead, we investigate an interesting scenario in which both best responses and relative best responses occur in the network game. In this case, it is not guaranteed that a homogeneous action profile is an equilibrium and the behavior may closer to real-world scenarios in which decision-makers value social information in different ways. And hence, the mixture of rationality and social learning can lead to more realistic outcomes. For simplicity, we assume that players always best respond or always relative best respond, and thus do not switch between the two decision rules. Although this is a simplification, it is a reasonable one that may be motivated by the empirical findings in [83] that suggest humans tend to *consistently* apply a decision rule under a variety of contexts. The following result follows immediately from the proofs of Theorem 1 and Theorem 2 and we omit its proof.

Corollary 3. *For a weighted potential game Γ_c, in which players consistently choose best responses as in Definition 10 or consistently choose h-relative best responses as in Definition 12, for every $\epsilon > 0$, and initial condition, every $\epsilon-$improvement path generated by* Eq. (4.8)*, converges to a ϵ-GNE in finite-time. The same holds for generalized ordinal potential games Γ_f, with $\epsilon = 0$.*

For the convergence analysis of a mixture of best responders and h-relative best responders no new theory is required. However, having both types of decision-makers in a network game can lead to significantly different behavior and equilibrium profiles that have not yet been studied in the context of network games. As more and more engineering systems take into account the complex behavior of humans, one may be interested in how different levels of social learning or different topologies of local information flows, affect the long-run behavior of economic decision making models. In the remainder of the chapter, we investigate the various effects that social learning through h-relative best response can have in economic models related to product adoption.

3.5 Competing products with network effects

Suppose there are two competing substitute products X and Y on the market and every player is using one of the two. Each product has an associated price $\gamma > 0$ and $\lambda > 0$ and individuals decide which product to use. Note that we are not modeling how a certain initial product adoption came to be, but we are interested if in the long run one of the technologies becomes dominant or not. However, it is worth mentioning that the adoption of a new product can be modeled in a very similar manner. Let $s_i = 1$ and $s_i = 0$ denote that player i uses product X and Y, respectively. Due to *network effects* [90] the utility that an individual experiences from these products partially depends on the number of individuals that are using it. Individuals may perceive this network effect differently but in general, a growing number of users increases the utility of the product. To this end, let $S = \sum_{i \in \mathcal{V}} \sigma_i$ denote the number of players in the network that are using product X. Then, the network effect X is modeled with an affine function, that is for all $i \in \mathcal{V}$

$$G_i(S) := aS + b_i, \quad a > 0,\ b_i \geq 0.$$

Because Y and X are substitutes, their network effects have a negative correlation. Such that, for all $i \in \mathcal{V}$

$$H_i(N - S) := d(N - S) + f_i, \quad d > 0,\ f_i \geq 0$$

The individual network effect parameters b_i and f_i may reflect how important beneficial network effects are for a player. For example, if b_i is relatively large, the player is eager to use product X even though the network effect is small. In a simplified model in which $d_i = 0$ and the player simply needs to choose to adopt a new product the players with a high b_i represent “early adopters” and players with a low b_i can represent “laggards” [91]. The utility that player $i \in \mathcal{V}$ obtains from using X or Y are given by

$$G_i(S) - \gamma, \text{ and } H_i(N - S) - \lambda.$$

Hence, the payoff of a player is

$$\pi_i(\sigma_i, \sigma_{-i}) = \left[G_i(S) - \gamma\right]\sigma_i + (1 - \sigma_i)\left[H_i(N - S) - \lambda\right].$$

Then, the following function is an exact potential function for the competing product game

$$P(\boldsymbol{\sigma}) = \sum_{i=1}^{N}(b_i + \lambda - dN - \gamma - f_i)s_i - d\sum_{i=1}^{N} s_i + (a + d)\left[\sum_{i=1}^{N} s_i^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{j \neq i}^{N} s_i s_j\right].$$

Because this competing product game is an exact potential game, Theorem 1 applies. Moreover, any mixture of best responders and h-relative best responders the fraction of the population using product X and Y will converge to a generalized Nash equilibrium in finite-time (Corollary 3).

Remark 3 (Mixed strategy extension)**.** *Because the competing products model is an exact potential game, it follows that its mixed-strategy extension, in which the players* *choose the fraction of time to use product X or Y, is also a potential game [42, Lemma 2.10]. And thus, the convergence results for h-relative best response dynamics in convex games are valid in this game. Such a setting can represent the dynamics of Example 2, in which the network effect of renewable energy can represent an increasingly cleaner environment.*

The addition of h-relative best responses is of particular interest in this model because they add a social influence to the competing product game that is not captured by best responses in which decisions of a player are solely based on the aggregate network effects and the cost and benefit parameters of the player. For relative best responses, the local information exchanges in the underlying social network of the players will affect their decisions.

Fig. 3.5 shows that when $h = 1$ the variation in the fraction of X adopters in the network is significantly larger than in myopic best response dynamics. These simulation results were obtained for 100 random initial conditions with $\pm 50\%$ adopters of product X. The slopes of the network effects are: $a = 0.15$ and $b = 0.12$. To introduce variation in the individual payoffs, the offsets b_i and f_i were randomly chosen between 0 and 10. The costs associated with the products are $\gamma = 3$ and $\lambda = 2$. The large variation in the standard deviation of the X adopters in the network is also typical for imitation dynamics and can be attributed to the variation in the initial action profiles, the stochasticity of the activation sequence and the large variety of generalized Nash equilibrium profiles in the product adoption game. From the blue line in Fig. 3.5 it can also be seen that, on average, the relative performance considerations in the 1-RBR dynamics allow for significantly higher adoption rates of product X that has a higher cost ($\gamma > \lambda$), but also a larger slope of the network effect ($a > b$). Naturally, these social effects are rather sensitive to the payoffs. In particular, for large networks, the network effect in the payoff can become dominant and an obvious best choice may arise that dominates under both types of dynamics. Fig. 3.6 shows another simulation on a similar network under the same conditions as in Fig. 3.5. One can observe that the 1-RBR dynamics have very similar qualitative behavior as imitation dynamics in which players imitate their best performing neighbor.

A typical feature of h-RBR dynamics is shown in Fig. 3.7. As h increases up to the point that all players employ relative best responses, the standard deviation in the fraction of X adopters tends to decrease. Interestingly, even for random initial conditions, the network structure causes significant differences in the behavior between best response and relative best response dynamics (shown in Fig. 3.5). However, these differences decrease when the connectivity in the network increases. In Fig. 3.8, the extreme case of a well-mixed network is shown and it can be seen that the behavior of the two types of dynamics are very similar.

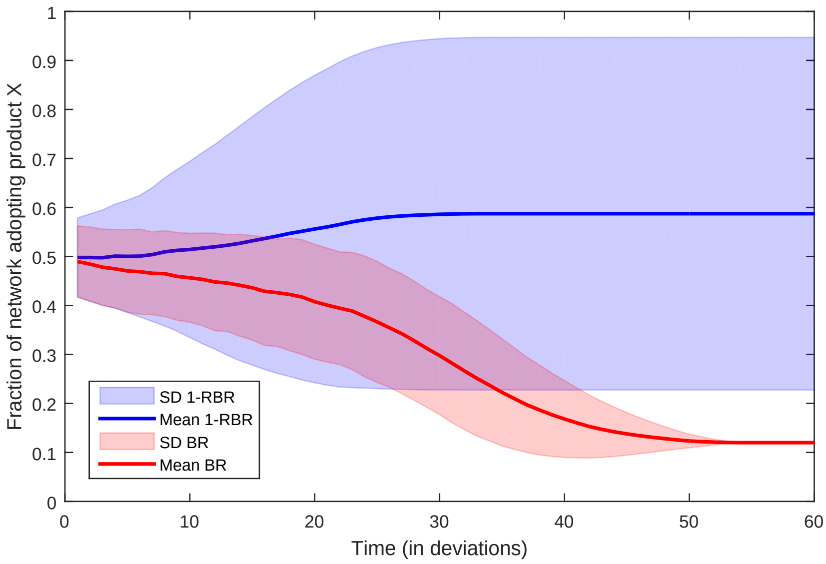


Figure 3.5: Simulations for the product adoption game with a *preferential attachment* network [92] with 50 players. The solid lines represent the mean of 100 iterations with random initial conditions. The shaded areas represent the standard deviation of the fraction of players adopting product X over all 100 iterations.

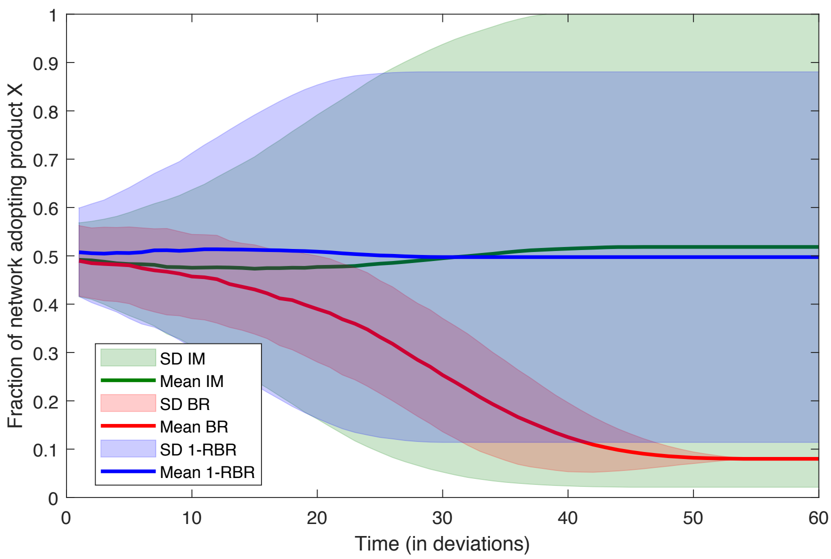


Figure 3.6: Another simulation of the product aoption game that compares myopic best response, imitate-the-best (indicated by IM) and 1-RBR dynamics on a preferential attachment network of size 50.

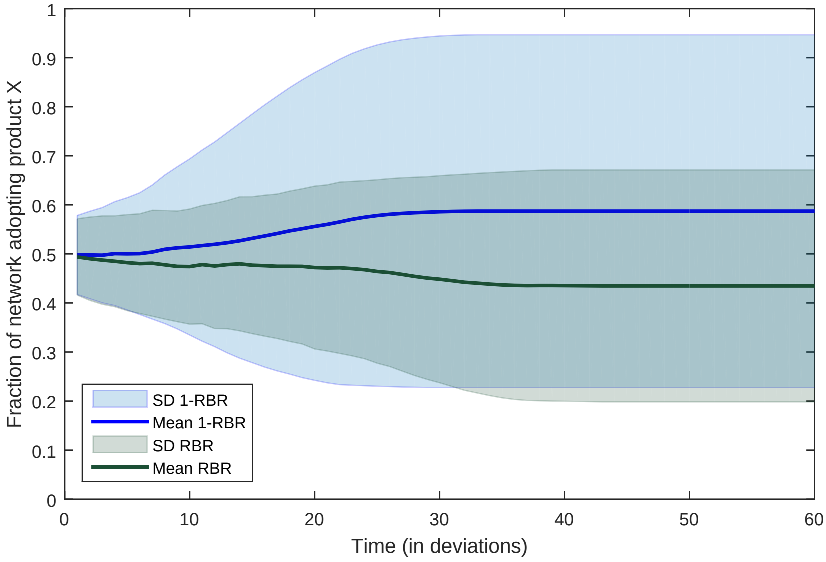


Figure 3.7: The effect of h on the fraction of players in the network that adopt X. Observe that the variation in the fraction reduces as h becomes larger.

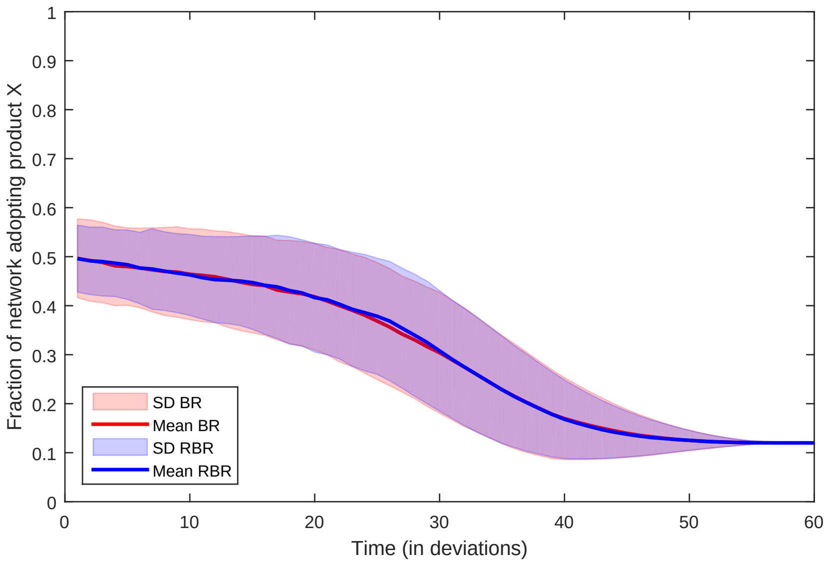


Figure 3.8: The product adoption game on a complete network with 50 players under best response and relative best response dynamics. Conditions are as described in the main text.

3.6 Final Remarks

We have introduced novel dynamics for finite and convex network games that result from an intuitive mix of rational best responses and social learning by imitation. It was shown that for a general class of games these dynamics converge to a generalized Nash equilibrium and that the corresponding decision-making process is "compatible" with best response dynamics. That is, any mix of best responders and h-relative best responders will eventually reach an equilibrium action profile. These results make it possible to rigorously study how relative performance considerations of "irrational" or *conforming* decision makers affect the behavior and equilibrium profiles of complex socio-technical and socio-economic processes. These effects are especially important for technological challenges that require increasingly complex models of large social systems that, in reality, are often affected by social learning effects that are not present in best responses.

In the next chapter, we will couple relative best response dynamics to imitation in finite games and study how *rational imitation* can significantly alter the decisions at equilibria of social dilemmas.

CHAPTER 4

Imitation, rationality and cooperation in spatial public goods games

> Imitation is not just the sincerest form of flattery– it's the sincerest form of learning.
>
> *George Bernard Shaw*

IMITATION and rationality are two seemingly paradoxical behaviors that are often observed in real-life decision-making processes. For example, companies can make investment decisions based on deliberate benefit-to-cost analysis or simply decide to invest because a successful competitor has done so already [93–95]. Likewise, the adoption of a product can be motivated by others, or because it provides some immediate benefit for an "innovator" [91]. Indeed, whereas rationality is often coupled to *innovation*, imitation is often linked to *social learning* [84]. Also in the literature of game theory, myopic best response dynamics are known as *innovative* dynamics because they can introduce actions that were not played before [37]. Dynamics based on imitations do not share this innovative feature. It is for this reason that decision-making processes based on imitations can have significantly different equilibrium profiles than those processes based on best responses.

The rationality of best responses is in line with economic theories that suggest *absolute optimization* is a natural result of evolutionary forces [70]. For these types of dynamics, the long-run collective behavior of non-cooperative network games, in which players interact exclusively with a fixed group of neighbors, have been studied extensively [42–45, 47–49, 69] More biologically inspired imitation dynamics on spatial structures defined by graphs have, amongst others, been studied in [34, 37, 68, 96]. One of the key drivers of *evolutionary graph theory* is to identify conditions and mechanisms under which *cooperation* can emerge, evolve and persist in social dilemmas in which the individual incentives are contradictory to the benefit of the system as a whole [8]. As mentioned before, in such social dilemmas, myopic optimizations tend to generate outcomes with payoffs that are far from the system optimum: a situation known as the tragedy of the commons [2, 50]. Spatial structure in the game-play interactions potentially overcomes this problem by means of *network reciprocity* through which cooperators can succeed by forming network clusters in which they help each other [8, 40]. However, for innovative rational dynamics in which players can introduce new actions at any given time, network reciprocity is far less effective in promoting cooperative actions. The main reason is that within a cluster of cooperators the best response of a player is to *defect*. Hence, the clusters can break down relatively easily if players can best respond and introduce new actions.

The convergence properties related to decision-making processes based on myopic optimizations are well understood. However, because the mechanisms for the evolution of cooperation tends to be less effective, more often than not at the equilibrium of a social dilemma, the players need to be satisfied with relatively low payoffs. And even though imitation based dynamics can lead to better outcomes, the mathematical study of their dynamics on spatial games is a challenging problem: first, because there typically exist a multitude of possible non-trivial outcomes [68] and second, because the existing optimization techniques, used in the analysis of rational best responses, are not applicable. And indeed, imitations can easily *prevent* the decision process to converge to an equilibrium at which all players are satisfied with their decisions [97, Ch. 10], [98]. We study the effects of rationality and imitation on the convergence properties and cooperation levels in a social dilemma model known as the spatial public goods game [96] and study the properties of *rational imitation*. Under rational imitations, players apply the *rationality principle* in their decisions to imitate a relatively successful neighbor or not. That is, actions are imitated only if they are expected to be efficient in terms of one's own success. This combination of rational decisions and imitation leads to beneficial dynamic features in the social dilemma that cannot be explained only by best responses or imitation. Hence, rational imitations can open the door to the design of novel methods for complex systems that have to rely on large scale cooperation for the maintenance of publicly available goods.

4.1 Spatial public goods games

In n-player games on networks, for each player i, a graph $\mathcal{G}$ defines a *group* of players $\bar{\mathcal{N}_i} = \mathcal{N}_i \cup \{i\}$, known as the closed *neighborhood* of player i, that play a *public goods game (PGG)* referred to as the game *centered around player i*. Therefore, in total, every player i participates in $|\mathcal{N}_i| + 1$ games; one centered around herself and $|\mathcal{N}_i|$ centered around her neighbors. Every player i chooses an action $\sigma_i \in \{0, 1\}$ that is either to *cooperate* ($\sigma_i = 1$), namely to contribute a fixed amount $c_i > 0$ to a publicly available resource called the *public good*, or *defect* ($\sigma_i = 0$), namely to contribute nothing. The player employs this same action in all of the games that she participates in [96]. This applies to the cases when players do not have the cognitive capabilities to discriminate between co-players, or when there is only one public good and the contribution scales with the degree of the player. In the next chapter, we will study a setting in which this "one-action" assumption is relaxed. Cooperators and defectors profit equally from the public good: all players participating in a game centered around player $j \in \mathcal{V}$, evenly share the *production* of that game defined by $p_j : \{0, 1\}^{|\bar{\mathcal{N}_j}|} \to \mathbb{R}$, which is a function of the actions of the players in the neighborhood $\bar{\mathcal{N}_j}$ denoted by the vector $\boldsymbol{\sigma_j} := \{\sigma_l : l \in \bar{\mathcal{N}_j}\}$. The local payoff of an player i upon the participation in this game is, hence,

$$\pi_{ij}(\boldsymbol{\sigma_j}) = \frac{p_j(\boldsymbol{\sigma_j})}{|\mathcal{N}_j| + 1} - c_i\sigma_i \quad \forall i \in \bar{\mathcal{N}_j}. \tag{4.1}$$

Production functions are, typically, non-decreasing in the number of cooperators, reflecting that contributions increase the public good. We relax this monotonicity assumption, allowing us to study the maintenance of artificially scarce goods known as *club goods*. The total payoff of player i is a weighted summation of the local payoffs she earns at each of the $|\mathcal{N}_i| + 1$ games:

$$\pi_i(\boldsymbol{\sigma}) = \sum_{j \in \bar{\mathcal{N}_i}} \lambda_j \pi_{ij}(\boldsymbol{\sigma_j}), \tag{4.2}$$

where $\lambda_j \in \mathbb{R}, j \in \mathcal{V}$, represents the relative importance of the game centered around player j, and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)^\top \in \{0, 1\}^n$ is the collective action profile of all the players. We denote the combined payoff function by $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)^\top$ and the spatial public goods played under the network $\mathcal{G}$ by $\Gamma = (\mathcal{G}, \boldsymbol{\pi})$. In an alternative spatial representation, the group structures are determined by a bipartite graph $\mathbb{B} = (\mathcal{M}, \mathcal{V}, \mathcal{K})$, with the player set $\mathcal{V}$, a set of non-empty group structures $\mathcal{M} \subset 2^{\mathcal{V}}$ and edge set $\mathcal{K} \subset \mathcal{V} \times \mathcal{M}$ (Fig. 4.1) [99]. The biadjacency matrix of $\mathbb{B}$, denoted by $\boldsymbol{B} = [b_{ji}] \in \mathbb{R}^{|\mathcal{M}| \times |\mathcal{V}|}$, is defined such that $b_{ji} = 1$ if and only if $(j, i) \in \mathcal{K}$ and zero otherwise, for all $i \in \mathcal{V}, j \in \mathcal{M}$. So the j^{th} row of the biadjacency matrix $\boldsymbol{B}$

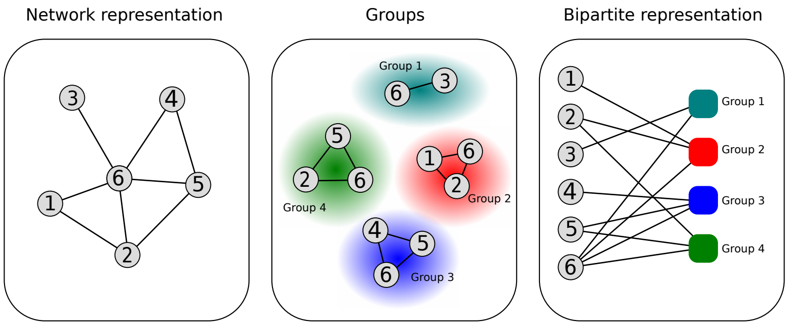


Figure 4.1: Group interaction on a given network can be represented by the neighborhood hypergraph of a network [96]. When the social interaction network is constructed from information of the group structure itself (middle), the interactions can alternatively be represented by a bipartite graph (right) in which the players are assigned to those groups in which they interact [99]. In this example, because of the central role of player 6, the network representation that is a one mode projection of the bipartite graph, induces different group structures than those in bipartite representation. Therefore the behavior of a spatial public goods game for the two representations differ. The figure is adapted from [99].

determines which players interact in the PGG played in group j. Hence, the number of players in a group $j \in \mathcal{M}$ equals $\sum_{i\in\mathcal{V}} b_{ji} > 0$. The payoff obtained in group $j \in \mathcal{M}$ and the total payoff of player $i \in \mathcal{V}$ are, thus,

$$\pi_{ij}(\boldsymbol{\sigma_j}) = \frac{p_j(\boldsymbol{\sigma_j})}{\sum_{i\in\mathcal{V}} b_{ji}} - c_i\sigma_i, \;\; \pi_i(\boldsymbol{\sigma}) = \sum_{j\in\mathcal{M}} b_{ji}\pi_{ij}(\boldsymbol{\sigma_j}). \tag{4.3}$$

We denote the spatial PGG with the bipartite representation $\mathbb{B}$ and payoff functions Eq. (4.3) by $\Gamma^{\mathrm{b}} = (\mathbb{B}, \boldsymbol{\pi})$.

Example 5 (Homogeneous linear public goods game)**.** *The simplest and most widely studied production function scales linearly with the number of cooperators in the game and every player has the same contribution $c > 0$, i.e.,*

$$p_i(\boldsymbol{\sigma}) = rc \sum_{j\in\mathcal{N}_i} \sigma_j, \quad 1 < r < N. \tag{4.4}$$

Here, r is the public good multiplier that can be seen as the benefit-to-cost ratio of the game. It is worth to mention that, even though the parameters c and r are the same for every player, a non-regular network structure introduces asymmetries in the players' payoffs and contributions as detailed in [96], and therefore, th corresponding asynchronous dynamics can lead to complex behaviors.

4.2 Rational and unconditional imitation update rules

Each player is associated with an action at time $t = 0$, and at every time $t \in \mathbb{Z}_{\geq 0}$, a single player i_t becomes active to update her action at time $t+1$ based on some *update rule*, resulting in a dynamical decision making process called an *asynchronous spatial public goods game*. We consider two update rules: *unconditional imitation* and *rational imitation*. The unconditional imitation rule dictates that player i active at time t updates her action at $t+1$ to that of one of the top h_i highest-earning in her neighborhood, $h_i \in \{1, \dots, |\bar{\mathcal{N}}_i|\}$:

$$\sigma_i(t+1) \in \mathcal{I}_i^{\mathrm{u}}(\boldsymbol{\sigma}(t)), \tag{4.5}$$

where $\mathcal{I}_i^{\mathrm{u}}$ provides the set of actions of the relatively successful players:

$$\mathcal{I}_i^{\mathrm{u}}(\boldsymbol{\sigma}) := \left\{ \sigma_j \,\middle|\, j \in \arg\max_{j \in \bar{\mathcal{N}}_i}{}^{h_i} \pi_j(\boldsymbol{\sigma}) \right\}. \tag{4.6}$$

When $h_i = 1$, unconditional imitation recovers "imitate-the-best" decisions [68], where only the most successful players can be imitated. Unconditional imitation may be seen as an *irrational* decision since players do not take into account their own *expected* payoff change that results from imitating their neighbors' actions. Arguably, imitation becomes more rational when the expected payoff change *is* taken into account. This is, for instance true under best-response dynamics in which players choose actions that optimize their own payoffs against the current actions of their opponents. Similarly, under rational imitation, players seek to improve their payoffs myopically, but are restricted to only copy their neighbors' actions. In other words, (similar to unconditional imitation) players copy their neighbors' actions (yet) only if it improves their own payoffs. More specifically, under rational imitation,

$$\sigma_i(t+1) \in \mathcal{I}_i^{\mathrm{r}}(\boldsymbol{\sigma}(t)), \tag{4.7}$$

where $\mathcal{I}_i^{\mathrm{r}}$ provides those actions of the top h_i-earning players in the neighborhood of player i who earn no less than player i herself:

$$\mathcal{I}_i^{\mathrm{r}}(\boldsymbol{\sigma}) := \{ y \in \mathcal{I}_i^{\mathrm{u}}(\boldsymbol{\sigma}) \cup \{\sigma_i\} \,|\, \pi_i(y, \boldsymbol{\sigma}_{-i}) \geq \pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) \}.$$

Note that a rational imitation differs from a "relative best response" [74, 75], in the sense that a rational imitation only requires the imitated action in the feasible action profile to be a better reply, not necessarily a best reply.

Remark 4. *The set $\mathcal{I}_i^{\mathrm{u}}(\boldsymbol{\sigma})$ may also be defined under the that players only take into account the actions of neighbors that receive a higher payoff than themselves (see also Assumption 5). This modification of the feasible action set does not affect the results presented in the remainder of the chapter.*

4.2.1 Asynchronous imitation dynamics

The players' activation sequence $\{i_t\}_{t=0}^{\infty}$ together with the imitation update rule govern the evolution of the players' actions over time, resulting in *asynchronous PGG dynamics*. Namely, for every time $t \in \mathbb{Z}_{\geq 0}$, there exists a unique player $i_t \in \mathcal{V}$ such that the collective action dynamics satisfy

$$\boldsymbol{\sigma}(t+1) \in (\mathcal{I}_{i_t}(\boldsymbol{\sigma}(t)), \boldsymbol{\sigma}_{-i_t}(t)), \tag{4.8}$$

with $\mathcal{I}_i(\cdot) = \mathcal{I}_i^{\mathrm{r}}(\cdot)$ for rational imitation dynamics and $\mathcal{I}_i(\cdot) = \mathcal{I}_i^{\mathrm{u}}(\cdot)$ for unconditional imitation dynamics. We assume Assumption 2, i.e. the activation sequence is *persistent*.

In the long-run, the dynamics either reach an equilibrium action profile in which all players are satisfied with their decisions *or* undergo perpetual oscillations in which a subset of the players do not reach a satisfactory decision and, not necessarily periodically, imitate each other's actions indefinitely [68]. We call the action profile $\boldsymbol{\sigma}^* \in \{0,1\}^N$ an *imitation equilibrium* if

$$\sigma_i^* \in \mathcal{I}_i(\boldsymbol{\sigma}^*) \quad \forall i \in \mathcal{V}. \tag{4.9}$$

In the following section, we study the asymptotic behavior of the asynchronous PGG dynamics under rational and unconditional imitation.

4.3 Finite time convergence of imitation dynamics

4.3.1 Rational Imitation

The imitation update rule Eq. (4.7) maps the action of the active player to a set of actions of size at most two. If the set include both cooperation and defection, the player can pick any of the two. We postulate the following assumption to ensure that players switch to another action only if they have an *incentive*, i.e., earn more.

Assumption 4 (Incentive to deviate)**.** *For player i active at time t, $\sigma_i(t) \neq \sigma_i(t+1)$ only if there exists an action $y \in \mathcal{I}_i^{\mathrm{r}}(\boldsymbol{\sigma})$ such that*

$$\pi_i(y, \boldsymbol{\sigma}_{-i}(t)) > \pi_i(\sigma_i(t), \boldsymbol{\sigma}_{-i}(t)).$$

The assumption is another reason why rational imitation can be considered to be a *rational* decision: should the player's expected payoff at the next time step not exceed its current payoff, the player would not deviate. This allows us to obtain the following general result.

Theorem 3 (Finite time convergence under rational imitation)**.** *Under Assumption 4, any asynchronous spatial PGG governed by the rational imitation update rule reaches an imitation equilibrium in finite time.*

Proof. We first show that the local game with the players in $\bar{\mathcal{N}_i}$ with payoff function Eq. (4.1) is an exact potential game. Consider the candidate potential function for a local interaction

$$\psi_i(\boldsymbol{\sigma}) = \frac{p_i(\boldsymbol{\sigma_i})}{|\mathcal{N}_i|+1} - \sum_{j\in\mathcal{N}_i\cup\{i\}} c_j\sigma_j. \tag{4.10}$$

The local payoff difference from a deviation of any player $j \in \mathcal{N}_i \cup \{i\}$ switching from $\sigma_j = 0$ is

$$\pi_{ji}(0,\boldsymbol{\sigma}_{-j}) - \pi_{ji}(1,\boldsymbol{\sigma}_{-j}) = \frac{p_i(0,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} - \frac{p_i(1,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} + c_j,$$

and the difference in the potential function is $\psi_i(0,\boldsymbol{\sigma}_{-j}) - \psi_i(1,\boldsymbol{\sigma}_{-j})$, which reads as

$$\begin{aligned}
&\frac{p_i(0,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} - \sum_{l\in\mathcal{N}_i\cup\{i\}} c_l\sigma_l - \frac{p_j(1,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} + \sum_{k\in\mathcal{N}_i\cup\{i\}} c_k\sigma_l \\
&= \frac{p_i(0,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} - \sum_{k\in\mathcal{N}_i\cup\{i\}} c_k\sigma_l - \frac{p_j(1,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} + \sum_{k\in\mathcal{N}_i\cup\{i\}\setminus\{j\}} c_k\sigma_l + c_j \\
&= \frac{p_i(0,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} - \frac{p_i(1,\boldsymbol{\sigma}_{-j})}{|\mathcal{N}_i|+1} + c_j.
\end{aligned}$$

It follows that $\pi_{ji}(0,\boldsymbol{\sigma}_{-j}) - \pi_{ji}(1,\boldsymbol{\sigma}_{-j}) = \psi_i(0,\boldsymbol{\sigma}_{-j}) - \psi_i(1,\boldsymbol{\sigma}_{-j})$. Naturally, the equality holds for the opposite switch as well. Moreover, observe that for all $v \notin \mathcal{N}_i \cup \{i\}$, $\psi_i(\sigma_v, \sigma_{-v}) - \psi_i(\sigma'_v, \sigma{-}v) = 0$. Indeed, when the unique deviator is not a member of the closed neighborhood, any payoffs of the players obtained in this local game do not change. We proceed to show that the function $P(\sigma) = \sum_{i=1}^{N} w_i\psi_i(\sigma)$ is a potential function for the aggregated payoff function Eq. (4.2). To see this, note that $\pi_j(\sigma_j,\boldsymbol{\sigma}_{-j}) - \pi_j(\sigma'_j,\boldsymbol{\sigma}_{-j})$ reads as

$$\begin{aligned}
\sum_{k\in\mathcal{N}_j} w_k\left(\pi_{jk}(\sigma_j,\boldsymbol{\sigma}_{-j}) - \pi_{jk}(\sigma'_j,\boldsymbol{\sigma}_{-j})\right) &= \sum_{k\in\mathcal{N}_j} w_k\left(\psi_k(\sigma_j,\boldsymbol{\sigma}_{-j}) - \psi_k(\sigma'_j,\boldsymbol{\sigma}_{-j})\right) \\
&= \sum_{i=1}^{N} w_i\left(\psi_i(\sigma_j,\boldsymbol{\sigma}_{-j}) - \psi_i(\sigma'_j,\boldsymbol{\sigma}_{-j})\right).
\end{aligned} \tag{4.11}$$

The last equality in Eq. (4.11) is because for all $k \notin \mathcal{N}_j$, $\psi_k(\sigma_j, \boldsymbol{\sigma}_{-j}) - \psi_k(\sigma_j', \boldsymbol{\sigma}_{-j}) = 0$. It follows that the spatial PGG is an exact potential game. To finish the proof we use the concept of the *Finite Improvement Property* (FIP) that is defined in the preliminaries in Chapter 2. Because of Assumption 3, for every $h = (h_i)_{i\in\mathcal{V}}$, the rational imitation dynamics generates improvement paths and because we have shown that the PGG is a potential game, by Lemma 1 each such improvement path terminates in a finite time. This completes the proof. □

Theorem 3 shows that for a general class of PGGs, i.e. with heterogeneous contributions and arbitrary production functions, the rational imitation dynamics are guaranteed to converge to an imitation equilibrium in finite time. For the bipartite representation of the PGG, we have the following result.

***Theorem* 4** (Finite time convergence in bipartite representations)**.** *Under Assumption 3, every asynchronous spatial PGG with a bipartite group structure and governed by the rational imitation update rule reaches an imitation equilibrium in finite time.*

Proof. The proof can be obtained by substituting the expressions for the local payoff and the total payoff in Eq. (4.3) into the payoff expressions in the proof of Theorem 3 and use the local potential function for the payoffs in group $j \in \mathcal{M}$

$$\psi_j(\boldsymbol{\sigma}) = \frac{p_j(\boldsymbol{\sigma_j})}{\sum_{i\in\mathcal{V}} b_{ji}} - \sum_{i\in\mathcal{V}} b_{ji} c_i \sigma_i, \tag{4.12}$$

and the potential function for the complete payoffs as $\sum_{j\in\mathcal{M}} \psi_j(\boldsymbol{\sigma})$. □

Remark 5. *It is worth to mention that the proofs of Theorems 3 and 4 imply that for these general classes of spatial PGG, best response dynamics will converge to a pure Nash equilibrium in finite time and the stationary distribution of log-linear learning in spatial public goods games can be characterized analytically for both representations of the PGG [45, Theorem 6.1].*

In Section 4.4, we will discuss how the imitation equilibria of rational imitation can significantly differ from the long-run behavior of rational innovative dynamics, and in some cases also from unconditional imitation dynamics. Before doing so, let us take a closer look at the convergence properties of the unconditional imitation dynamics for the spatial public goods game in which the group structures are determined by a neighborhood hypergraph.

4.3.2 Unconditional imitation

The behavior of decision processes based on unconditional imitations is a challenging open problem because the generated paths in the combined action profile are not necessarily *improvement paths*: by copying the action of a successful neighbor, a player may decrease its payoff, even if all other players do not change their actions. A second complicating factor is that imitations are limited to direct neighbors whereas the payoffs of the players also depend on two-hop neighbors. This creates an asymmetry in the spatial structure: the interaction graph that determines one's payoff and the replacement graph that determines one's feasible action set are, in general, not equal. Indeed, equilibration is not guaranteed under unconditional imitation for arbitrary spatial structures. For example, even for the relatively simple homogeneous linear PGG in Example 5, imitating the best performing neighbors can lead to persistent oscillations (Fig. 4.2). Nevertheless, we will discuss in Section 4.4 how these 'inconvenient' properties of imitation dynamics can be beneficial for the maintenance of publicly available goods. Here, our goal is to identify spatial structures that *do* allow the players in the homogeneous linear PGG to reach a satisfactory decision. We restrict our analysis to *'imitate the best'* unconditional imitation dynamics, i.e., $h_i = 1$ for all $i \in \mathcal{V}$. Similar to the rational imitation case, here, we restrain the active player to arbitrarily switch actions; that is, imitation occurs only if the target action is more successful. Decision rules that have this property are called *payoff monotone* [37, 100, 101].

Assumption 5 (Payoff monotone [37, 100]). *For player i active at time t, $\sigma_i(t) \neq \sigma_i(t+1)$ only if there exists an player $j \in \mathcal{N}_i$ with $\sigma_j \in \mathcal{I}_i^{\mathrm{u}}(\boldsymbol{\sigma}(t))$ such that*

$$\pi_j(\boldsymbol{\sigma}(t)) > \pi_i(\boldsymbol{\sigma}(t)).$$

It can be easily shown that if Assumption 5 holds and the network is fully connected (complete), then the linear homogeneous PGG converges to full defection for every initial action profile in which defectors exist. This is in line with experimental results in [78], that indicate that focusing on the success of others leads to selfish behavior in complete network games. This observed highly defective tendency does not necessarily occur in more complex spatial structures. We proceed to one of the minimally-connected network structures: a star.

Lemma 4. *Consider a linear homogeneous PGG played on a star network. If the central player defects at some time instance, then the action profile generated by 'imitate the best' unconditional imitation dynamics will reach the full-defection imitation equilibrium in finite time.*

Proof. Let player 1 represent the central player and players l_c and l_d represent the cooperating and defecting leaves, respectively. The following result can be derived directly from Eq. (4.1), Eq. (4.2) and Eq. (4.4).

Lemma 5. *Consider a star network consisting of one central player and p cooperating and q defecting leaf players. The players' accumulated payoffs are given by*

$$\pi_1 = \frac{(\sigma_1 + p)r}{N} + \frac{\sigma_1 + 1}{2}pr + \frac{\sigma_1}{2}qr - (N+1)\sigma_1, \tag{4.13}$$

$$\pi_{l_c} = \frac{(\sigma_1 + p)r}{N} + \frac{1 + \sigma_1}{2}r - 2, \tag{4.14}$$

$$\pi_{l_d} = \frac{(\sigma_1 + p)r}{N} + \frac{\sigma_1}{2}r. \tag{4.15}$$

We now continue proving the main statement in Lemma 4. We prove by induction on m defined as the number of cooperating leaves when the central player is defecting. The result is trivial for $m = 0$, i.e., when all the players are defecting. Assume the result holds for $m = p - 1, p \geq 1$. Consider some time k_0 when the central player is defecting and there are p cooperating leaves in the network, i.e., $m = p$. If the active player at k_0 is a defecting leaf, she will not switch since her only neighbor is the central player who is also defecting. Hence, the state at $k_0 + 1$ will be the same as the initial state. So consider the first time $k_1 \geq k_0$ when a cooperating leaf is active. This time exists due to the persistent activation assumption. From Lemma 5, the payoffs at k_1 of the active player l_c and her neighbor, player 1, are given by

$$\left.\begin{aligned} \pi_1 &= \frac{pr}{N} + \frac{1}{2}pr \\ \pi_{l_c} &= \frac{pr}{N} + \frac{1}{2}r - 2. \end{aligned}\right\} \overset{p\geq 1}{\Longrightarrow} \pi_1 > \pi_{l_c}.$$

Therefore, the cooperating leaf will switch to defection at $k_1 + 1$, resulting in a new state where the central player is still defecting and there are $p - 1$ cooperating leaves. This is the case with $m = p - 1$, which completes the proof. □

Next, we consider to the case when the central player is cooperating and provide sufficient conditions for reaching the full-cooperation equilibrium and a mixed equilibrium in which cooperators and defectors coexist. We refer to the non-central players as leaf players.

Lemma 6. *Consider a linear homogeneous PGG played on a star network. Assume that initially the central player is cooperating and there are $p \geq 0$ cooperating and $q \geq 1$ defecting leaf players. Then*

- *if $r < \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$, then the network will reach the full-defection imitation equilibrium;*
- *if $r = \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$, the network is already at a mixed imitation equilibrium;*

- *if* $r > \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$, *then the network will reach the full-cooperation imitation equilibrium.*

Proof. Case 1: $r < \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$. It follows from Lemma 5 that $\pi_1 < \pi_{l_d}$. Now in case there are no cooperating leaves in the network, the central player switches to defection in the next time step, and hence, the networks reaches the full-defection equilibrium. So consider the case when there is at least one cooperating leaf in the network, i.e., $p \geq 1$. Then since $q \geq 1$, it holds that

$$3 < 3p + q \Rightarrow \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}} < 4 \Rightarrow r < 4.$$

Thus, from Lemma 5 it follows that $\pi_{l_c} < \pi_{l_d}$. Now since $\pi_1 < \pi_{l_d}$, it can be concluded that only the central player may switch at the next time step. Due to the persistent activation assumption, there exists some time that the central player becomes active, and the fist time when that happens, she will switch to defection. Then in view of Proposition 4, the network will reach the full-defection equilibrium state.

Case 2: $r = \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$. Then $\pi_1 = \pi_{l_d}$. Hence, neither the central player nor any of the defecting leaves will switch actions at the next time step. Trivially the same holds for every cooperating leaf, resulting in an equilibrium state.

Case 3: $r > \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$. Then $\pi_1 > \pi_{l_d}$. Hence, since the cooperating leaves do not switch actions, the first time that a defecting leaf is active, she will switch to cooperation. So the new number of cooperating and defecting leaves will be $\bar{p} = p + 1$ and $\bar{q} = q - 1$. Then the condition $r > \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$ for the new state becomes

$$r > \frac{\bar{p}+\bar{q}+1}{\bar{p}+\frac{1}{2}\bar{q}-\frac{1}{2}} = \frac{p+q+1}{p+\frac{1}{2}q},$$

which holds since $r > \frac{p+q+1}{p+\frac{1}{2}q-\frac{1}{2}}$. Hence, again a defecting leaf will switch to cooperation. Therefore, eventually the network will reach a full-cooperation equilibrium state. This completes the proof. □

***Theorem* 5** (Finite time convergence star networks)**.** *Every linear homogeneous PGG played on a star network reaches an imitation equilibrium in finite time.*

Proof. The proof follows from Proposition 4 and 6. □

After establishing convergence of a completely connected and a minimally connected network, we now proceed to a network that has a regular and close-to-minimal connectivity: a ring. Before stating the main result, let us postulate the following assumption that we assume to hold for some values of the public goods multiplier r.

Assumption 6 (Pairwise persistence)**.** *For any pair of players $i, j \in \mathcal{V}$ and each time k, there exists some finite time $k' > k$ such that i and j are activated consecutively at k' and $k'+1$.*

Although stronger than the persistent activation assumption, the pairwise persistent activation assumption still holds almost surely in most stochastic settings, particularly when players are activated independently, e.g., according to Poisson clocks or by a stochastic process where at each time step, one random player becomes active to alter its current action [46].

***Theorem* 6** (Finite time convergence ring networks)**.** *Consider a linear homogeneous PGG played on a ring network. If the public goods multiplier r belongs to the interval $\left[0, \frac{9}{2}\right]$ the 'imitate the best' unconditional imitation dynamics reach an imitation equilibrium in finite time. The same holds for $r > \frac{9}{2}$, but when the pairwise persistent activation Assumption 6 holds.*

Proof. We first show that the behavior of the homogenous linear PGG under the imitate the best unconditional imitation dynamics, although depending on the public-goods multiplier r, is the same for different values of r in certain ranges. In order to do this we first introduce the notation $\boldsymbol{\sigma}(k)_{|_{r=g}}$, which denotes the state vector at time k, given that $r = g \in \mathbb{R}_{\geq 0}$.

Lemma 7. *Given a ring network, its initial action vector and activation sequence, for every public goods multiplier r_1 and r_2 taken from one of the following intervals*

$$\left(0, \frac{9}{5}\right), \left(\frac{9}{5}, \frac{9}{4}\right), \left(\frac{9}{4}, \frac{9}{3}\right), \left(\frac{9}{3}, \frac{9}{2}\right), \left(\frac{9}{2}, \frac{9}{1}\right), \left(\frac{9}{1}, \infty\right),$$

it holds that for all time $k \geq 0$ $\boldsymbol{\sigma}(k)_{|_{r=r_1}} = \boldsymbol{\sigma}(k)_{|_{r=r_2}}$.

Proof. We prove by strong induction. The statement is trivial for $k = 0$. Assume that the statement is true for all $k \leq t$ for some $t \in \mathbb{Z}_{\geq 0}$. Let i denote the active player at time t. It suffices to show that all r that belong to one of the above 6 intervals yield the same $\sigma_i(t+1)$. According to the unconditional imitation update rule with $h_i = 1$ for all $i \in \mathcal{V}$ in Eq. (4.6), player i's action at the next time step $\sigma_i(t+1)$ depends on the payoffs of player i and her neighbors at time t. The payoff of player i is determined by the action of herself, her neighbors and the neighbors of her neighbors (recall that player i also participates in the games centered around her neighbors). Due to the ring topology, this implies that π_i is determined by $\sigma_{i-2}, \sigma_{i-1}, \sigma_i, a_{i+1}$ and σ_{i+2}. By following the same argument for each of the neighbors of player i, i.e., $i-1$ and $i+1$, we conclude that the payoffs of player i and her neighbors are determined by the vector

$$s^i = \begin{pmatrix} \sigma_{i-3} & \sigma_{i-2} & \sigma_{i-1} & \sigma_i & \sigma_{i+1} & \sigma_{i+2} & a_{i+3} \end{pmatrix}. \tag{4.16}$$

Therefore, $\sigma_i(t+1)$ is completely determined by the actions in $s^i(t)$. Clearly s^i allows for 2^7 different action profiles. Some of the possible action profiles keep the action of player i unchanged at $t+1$, some others make player i switch and the rest of the possible action profiles require r to fulfill a certain condition in order for the action of player i to change. For example, if $s^i(t) = (1,1,1,1,1,1,1)$, then $\sigma_i(t+1) = \sigma_i(t)$. Moreover, if $s^i(t) = (1,0,0,1,0,0,1)$, then we obtain the following payoffs:

$$\pi_{i-1} = r, \quad \pi_i = r-3, \quad \pi_{i+1} = r.$$

Hence, $\sigma_i(t+1) \neq \sigma_i(t)$ since $\pi_{i-1}, \pi_{i+1} > \pi_i$, implying that player i's action changes regardless of r. However, if $s^i(t) = (1,0,1,0,1,0,1)$, then we obtain the following payoffs:

$$\pi_{i-1} = \frac{5}{3}r - 3, \quad \pi_i = \frac{4}{3}r, \quad \pi_{i+1} = \frac{5}{3}r - 3.$$

Hence, $\sigma_i(t+1) \neq \sigma_i(t)$ if and only if $\pi_{i-1} = \pi_{i+1} > \pi_i$, resulting in $r > 9$. By investigating all 128 values of s^i, we obtain the following critical values of r, so that for a given $s^i(t)$, all values of r between any two consecutive critical values result in the same $\sigma_i(t+1)$:

$$0, \quad \frac{9}{5}, \quad \frac{9}{4}, \quad \frac{9}{3}, \quad \frac{9}{2}, \quad 9.$$

This proves the statement for $k = t+1$, which completes the proof. □

We now continue with proving the statement in Theorem 6.

For ring networks consisting of 5 players or fewer, the result can be verified by exhausting all the cases. For ring networks consisting of more than 5 players, we show the result only for the following two cases. Other cases can be handled similar to Case 1. Based on Lemma 7, we prove the theorem for each of the following cases:

Case 1: $0 < r < \frac{9}{5}$. In view of Lemma 7, we only need to prove the result for just one value of r in this range, say $r = 1$. Consider the function $n_{1111}(\boldsymbol{\sigma}) : \{0,1\}^n \to \mathbb{Z}_{\geq 0}$ defined as the number of 4 consecutive cooperators in the whole network:

$$n_{1111}(\boldsymbol{\sigma}) = |\{j \in \mathcal{V} \mid \sigma_j = \sigma_{j+1} = \sigma_{j+2} = \sigma_{j+3} = 1\}|$$

where $|\mathcal{X}|$ denotes the cardinality of the set $\mathcal{X}$. We show that n_{1111} never decreases over time, i.e., at every time $K \geq 0$,

$$\Delta n_{1111}(K) = n_{1111}(\boldsymbol{\sigma}(K+1)) - n_{1111}(\boldsymbol{\sigma}(K)) \geq 0. \tag{4.17}$$

Let player i be active at K. Then all consecutive quadruple actions that may change in number at $K+1$ are

$$(\sigma_{i-3}, \sigma_{i-2}, \sigma_{i-1}, \sigma_{\boldsymbol{i}}), (\sigma_{i-2}, \sigma_{i-1}, \sigma_{\boldsymbol{i}}, \sigma_{i+1}),$$
$$(\sigma_{i-1}, \sigma_{\boldsymbol{i}}, \sigma_{i+1}, \sigma_{i+2}), (\sigma_{\boldsymbol{i}}, \sigma_{i+1}, \sigma_{i+2}, \sigma_{i+3}).$$

Therefore, it suffices to show that the number of these four quadruples equaling $(1,1,1,1)$ at $K+1$ is not less than the number of quadruples at K. Since all of these four quadruples are included in s^i defined in Eq. (4.16), it suffices to show that the number of quadruples $(1,1,1,1)$ in $s^i(K+1)$ is no less than that in $s^i(K)$. Again, the vector $s^i(K)$ allows for 2^7 different states. On the other hand, for each of these states, $s^i(K+1)$ can be determined uniquely based on $s^i(K)$ as discussed in the proof of Lemma 7 (this is because only player i's action may change at $K+1$, which is uniquely determined by the actions of herself and her three left and right neighbors at K). It is worth to mention that if the action that receives the maximum payoff in the neighborhood is not unique, then the binary action set implies that the player will not switch. Stack all 2^7 different states of $s^i(K)$ and $s^i(K+1)$ in two $2^7 \times 7$ binary-matrices S^- and S^+ so that for every row $j = 1,2,3,\ldots,2^7$, if $s^i(K) = S^-_j$, then $s^i(K+1) = S^+_j$ where X_j represents the jth row of matrix X. For every $j = 1,2,3,\ldots,2^7$, delete the jth rows of S^+ and S^-, if they are the same, to obtain S^+_0 and S^-_0. Then the rows of S^-_0 represent all possible values of $s^i(K)$ that will result in player i switching her action. One can check that the number of 4 consecutive 1's in every row of S^+_0 is **no less than** that in the same row in S^-_0. This implies that the number of quadruplets $(1,1,1,1)$ in $s^i(K+1)$ is no less than $s^i(K)$, regardless of what value $s^i(K)$ takes. Consequently, Eq. (4.17) holds. Hence, whenever an player switches, the function n_{1111} either increases or remains constant. Since n_{1111} is bounded, this yields the existence of some time k_1 at which n_{1111} becomes fixed and never changes afterwards. Consider the matrices S^-_1 and S^+_1 that are obtained from S^-_0 and S^+_0 after deleting each row j from both of them if the number of $(1,1,1,1)$s in the jth row of S^-_0 is less than that in S^+_0. Since n_{1111} is fixed after k_1, no switching that results in a change in the number of quadruples $(1,1,1,1)$ may take place after k_1. Hence, for $k \geq k_1$, $s^i(k)$ equals one of the rows of S^-_1 and $s^i(k+1)$ equals the corresponding row in S^+_1. Now one can check that the number of quadruples $(1,1,0,1)$ in every row of S^+_1 is **no less than** that in the same row in S^-_1. Hence, the function n_{1101} defined as the number of $(1,1,0,1)$s in the network never decreases after k_1. Therefore, similar to the argument above, there exists some finite time k_2 when n_{1101} becomes fixed and never changes afterwards. So after k_2, the number of quadruples $(1,1,1,1)$ and $(1,1,0,1)$ will remain constant.

Next, we obtain S^-_2 and S^+_2 by deleting all rows j from both S^-_1 and S^+_1 where the number of $(1,1,0,1)$s in the jth row of S^+_1 is more than that in S^-_1. Then by following the above process, one can show the existence of some time k_5 after which the number of each of the quadruples $(1,1,1,1)$, $(1,1,0,1)$, $(1,0,1,1)$, $(0,1,1,1)$, $(0,1,0,1)$ and $(1,1,1,0)$ becomes fixed. Correspondingly, we obtain S_5^- and S_5^-. Then one can check that the number of quadruples $(0,0,0,0)$ in every row of S_5^+ is **no more than** that in the same row in S_5^-. Hence, the function $-n_{0000}$ where n_{0000} is defined as the number of quadruples $(0,0,0,0)$ in the network never decreases after k_5. Therefore, there exists some finite time k_7, after which n_{0000} remains constant.

Following this approach and by using consecutively the functions n_{0011}, n_{0110}, n_{1100}, n_{1001}, $-n_{0010}$, $-n_{0100}$, $-n_{1000}$ and $-n_{0001}$, one can show the existence of some time k_{15} after which, the number of each corresponding quadruple becomes fixed. Moreover, we obtain S_{15}^- and S_{15}^+ as explained above, yet this time they both become an empty matrix. This implies that after k_{15}, no more switches of actions may take place in the network. Hence, the network will reach a stationary state at k_{15}, which must be an equilibrium due to the persistent activation assumption.

Case 2: $\frac{9}{5} \leq r \leq \frac{9}{4}$. In view of Lemma 7, we only need to prove the result for just one value of r in this range, say $r = 2$. Similar to the previous case and by using the exact same potential-like functions, one can show that the network reaches an equilibrium.

Case 3: $\frac{9}{4} \leq r \leq \frac{9}{3}$. This case can be proven exactly as case 1.

Case 4: $\frac{9}{3} \leq r \leq \frac{9}{2}$. In view of Lemma 7, we only need to prove the result for just one value of r in this range, say $r = 3.5$. Similar to the previous case and by using consecutively the functions n_{1110}, n_{1101}, n_{1011}, n_{0111}, n_{0101}, n_{0011}, n_{1100}, n_{1010}, n_{0110}, $-n_{1111}$, n_{1001}, n_{0010}, n_{0100} and $-n_{0000}$, one can show the existence of some time k_{14} when the network reaches an equilibrium.

Case 5: $\frac{9}{2} \leq r \leq 9$. This case can be proved similar to case 6.

Case 6: $r > 9$. In view of Lemma 7, we only need to prove the result for just one value of r in this range, say $r = 10$, which we do in two steps. First, we follow a similar approach to that in the previous cases. However, this time, instead of particular quadruples we investigate the number of particular quintuplets. We start with n_{11111} that is the number of 5 consecutive 1's in the network. Similar to above, in order to inspect the evolution of n_{11111}, we consider some time K and denote the active player at K by i. In order to count the difference of the quintuplets $(1,1,1,1,1)$ at and after time K, we need to investigate the actions of 4 players before and after player i in the ring, resulting in the action vector

$$q^i = (a_{i-4},\ \sigma_{i-3},\ \sigma_{i-2},\ \sigma_{i-1},\ \sigma_i,\ \sigma_{i+1},\ \sigma_{i+2},\ \sigma_{i+3},\ \sigma_{i+4}).$$

Then similar to S_0^- and S_0^+, we construct the $2^9 \times 9$ binary-matrices Q_0^- and Q_0^+. Now one can check that the number of 5 consecutive 1's in every row of Q_0^- is non-less than that in the same row in Q_0^+. Hence, there exists some time t_1, at which $-n_{11111}$ becomes fixed. By following this approach and using consecutively the functions $-n_{01111}$, $-n_{11110}$, $-n_{01110}$, n_{11101}, n_{10111}, $-n_{11100}$, $-n_{01100}$, $-n_{00111}$, n_{00101}, $-n_{00110}$, $-n_{11010}$, n_{11001}, n_{10101}, n_{10001}, $-n_{00001}$, n_{01010}, $-n_{10000}$, $-n_{11000}$, $-n_{00000}$, $-n_{10011}$, n_{00011}, n_{00010}, one can show the existence of some time t_{23}, after which the number of all corresponding quintuplets is fixed. Moreover, we obtain Q_{23}^- and Q_{23}^+ as follows:

$$Q_{23}^- = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

$$Q_{23}^+ = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Therefore, whenever a player i switches after time t_{23}, the actions of herself and her four left and right neighbors must be in the form of one of the rows in Q_{23}^- before the switch and becomes in the form of one of the rows in Q_{23}^+ after the switch.

As the second step of the proof, we show that only a finite number of these switches is possible. First we prove that only a finite number of switches can happen when the actions of the active player and her four left and right neighbors are in the form of the first row in Q_{23}^- before the switch. Assume on the contrary that there are infinite number of these types of switches. Then there exists some player $i \in \mathcal{V}$ such that q^i equals the first row of Q_{23}^- for an infinite time sequence $\mathcal{K} = \{k^1, k^2, \ldots\}$. On the other hand, due to the pairwise persistent assumption, there exists some time $k^j \in \mathcal{K}$ such that player $i-1$ is active at k^j+1. The actions of player $i-1$ and her three left and right neighbors at k^j+1 are

$$s^{i-1}(k^j+1) = (1,0,0,1,0,0,1).$$

Hence, player $i-1$ switches actions at k^j+2 resulting in

$$s^{i-1}(k^j+2) = (1,0,0,0,0,0,1).$$

Correspondingly, we have

$$q^{i-1}(t^j+1) = (*,1,0,0,1,0,0,1,*),$$
$$q^{i-1}(t^j+2) = (*,1,0,0,0,0,0,1,*)$$

where $*$ can be either 0 or 1. However, neither $q^{i-1}(t^j+1)$ equals any of the rows of Q_{23}^-, nor $q^{i-1}(t^j+2)$ equals any of the rows of Q_{23}^+. This is in contradiction with the fact that the actions of every player who switches and her four left and right neighbors are in the form of one of the rows in Q_{23}^- before the switch. So there exists some time $k_{24} \geq t_{23}$, after which no player i whose corresponding q^i is in the form of the first row of Q_{23}^{-} takes place. Similarly, the same can be shown for the second row. Therefore, after some finite time, no player will switch actions, and hence, the network will be fixed at some state. On the other hand, due to the persistent activation assumption, that state must be an equilibrium since every player gets the chance to update her action infinitely many times. This completes the proof. □

The proof of Theorem 6 is algorithmic and can be generalized to symmetric spatial structures, e.g., regular graphs, but is less useful in the convergence analysis for games on irregular networks. A key feature of the proof is the exploitation of the fact that the behavior of the homogeneous linear PGG under the 'imitate the best' unconditional imitation dynamics is equivalent for different values of r in certain ranges. This enabled us to significantly decrease the computations necessary to show finite time convergence for every $r \geq 0$.

4.4 Cooperation, convergence and imitation

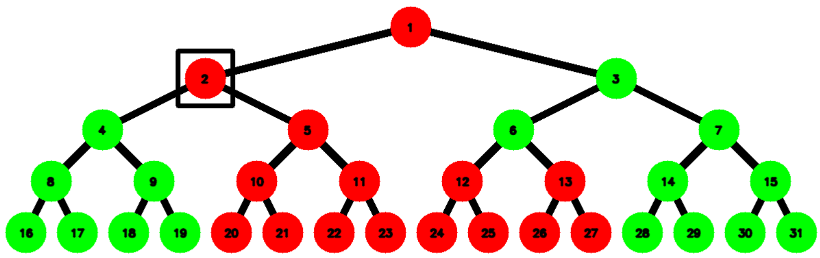
We have seen that for the spatial PGG, rational imitation dynamics converge to an imitation equilibrium regardless of the spatial structure and heterogeneity in the payoff parameters. When the rationality of having incentives to deviate is broken, as in unconditional imitation dynamics, convergence of the decision process can only be guaranteed for specific spatial structures. Thus, even under payoff monotonicity assumptions, in general, one cannot expect that players reach a decision they are satisfied with. However, an important aspect of imitation dynamics even in the absence of convergence and mechanisms such as punishment, reward and voluntary participation [102], unconditional imitations can allow for *cooperative* actions to exist in the imitation equilibrium of a social dilemma game (Lemma 6, Fig. 4.2 and [96]). Hence, under unconditional imitation dynamics the maintenance of a publicly available resource (i.e. the public good in a PGG), can be assured with relative ease.

For the homogeneous linear PGG, the best response depends solely on the public goods multiplier r, the degree of a player and the degree of its neighbors. In this case, cooperation is promoted (resp. impeded) for players with a degree that is higher (resp. lower) than their neighbors' degrees. For regular networks, in which all players have the same degree, say $d \geq 1$, cooperation can only exist at the Nash equilibrium if $r > d + 1$. This simple condition, however, implies that cooperation is a *dominant pure strategy* in each local interaction, and hence, at least in the spatial PGG, *network reciprocity* [8] is ineffective under such rational and innovative decision processes. Take, for example, the simple 2-regular tree depicted in Fig. 4.2. For a public goods multiplier $r = \frac{5}{2}$, the unique Nash equilibrium is full defection. Under 'imitate the best' unconditional imitation dynamics, the action profile has persistent oscillations with a high number of cooperators in the oscillations action profiles. An example of such oscillations is shown in Fig. 4.2: starting from the action profile in (a) either the players labeled as 2 and 5, or 3 and 6 persistently imitate each other's actions. Thus, even though players cooperate, they are not necessarily *satisfied* with their decision and keep changing their actions. This behavior can also occur in matrix games on networks [97].

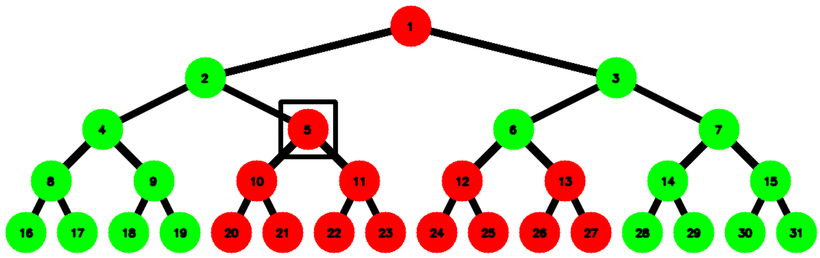
Interestingly, for a *rational* imitation process with $h_i = 1$ for all i, the action profile in Fig. 4.2(a) is, in fact, an imitation equilibrium that coincides with a *generalized Nash equilibrium* [88] in which each players selects a *relative* best response. Thus it is not the rationality of selfish players that is necessarily detrimental to the cooperation levels in spatial PGG, but rather their ability to *innovate* rationally. This example shows that *rational* imitation can facilitate cooperative decisions without compromising the finite time convergence of the decision process, and hence, rational imitations of selfish players can facilitate cooperative decisions without requiring any punishment of defectors [12], reputation considerations [12] or the possibility of players to waive participation in the game and opt for a more self-sustaining action [102]. Aside from this specific example, extensive simulations on arbitrarily connected networks support this finding. Thus, it is not always the irrationality of imitations that allows cooperation to exist, but rather the combination of imitations and the ability of players to predict, via the (lack of) success of others, when their own defective motives will become detrimental to their *own* success. In the folowing we will show via simulations, how rational imitation can even result in even higher cooperation levels than unconditional imitation.

4.4.1 Simulations on a bipartite graphs

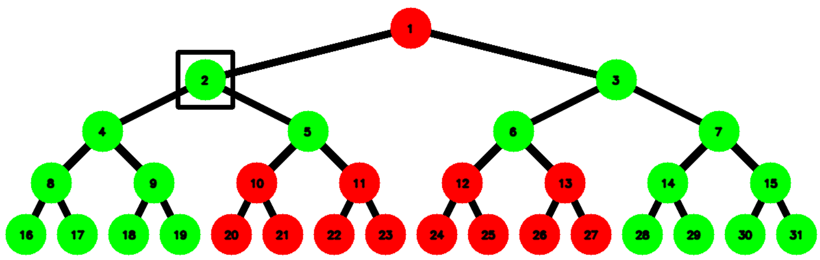
Let us set $h_i = 1$ for all players and simulate imitate the best unconditional imitation and rational imitation dynamics for a homogeneous linear PGG with $c = 1$ and a variable public goods multiplier r. By varying r we are interested in how the public goods parameter, acting like a benefit-to-cost ratio, influences the number of cooperators in the imitation equilibrium. Group structures are determined by the neighborhoods of a bipartite graph with two independent and disjoint sets each containing 14 players so that the total number of players is 28. We vary the number of connections of a player by varying the probability of a player to be connected to another player in the other disjoint independent player set. Imitate the best unconditional imitation dynamics need not converge to an imitation equilibrium. In this case, we let the action profile evolve for 10^4 time steps and determine the fraction of cooperators by the number of cooperators at the final time step $t = 10^4$.



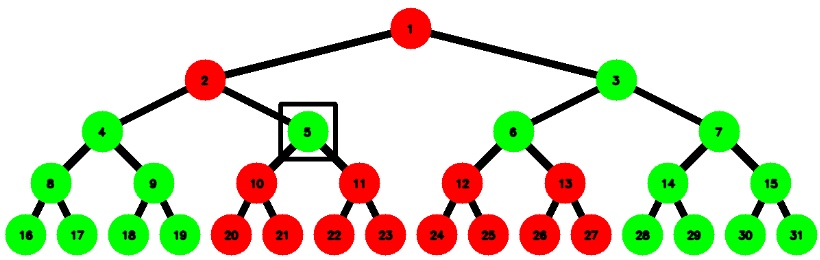
(a) Player 2 imitates its best performing neighbor 4.



(b) Player 5 imitates its best performing neighbor 2.



(c) Player 2 imitates its best performing neighbor 1.



(d) Player 5 imitates its best performing neighbor 2, and the action profile returns to (a). A similar imitation cycle exists for players 3 and 6 when in the action profile (c) player 3 imitates its best performing neighbor 1.

Figure 4.2: Persistent imitation oscillations in a spatial PGG on a 2-regular tree under asynchronous unconditional imitation dynamics with parameters $c = 1$ and $r = 2.5$. Green vertices represents a cooperators, red vertices represent defectors. The square indicates the unique deviator.

To get a feeling for how rational imitation can facilitate cooperation we initialize one of the independent sets as cooperators and the other as defectors. The simulation results are shown in Fig. 4.3. The plots are obtained by averaging over 100 random activation sequences. In the top sub-figures of Fig. 4.3 one can see that if the average degree of the players in the network is relatively high e.g. 10 and 7, rational imitation can facilitate half the network to cooperate for a large range of public goods multiplier values, whereas unconditional imitation dynamics result in significantly lower proportion of cooperators. When the average degree of the players is reduced, this promoting effect of rational imitation is less noticeable, and the proportion of

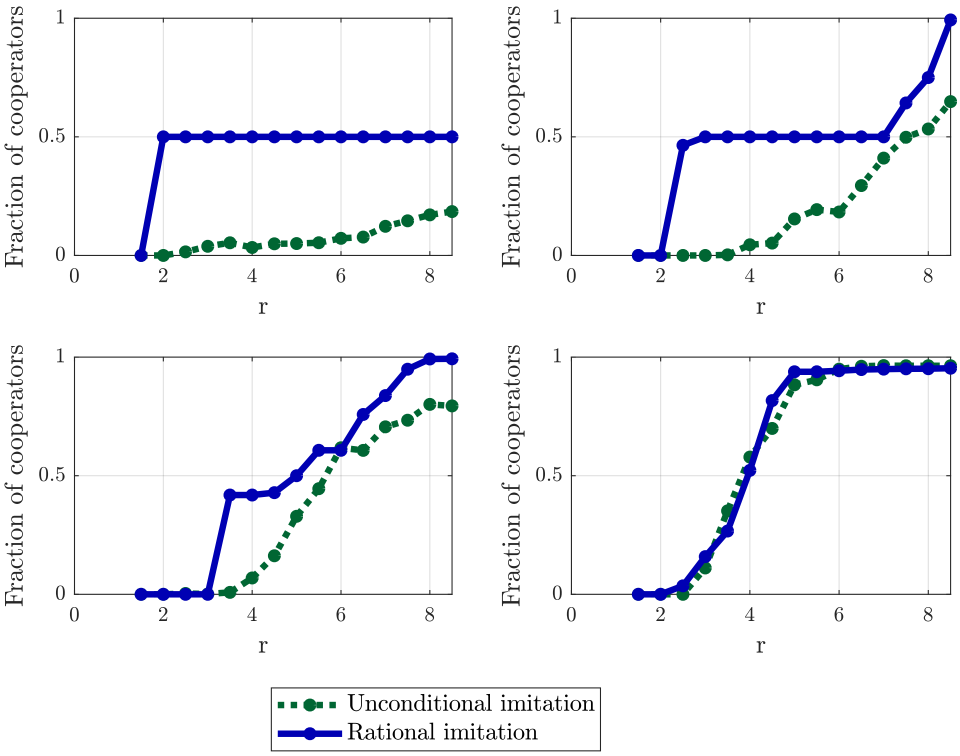


Figure 4.3: Simulations for a homogeneous linear PGG on a bipartite network with a clustered initial condition. The four subgraphs correspond to simulation results obtained for different levels of connectivity between the two independent and disjoint sets of vertices with an average degree of: 10, 7, 2.5 and 3 (clockwise starting from top left).

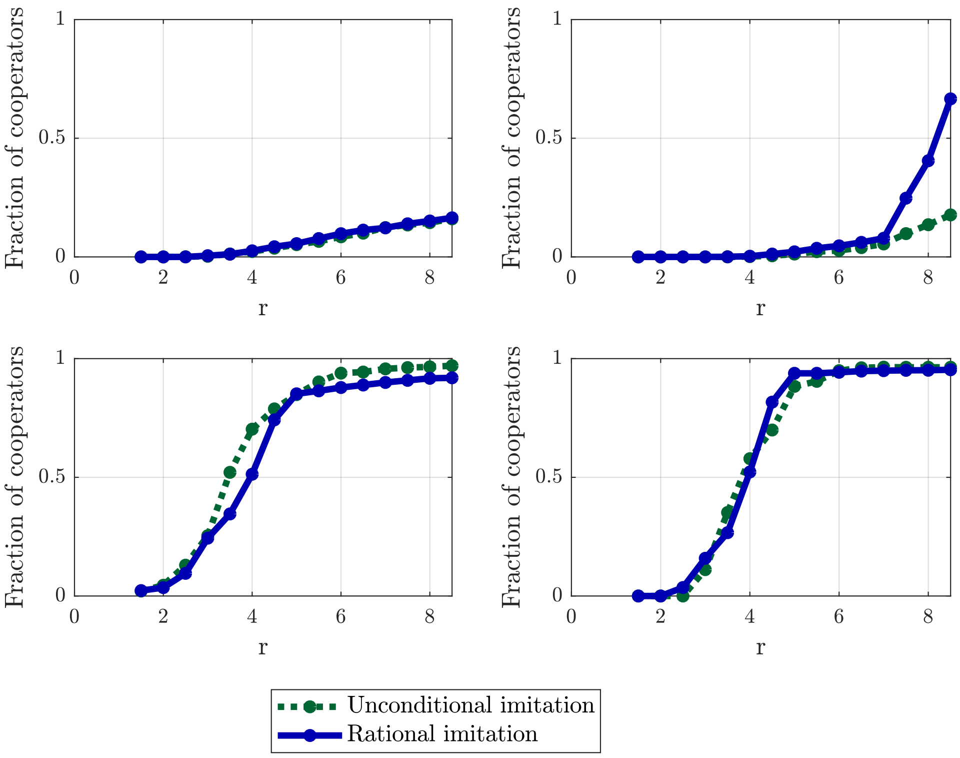


Figure 4.4: Simulations for a homogeneous linear PGG on a bipartite network with initial actions that are determined by a discrete uniform distribution. The average degrees of the players in the networks correspond with those in Fig. 4.3.

cooperators of rational and unconditional imitations are similar (bottom sub-figures of Fig. 4.3).

When the initial action profile is random but equal for both dynamics, the number of cooperators are similar for both unconditional and rational imitation, see Fig. 4.4. The proportion of cooperators in these simulations are obtained by averaging over 500 random initial conditions and activation sequences. In this case, the rational imitation dynamics promote cooperation more than unconditional imitation dynamics for larger values of public goods multipliers and an average degree of 7. These simulations illustrate that rational imitation dynamics of spatial PGG in which the players have relatively high degrees, the rational imitation dynamics can facilitate initially clustered cooperators better than unconditional imitation. It is in these cases that network reciprocity and rational imitations can optimally maintain publicly available goods.

4.5 Final Remarks

We have shown that rational imitation dynamics in a general class of asynchronous spatial PGG converge to an imitation equilibrium in a finite time. By means of a counter-example we have shown that this general case of convergence is not guaranteed when imitation is *unconditional*. For regular spatial structures and production functions, however, we have proven convergence either directly from the payoff functions or by using an algorithmic proof technique that takes advantage of the regularity of the network. We have shown that in the case of rational imitation, convergence is also guaranteed when the group structures are determined by a bipartite graph. Such a representation of a social dilemma can, for instance, be used when the group structures are obtained from data that does not contain information about the entire social network. Next to convergence, we have provided evidence that in contrast to best response dynamics, rational imitation can effectively facilitate the evolution of cooperation via network reciprocity. Our results indicate that through the combination of rationality and imitation, beneficial dynamic features can arise that are able to sustain the availability of a publicly available good, providing new insights in the design of solutions to the tragedy of the commons.

CHAPTER 5

Strategic differentiation in finite network games

Masked, I advance.

René Descartes

SIMPLE decisions or actions taken by interacting individuals can lead to surprisingly complex and unpredictable population-level outcomes. In particular, when individual decisions or actions are based on personal interest, the long run collective behavior, characterized by these selfish decisions, can be detrimental for the population as a whole. Mathematical models of such systems require trade-offs between the complexity of micro-dynamics and the accuracy with which the model can describe a macro-behavior.

Evolutionary game theory has proven to be a valuable tool in providing mathematical models for such complex dynamical systems. In its original application to evolving biological populations, rational reasoning or decision-making is not needed; rather, competing strategies propagate through a population via natural selection. Economists later adapted this evolutionary game model for the mathematical modeling of individual decision-makers with *bounded rationality* [41]. To reach decisions they are satisfied with, players may thus rely on simple rules. Relative best responses and rational imitations, that we have studied in the previous chapters, are examples of such myopic decision rules. As we have seen, these decision-making models can be extended to include *groupwise interactions* [96], that are important to study because many biological and social interactions involve more than two individuals whose collective decisions can have a variety of behaviors even in well-mixed populations [103].

A common assumption in the existing models for finite network games is that players do not distinguish between their opponents. In some sense the opponents are *anonymous* and hence, there is no difference in the actions employed against each of them. Indeed, in the previous chapters, it was assumed that all players employ the same action against their opponents. However, to create a competitive advantage, in real life competitive settings, it is often crucial to identify the rivals [104]. Avoiding 'blindspots' in a *competitive decision-making process*, i.e. those decisions that require taking into account the decisions of competitors, is a major topic in the strategic decision-making literature [105]. In such competitive environments, decision-makers *are* likely to distinguish their opponents, and consequently, they may employ different actions against them.

In this chapter, the mechanism of *strategic differentiation* is introduced through which a subset of players in the network, called *differentiators*, can employ different actions against different opponents. We will connect strategic differentiation to the theory of potential games and their generalizations and show that for the class of weighted potential games the effect of strategic differentiation on any network topology can be studied analytically using the potential function of the original game. In the following, we will distinguish groupwise and pairwise games on networks. To make the difference clear we introduce some additional notation. let $\pi_{ij}(\sigma_i, \sigma_j)$ denote the payoff that player i obtains from action $\sigma_i \in \mathcal{A}_i$ in the pairwise interaction against opponent $j \in \mathcal{N}_i$ with action $\sigma_j \in \mathcal{A}_j$. The total payoff that player i obtains in a single round of play with pairwise interactions is

$$\pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) = \sum_{j \in \mathcal{N}_i} w_{ij} \pi_{ij}(\sigma_i, \sigma_j), \tag{5.1}$$

with $w_{ij} \in \mathbb{R}$ denoting the weight associated to the local interaction between i and j. We refer to a non-cooperative game with a payoff function of the form Eq. (5.1) as a *pairwise network game*. An example of such a game is the famous spatial prisoner's dilemma game. As we have seen in the previous chapter, players may also interact in groups with a size greater than two, and thus the local interactions form a *multiplayer* game. The spatial public goods game studied in the previous chapter is an example of such a *groupwise network game*. In general, the payoffs of groupwise network games cannot be represented by the corresponding sum of pairwise interactions, however the total payoff that player i obtains in a single round of play is again a weighted sum of the local payoffs,

$$\pi_i(\sigma_i, \boldsymbol{\sigma}_{-i}) = \sum_{j \in \bar{\mathcal{N}}_i} w_j \pi_{ij}(\sigma_i, \boldsymbol{\sigma}_{-i}), \tag{5.2}$$

with $w_j \in \mathbb{R}$ denoting the weight associated to the multiplayer game with exactly those players in $\bar{\mathcal{N}}_j$. Note that in Eq. (5.2) the single round local payoffs depend on $|\mathcal{N}_j + 1| \geq 2$ actions and the network structure imposes an interdependence in the payoffs of players that are connected via an undirected path with length two, sometimes referred to as the 2-hop neighbors.

5.1 Strategic Differentiation

In a network game with strategic differentiation, a *differentiator* can employ a separate pure action for each neighbor; see figure 5.1 for an example of a pairwise game with a single differentiator on a ring network. Let $\mathcal{D}$ be a non-empty subset of $\mathcal{V}$ denoting the set of differentiators in the network, and let $\mathcal{F} := \mathcal{V} \setminus \mathcal{D}$ denote the set of non-differentiators. In a groupwise network game, the actions of a player $i \in \mathcal{D}$ is a *vector* $s_i \in \mathcal{S}_i := \mathcal{A}_i^{|\mathcal{N}_i|+1}$: a separate action can be chosen in each of the multiplayer game with the closed neighborhoods that the player belongs to. The action space of differentiators is indicated by $\boldsymbol{\mathcal{S}_{\mathcal{D}}} := \prod_{i \in \mathcal{D}} \mathcal{S}_i$. When the game interactions are pairwise, the dimension of the action vector of player i is reduced by one because in this case players only interact with their $|\mathcal{N}_i|$ neighbors. For some $j \in \mathcal{N}_i$, we indicate by $s_{ij} \in s_i$ the action that player $i \in \mathcal{D}$ employs in the local pairwise (resp. groupwise) game played against j (resp. $\mathcal{N}_j$). Note that for all $i \in \mathcal{D}$ and $j \in \bar{\mathcal{N}}_i$ we assume that $s_{ij} \in \mathcal{A}_i$, i.e. each action employed by a differentiator is in its own action set. The action space of players who do not differentiate is indicated by $\boldsymbol{\mathcal{A}_{\mathcal{F}}} := \prod_{j \in \mathcal{F}} \mathcal{A}_j$. Without loss of generality, label differentiators by $\mathcal{D} = \{1, \ldots |\mathcal{D}|\}$ and the non-differentiators by $\mathcal{F} = \{|\mathcal{D} + 1|, \ldots, n\}$. Then the action space of the

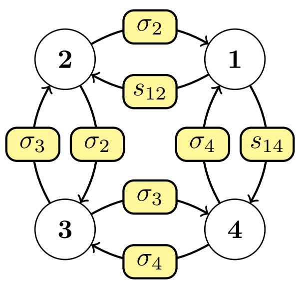


Figure 5.1: Graphical interpretation of a pairwise non-cooperative game on a network with strategic differentiation. The label of outgoing edges indicates the action played in the local pairwise interaction. In this example $\mathcal{D} = \{1\}$ and $\mathcal{F} = \{2, 3, 4\}$.

networked game with strategic differentiation is given by

$$\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{S}}_{\boldsymbol{\mathcal{D}}} \times \boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{F}}},$$

An action profile in the action space of a game with strategic differentiation is indicated by $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$. As before, $\boldsymbol{s}_{-i} \in \boldsymbol{\mathcal{S}}_{-i}$ indicates the action profile of all players except player i. In a game with strategic differentiation we denote the local payoff function for the interaction between i and (the neighbors of) $j \in \bar{N}_i$ by $u_{ij} : \boldsymbol{\mathcal{S}} \to \mathbb{R}$. Similarly, $u : \boldsymbol{\mathcal{S}} \to \mathbb{R}^n$ denotes the combined payoff vector of the game with strategic differentiation. For pairwise interactions the payoffs of a differentiator $i \in \mathcal{D}$ is

$$u_i(s_i, \boldsymbol{s}_{-i}) = \sum_{j \in \mathcal{N}_i \cap \mathcal{D}} w_{ij} u_{ij}(s_{ij}, s_{ji}) + \sum_{h \in \mathcal{N}_i \cap \mathcal{F}} w_{ih} u_{ih}(s_{ih}, \sigma_h). \tag{5.3}$$

The payoff of a non-differentiator $k \in \mathcal{F}$ is

$$u_k(s_k, \boldsymbol{s}_{-k}) = \sum_{m \in \mathcal{N}_k \cap \mathcal{D}} w_{km} u_{km}(\sigma_k, s_{mk}) + \sum_{v \in \mathcal{N}_k \cap \mathcal{F}} w_{kv} u_{kv}(\sigma_k, \sigma_v) \tag{5.4}$$

For games with strategic differentiation and *groupwise* interactions, as in Eq. (5.2), the local payoff function u_{ij} will depend on more than two actions. For $i \in \mathcal{D}$,

$$u_i(s_i, \boldsymbol{s}_{-i}) = \sum_{k \in \bar{\mathcal{N}}_i} w_k u_{ik}(s_{ik}, \boldsymbol{s}_{-i}). \tag{5.5}$$

And the payoff of a non-differentiator $j \in \mathcal{F}$ is,

$$u_j(s_j, \boldsymbol{s}_{-j}) = \sum_{k \in \bar{\mathcal{N}}_j} w_k u_{jk}(\sigma_j, \boldsymbol{s}_{-j}) \tag{5.6}$$

We are now ready to formally define a network game with strategic differentiation.

Definition 17 (Strategically differentiated game)**.** *A network game with strategic differentiation is defined by the triplet* $\Xi := (\mathbb{G}, \boldsymbol{\mathcal{S}}, u)$*. If* $\pi_{ij} = u_{ij}$ *for all* $(i, j) \in \mathcal{E}$*, then* Ξ *is said to be the strategically differentiated version of* $\Gamma = (\mathbb{G}, \boldsymbol{\mathcal{A}}, \pi)$*.*

Example 6 (Strategic differentiation in a groupwise game)**.** *As an example of a groupwise game with strategic differentiation consider a linear public goods game in which players need to determine whether or not to contribute to a public good that their opponents can profit from. This decision is modeled by the pure action set* $\mathcal{A}_i = \{0, 1\}$ *for all* $i \in \mathcal{V}$*. A differentiator* $i \in \mathcal{D}$ *may choose to contribute to one good but withhold from contributing to another. Hence,* $s_i \in \{0, 1\}^{|\mathcal{N}_i|+1}$*. For some* $j \in \bar{\mathcal{N}}_i$*, when* $s_{ij} = 1$ *(resp.* $s_{ij} = 0$ *) player* i *is cooperating (resp. defecting) in the local game against the group of opponent players* $\bar{\mathcal{N}}_j$*. Let* $c_{ij} \in \mathbb{R}_{>0}$ *denote the contribution of a cooperating* *player i in the game against the neighbors of $\bar{\mathcal{N}}_j \subseteq \mathcal{D}$. In a public goods game, the contributions get multiplied by an enhancement factor $r \in [1, n]$, which can be seen as a benefit-to-cost ratio of the local interaction. The payoff that player i in this local game is,*

$$u_{ij}(s_{ij}, \boldsymbol{s}_{-i}) = \frac{r(\sum_{k \in \bar{\mathcal{N}}_j} s_{kj} c_{kj} + s_{jj} c_{jj})}{d_j + 1} - c_{ij} s_{ij}. \tag{5.7}$$

5.2 Rationality in games with strategic differentiation

A best response of a differentiator is a vector of actions for which each element is a best response in the corresponding local game.

Definition 18 (Differentiated Best Response)**.** *For a differentiator $i \in \mathcal{D}$, the action $s_i \in \mathcal{S}_i$ is a strategically differentiated pure best response against $\boldsymbol{s}_{-i}$ if for all $s_{ik} \in s_i$*

$$s_{ik} \in \arg\max_{x \in \mathcal{A}_i} u_{ik}(x, \boldsymbol{s}_{-i}) \tag{5.8}$$

Based on the definition of a differentiated best response, a Nash equilibrium in a strategically differentiated game is naturally defined as follows.

Definition 19 (Differentiated Pure Nash equilibrium)**.** *An action profile $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$ is a differentiated pure Nash equilibrium of Ξ if for all $j \in \mathcal{F}$, $\sigma_j \in \boldsymbol{s}$ is a best response and for all $i \in \mathcal{D}$, $s_i \in \boldsymbol{s}$ is a strategically differentiated pure best response.*

When $\mathcal{D} = \emptyset$ the original definition of a pure Nash equilibrium is recovered. Best responses of differentiators are thus defined as vectors of actions for which each element is *locally* optimal. Herein lies the main distinguishing feature of best replies in games without strategic differentiation: a best reply x_i^* over the aggregated payoff $\pi_i(x_i, x_{-i})$ might not optimize the payoffs of each separate local game with payoff π_{ij}. Hence, a strategically differentiated Nash equilibrium can contain actions that are not present in the Nash equilibrium of the game without strategic differentiation.

Let us now consider myopic best response dynamics in games with strategic differentiation: the action s_{ij} is chosen such that it maximizes u_{ij}, ceteris paribus.

Definition 20 (Differentiated myopic best response dynamics)**.** *If a player i updates using differentiated best responses the resulting (myopic) best response dynamics are*

$$s_{ij}(t+1) \in \arg\max_{y \in \mathcal{A}_i} u_{ik}(y, \boldsymbol{s}_{-i}(t)). \tag{5.9}$$

When all differentiators update their actions according to the differentiated best response dynamics Eq. (5.9) and the others according to myopic best response dynamics, we indicate the "evolutionary" game with strategic differentiation by (Ξ, β).

Remark 6. *Differentiated myopic best response dynamics are an unconstrained version of myopic best responses in the sense that the local actions are optimized over the local payoffs without requiring that the employed actions are equal. It follows that when $\mathcal{D} = \mathcal{V}$, for innovative update dynamics like myopic best response the effect of the network structure on the equilibria of the network game is lost. That is, the equilibrium action profiles of the networked game would, in this case, correspond to a collection of separate Nash equilibria of the local games played on the network. When $\mathcal{D} \subset \mathcal{V}$ the network structure remains important to the myopic best response dynamics. Moreover, the differentiators may obtain an advantage over their opponents that are not able to differentiate their actions because for each $\sigma_i \in \mathcal{A}_i$, $\boldsymbol{s}_{-i} \in \boldsymbol{\mathcal{S}}_{-i}$*

$$\exists s_i \in \mathcal{S}_i \ : \ u_i(s_i, \boldsymbol{s}_{-i}) \geq u_i(\sigma_i, \boldsymbol{s}_{-i}).$$

Hence, in terms of payoffs, players that differentiate their actions rationally are always at least as successful as they would have been not differentiating. The benefit that differentiators can get over non-differentiators implies that especially for evolutionary update dynamics in which the most successful players are imitated, the existence of differentiators can have a significant impact on the evolution of the actions in the network.

5.3 Potential functions for network games with strategic differentiation

In this section, we describe conditions on the local interactions of network games that ensure that their strategically differentiated versions have pure Nash equilibria and convergence of differentiated myopic best response dynamics is guaranteed. For this, we apply the theory of potential games to strategically differentiated games. Consider the following definition derived from ordinal potential games [42].

Definition 21 (Differentiated ordinal potential game)**.** *Ξ is a strategically differentiated ordinal potential game if there exists an ordinal potential function $P : \boldsymbol{\mathcal{S}} \to \mathbb{R}$ such that for all $\sigma_j, \sigma_j' \in \mathcal{A}_j$, $s_i, s_i' \in \mathcal{S}_i$, $\boldsymbol{s}_{-i} \in \boldsymbol{\mathcal{S}}_{-i}$ and $\boldsymbol{s}_{-j} \in \boldsymbol{\mathcal{S}}_{-j}$ the following holds:*

$$u_i(s_i, \boldsymbol{s}_{-i}) - u_i(s_i', \boldsymbol{s}_{-i}) > 0 \Leftrightarrow P(s_i, \boldsymbol{s}_{-i}) - P(s_i', \boldsymbol{s}_{-i}) > 0, \qquad \forall i \in \mathcal{D},$$
$$u_j(\sigma_j, \boldsymbol{s}_{-j}) - u_j(\sigma_j', \boldsymbol{s}_{-j}) > 0 \Leftrightarrow P(\sigma_j, \boldsymbol{s}_{-j}) - P(\sigma_j', \boldsymbol{s}_{-j}) > 0, \qquad \forall j \in \mathcal{F}.$$

Note that if $\mathcal{D} = \emptyset$, the original definition of an ordinal potential game introduced in [42] is recovered from Definition 21. It is well known that every *finite* ordinal potential game has a pure Nash equilibrium. This property is generalized to strategically differentiated games in the following lemma.

Lemma 8. *Every finite differentiated ordinal potential game possesses a differentiated pure Nash equilibrium.*

Proof. First assume $\boldsymbol{s}^* \in \boldsymbol{\mathcal{S}}$ is a differentiated Nash equilibrium for Ξ. Then, for all $i \in \mathcal{D}$, $s_i^* \in \boldsymbol{s}^*$ is such that

$$u_i(s_i^*, \boldsymbol{s}_{-i}^*) - u_i(s_i, \boldsymbol{s}_{-i}^*) \geq 0, \ \forall s_i \in \mathcal{S}_i.$$

For all $j \in \mathcal{F}$ the actions of the non-differentiators $\sigma_j^* \in \boldsymbol{s}^*$ are such that

$$u_i(\sigma_j^*, \boldsymbol{s}_{-i}^*) - u_i(\sigma_j, \boldsymbol{s}_{-i}^*) \geq 0, \ \forall \sigma_j \in \mathcal{A}_j.$$

By the definition 21 of a differentiated ordinal potential game it follows that

$$\begin{aligned} &P(s_i^*, \boldsymbol{s}_{-i}^*) - P(s_i, \boldsymbol{s}_{-i}^*) \geq 0, \forall i \in \mathcal{D} \text{ and } \forall s_i \in \mathcal{S}_i, \\ &P(\sigma_j^*, \boldsymbol{s}_{-j}^*) - P(\sigma_j, \boldsymbol{s}_{-i}^*) \geq 0, \forall j \in \mathcal{F} \text{ and } \forall \sigma_j \in \mathcal{A}_j. \end{aligned}$$

Hence $\boldsymbol{s}^*$ is also a maximum point in P. Similarly one can show that each maximum point in P is a differentiated Nash equilibrium of Ξ. Since for finite games $\boldsymbol{\mathcal{S}}$ is bounded, a maximum of P always exists. This completes the proof. □

One can show that if Ξ is a differentiated ordinal potential game, then (Ξ, β) will always terminate in a differentiated Nash equilibrium. Instead, we now focus on finding conditions on the local interactions in groupwise games on networks that ensure the convergence properties of Γ under best responses are preserved in its strategically differentiated version Ξ. This is especially useful when one already has a potential function for the original game on a network and is interested in comparing the behavior of the game with strategic differentiation. Before doing so, consider the following definition.

Definition 22 (Differentiated weighted potential games)**.** *Ξ is a strategically differentiated weighted potential game if there exists a potential function $\bar{P} : \boldsymbol{\mathcal{S}} \to \mathbb{R}$ and weights $\alpha_i, \alpha_j \in \mathbb{R}_{>0}$, such that for all $\sigma_j, \sigma_j' \in \mathcal{A}_j$, $s_i, s_i' \in \mathcal{S}_i$, $\boldsymbol{s}_{-i} \in \boldsymbol{\mathcal{S}}_{-i}$ and $\boldsymbol{s}_{-j} \in \boldsymbol{\mathcal{S}}_{-j}$ the following holds:*

$$\begin{aligned} &u_i(s_i, \boldsymbol{s}_{-i}) - u_i(s_i', \boldsymbol{s}_{-i}) = \alpha_i \left[\bar{P}(s_i, \boldsymbol{s}_{-i}) - \bar{P}(s_i', \boldsymbol{s}_{-i})\right)], && \forall i \in \mathcal{D} \\ &u_j(\sigma_j, \boldsymbol{s}_{-j}) - u_j(\sigma_j', \boldsymbol{s}_{-j}) = \alpha_j \left[\bar{P}(\sigma_j, \boldsymbol{s}_{-j}) - \bar{P}(\sigma_j', \boldsymbol{s}_{-j})\right], && \forall j \in \mathcal{F}. \end{aligned}$$

Note that if $\mathcal{D} = \emptyset$, the original definition of a weighted potential game introduced in [42] is recovered.

The following result relates the fundamental properties of weighted potential games to their strategically differentiated version.

Theorem 7. *In Γ, if for all players $i \in \mathcal{V}$ there exists for each local payoff function $\pi_{ij} : \boldsymbol{\mathcal{A}} \to \mathbb{R}$, $j \in \mathcal{N}_i$, a weighted potential function $\rho_j : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ with a common weight $\alpha_i \in \mathbb{R}_{>0}$ for player i, then (Ξ, β) converges to a differentiated pure Nash equilibrium.*

Proof. For all $j \in \mathcal{F}$, let $\bar{s}_j := (\sigma_j, \dots, \sigma_j) \in \mathcal{A}_j^{|\mathcal{N}_j+1|}$ such that each element in $\bar{s}_j$ is equal to $\sigma_j \in \mathcal{A}_j$. For all differentiators $i \in \mathcal{D}$ let $\bar{s}_i := s_i$. Then, for all $j \in \mathcal{F}$ the payoff in the strategically differentiated game can be written as

$$u_j(\sigma_j, \boldsymbol{s}_{-j}) = \sum_{k \in \bar{\mathcal{N}}_j} w_k u_{jk}(\bar{s}_{jk}, \bar{s}_{-jk}),$$

where $\bar{s}_{-jk}$ denotes the set of actions in the game centered around k, chosen by the neighbors of k different from j, i.e., $\bar{s}_{-jk} := \{\bar{s}_{vk} \in \mathcal{A}_v : v \neq j \ \wedge \ v \in \mathcal{N}_k\}$. For differentiators $i \in \mathcal{D}$, the payoff in the strategically differentiated game is

$$u_i(s_i, \boldsymbol{s}_{-i}) = \sum_{k \in \bar{\mathcal{N}}_i} w_k u_{ij}(\bar{s}_{ik}, \bar{s}_{-ik}).$$

By assumption, in any local game with the neighbors of some $k \in \mathcal{V}$, for all $i \in \bar{\mathcal{N}}_k$ there exists a function $\rho_k : \boldsymbol{\mathcal{A}} \to \mathbb{R}$ and weights $\alpha_i \in \mathbb{R}_+$, such that for every $\sigma_i, \sigma_i' \in \mathcal{A}_i$ and every $\boldsymbol{\sigma}_{-i} \in \boldsymbol{\mathcal{A}}_{-i}$ the following equality holds,

$$\pi_{ik}(\sigma_i, \boldsymbol{\sigma}_{-i}) - \pi_{ik}(\sigma_i', \boldsymbol{\sigma}_{-i}) = \alpha_i \left[\rho_k(\sigma_i, \boldsymbol{\sigma}_{-i}) - \rho_k(\sigma_i', \boldsymbol{\sigma}_{-i})\right].$$

For all the non-differentiators $j \in \bar{\mathcal{N}}_k \cap \mathcal{F}$, it follows that for every $\sigma_j, \sigma_j' \in \mathcal{A}_j$ and $\bar{s}_{-jk} \in \prod_{v \in \bar{\mathcal{N}}_k \setminus \{j\}} \mathcal{A}_k$ it holds that

$$\begin{aligned} u_{jk}(x_j, \bar{s}_{-jk}) - u_{jk}(x_j', \bar{s}_{-jk}) &= \alpha_j \left[\rho_k(x_j, \bar{s}_{-jk}) - \rho_k(x_j', \bar{s}_{-jk})\right] \\ &= \alpha_j \left[\rho_k(\bar{s}_{jk}, \bar{s}_{-jk}) - \rho_k(\bar{s}_{jk}', \bar{s}_{-jk})\right]. \end{aligned}$$

Similarly for the differentiators $i \in \mathcal{D} \cap \bar{\mathcal{N}}_k$, since $\bar{s}_{ik} \in \mathcal{A}_i$ and $\bar{s}_{-ik} \in \mathcal{A}_{-ik}$, from the existence of the local weighted potential function ρ_k it follows that for all $\bar{s}_{ik}, \bar{s}_{ik}' \in \mathcal{A}_i$, $\bar{s}_{-ik} \in \mathcal{A}_{-ik}$

$$u_{ik}(\bar{s}_{ik}, \bar{s}_{-ik}) - u_{ik}(\bar{s}_{ik}', \bar{s}_{-ik}) = \alpha_i \left[\rho_k(\bar{s}_{ik}, \bar{s}_{-ik}) - \rho_k(\bar{s}_{ik}', \bar{s}_{-ik})\right].$$

Let $\bar{s}_{\{\bar{\mathcal{N}}_v\}} := \{\bar{s}_{kv} \in \mathcal{A}_k : k \in \bar{\mathcal{N}}_v\}$ denote the set of actions employed by the players in the local interaction with the closed neighborhood of l. The difference in the payoffs of a unique deviator $i \in \mathcal{V}$ switching from action $\bar{s}_i$ to $\bar{s}_i'$ is given by

$$\begin{aligned} u_i(\bar{s}_i, \bar{s}_{-i}) - u_i(\bar{s}_i', \bar{s}_{-i}) &= \sum_{j \in \bar{\mathcal{N}}_i} w_j \left[u_{ij}(\bar{s}_{ij}, \bar{s}_{-ij}) - u_{ij}(\bar{s}_{ij}', \bar{s}_{-ij}) \right] \\ &= \sum_{j \in \mathcal{N}_i} w_j \left[\alpha_i \left(\rho_j(\bar{s}_{ij}, \bar{s}_{-ij}) - \rho_j(\bar{s}_{ij}', \bar{s}_{-ij}) \right) \right] \\ &= \alpha_i \sum_{j=1}^{n} w_j \left(\rho_j(\bar{s}_{\{\mathcal{N}_j\}}) - \rho_j(\bar{s}_{\{\mathcal{N}_j\}}') \right). \end{aligned} \tag{5.10}$$

The last equality in Eq. (5.10) follows from the fact that when the unique deviator is not a member of some closed neighborhood $\mathcal{N}_h$, then $\rho_h(s_{\{\mathcal{N}_h\}}) - \rho_h(s_{\{\mathcal{N}_h\}}') = 0$. This implies that

$$\bar{P} = \sum_{j \in \mathcal{V}} w_j \rho_j(\bar{s}_{\mathcal{N}_j}),$$

with weights α_i is a weighted potential function and thus Ξ is a strategically differentiated weighted potential game. The convergence of the differentiated myopic best response then follows from the argument used in traditional potential games. Clearly, since $\bar{S}$ is finite, $\bar{P}$ is bounded. Moreover it is *increasing* along the trajectory generated by asynchronous myopic best responses of non-differentiators and asynchronous differentiated myopic best responses of differentiators. This implies convergence of the differentiated myopic best response action update dynamics to a differentiated pure Nash equilibrium. □

The proof of Theorem 7 can be easily adjusted to show that the same statement holds for strategically differentiated pairwise games on networks with $w_{ij} = w_{ji}$ for all $(i, j) \in \mathcal{E}$. The following corollary of Theorem 7 applies to the class of *exact* potential games in which $\alpha_i = 1$ for all $i \in \mathcal{V}$.

Corollary 4. *If the local groupwise interactions of* Γ *are potential games, then* (Ξ, β) *converges to a strategically differentiated Nash equilibrium.*

Remark 7. *Theorem 7 and its corollary hold because there always exists a weighted potential function for payoffs that are a linear combination of local payoffs obtained from either potential games or weighted potential games with the same action sets and fixed weight vectors [106]. Hence, conditions on the local game interactions extend to the entire network game and its strategically differentiated version. This linear combination property does not hold for ordinal potential games [42]. Thus, assuming that the entire network game* Γ *is an ordinal potential game may not be sufficient for convergence of its strategically differentiated version* (Ξ, β)*. Up to now, we have not been able to find conditions for ordinal potential games that ensure that their differentiated versions share their fundamental convergence properties.*

5.4 The free-rider problem with strategic differentiation

In many social situations, individual members of a group can benefit from the efforts of other group members. When the individuals tend to be selfish the possibility to profit from others naturally results in trying to balance out one's efforts and rewards. In economics, the *free-rider* problem describes a situation in which a good or service becomes under-provided or even depleted as a result of selfish individuals profiting from a good without contributing to it. Here, we seek to determine how strategic differentiation can result in more desirable outcomes in which contributions to a good are preserved in the long run. The problem of finding a pure Nash equilibrium in a finite potential game is PLS-complete [107]. Therefore, we investigate the effect of strategic differentiation on the equilibrium action profiles of multiplayer games on networks via simulation. Unless stated otherwise, all simulations are conducted on a threshold public goods game which is a non-linear version of the public goods game described in Example 6. The non-linearity in the payoff function is created by requiring a minimum number of cooperators in order for the players to obtain a benefit from the local interaction. For all $i \in \mathcal{V}$, let τ_i denote this threshold value of cooperators in the local game played by the players in $\mathcal{N}_i$. The payoffs in the local game of the threshold public goods game are given by,

$$\forall j \in \mathcal{N}_i : \; u_{ji}(\boldsymbol{\sigma}) = \frac{r \sum_{j \in \mathcal{N}_i} (\sigma_j c_{ji})}{|\mathcal{N}_i| + 1} \theta_i(\sigma, \tau_i) - c_{ji}\sigma_j,$$

with $\theta_i(\sigma, \tau_i)$ defined by

$$\theta_i(\boldsymbol{\sigma}, \tau_i) = \begin{cases} 1, & \text{if } \sum_{j \in \mathcal{N}_i} \sigma_j \geq \tau_i, \\ 0, & \text{otherwise.} \end{cases}$$

This model is well established in the fields of economics, sociology and evolutionary biology and captures the free-rider problem because defectors can benefit from contributions of cooperators [96, 108, 109], and it is known that the cooperator thresholds in the model can promote moderate levels of cooperation at an equilibrium [110]. In the absence of thresholds (i.e. $\tau_i = 1$ for all $i \in \mathcal{V}$), it can be shown that the best response set for a player is solely determined by the degree distribution of the network and the public goods multiplier r, and thus, is static. The addition of thresholds thus allows for richer and more complex decision-making dynamics under best responses. In all simulations we start with 50% cooperators which are randomly assigned to the nodes on the network. The local contributions are determined by a players degree: for all $i \in \mathcal{V}$, $j \in \bar{\mathcal{N}}_i$, $c_{ij} = \frac{1}{|\mathcal{N}_i|+1}$. Hence, the total contributions that a player can make is

$\sum_{j\in\bar{\mathcal{N}}_i} c_{ij} = 1$. This corresponds to a set-up known as *fixed costs per individual* [96]. The total number of contributions in an equilibrium action profile $\hat{s}$ is determined as,

$$0 \leq \sum_{i\in\mathcal{D}} \sum_{j\in\mathcal{N}_i} c_{ij} s_{ij} + \sum_{h\in\mathcal{F}} \sum_{l\in\mathcal{N}_h} c_h \sigma_h \leq n.$$

5.4.1 Differentiated Best Response

In this section, we investigate the effect of differentiators on the existence of cooperation in a differentiated Nash equilibrium. For every player the threshold is equal to two, i.e., $\tau_i = 2$ for all $i \in \mathcal{V}$. The considered network has size 50 and was formed by a preferential attachment process and thus has high degree heterogeneity. When initially there are no differentiators in the network, the total contributions in the Nash equilibrium is close to zero. Because $r = 2.4 < \bar{d} := \frac{\sum_{i\in\mathcal{V}} |\mathcal{N}_i|}{N}$, this is consistent with a rule for the emergence of cooperation in games on networks proposed in [40]. When differentiators are added to the network the level of cooperation in equilibrium changes significantly. For low values of r, within the two-hop neighborhood of differentiators located at high degree nodes, cooperation starts to exist in the differentiated Nash equilibrium (Fig. 5.2). However, when the differentiators are placed on low degree nodes (*i.e.*, $|\mathcal{N}_i| \leq \bar{d}$) the total number of contributions in equilibrium tends to be lower than in the Nash equilibrium without differentiators. The same qualitative effects of strategic differentiation in a set-up known as *fixed cost per game*, in which the total contribution that a player can make is equal to their degree in the network. This illustrates that cooperation can be promoted if individuals with a large social network (i.e. hubs) can differentiate their actions. An explanation is that, in traditional network games, cooperating hubs can be taken advantage of by many players and therefore tend to defect when they cannot differentiate. Indeed, when the groupsize of a multiplayer game becomes larger the emergence of cooperation becomes more difficult [111]. When players located at hubs can differentiate however, they can cooperate against cooperators and defect against defectors, thereby promoting the emergence of cooperation in their neighborhood. On the other hand, when low degree players can differentiate, *network reciprocity* [8] becomes less effective because cooperators surrounded by other cooperators can start to free-ride in their separate local games. When the network has a narrow degree distribution as in small-world networks or regular networks, the effect of differentiators on the emergence of cooperation in the equilibrium action profile is not as pronounced and more differentiators are needed to make a significant change to the equilibrium action profile.

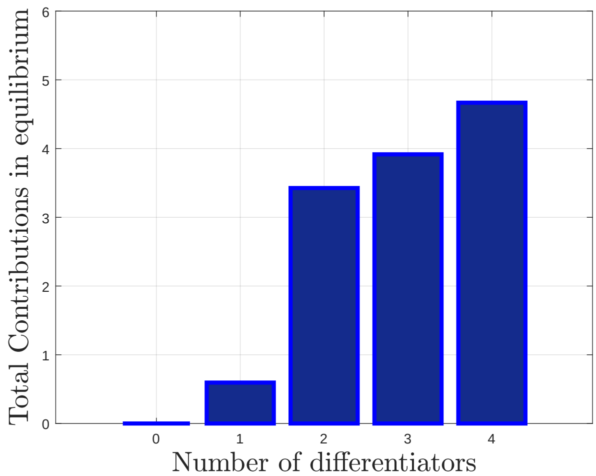


Figure 5.2: Each plot is obtained by averaging 10 trajectories generated by the differentiated myopic best response dynamics for the same initial condition with differentiators placed on high degree nodes. In both cases $n = 50$ and the network is generated by a preferential attachement process [92].

5.4.2 Differentiated Imitation

We have seen that the existence of differentiators and social influence in network games can promote the emergence of cooperation at an equilibrium action profile of a non-cooperative game on a network in which players seek to optimize their payoffs by playing best responses. Now we assume the players update their actions according to an *imitation* process in which each differentiator updates his/her action in a local game by imitating an action of the best performing players in that local game. The players who do not differentiate, update their actions by imitating one of their best performing neighbors [112]. For these imitation based dynamics, the effect of differentiators located on high degree nodes in the network is remarkable. For a *neutral* benefit to cost ratio ($r = 1$), increasing the number of differentiators tends to increase the level of cooperation in the equilibrium. When there are only four differentiators located at high degree nodes, almost half of the players cooperate at the equilibrium action profile. This behavior was consistent for different activation sequences. Such levels of cooperation cannot be seen without strategic differentiation. As in the differentiated best response dynamics, when differentiators employ cooperation against groups of cooperators and defection against groups of defectors, network reciprocity allows clusters of cooperators to emerge around differentiators in the equilibrium action profile even for very low values of r. More importantly, the concept of *false action attribution* seems to be crucial for large scale cooperation in the equilibrium of a social dilemma with imitation based update rules and strategic differentiation. In these games, the players who differentiate can obtain high payoffs when they defect against some of their cooperating opponents. Other players in the network observe these high payoffs and imitate the action that the differentiator employs in *their* local game. False action attribution occurs when that local action happens to be cooperation: a defecting neighbor of the differentiator is then likely to switch to cooperation, even though the differentiator obtained the payoffs mostly by defecting. This suggests that when the number of differentiators increases, not the payoff parameters, but the initial action profile and the network structure become determinative for the frequency of the actions in equilibrium. Indeed, when there are many differentiators in the network the influence of the benefit to cost ratio r on the total contributions in equilibrium is suppressed. This effect is similar to *topological enslavement* [113] seen in evolutionary games on multiplex networks in which hubs dominate the game dynamics. When the differentiators are placed on low degree nodes, these effects are mitigated.

5.5 Final Remarks

We have shown how network games can be extended to include a subset of players that can employ different actions against different opponents. When the local games in the network admit a weighted potential function convergence of the strategically differentiated version under myopic best response dynamics is guaranteed. For both imitation and best response like dynamics, the topology of the network, the existence, and location of differentiators in the networks can crucially alter the action profile at an equilibrium of groupwise games. When differentiators are plentiful the equilibrium action profile becomes less sensitive to changes in the values of the payoff parameters. The convergence results in this framework can be combined with those of Chapter 3 and 4. The combination of relative best responses, rational imitation and strategic differentiation allows us to study the behavior of many classic games from a novel perspective.

This chapter concludes the first part of the thesis. We have mainly focused on network games and structural solutions to social dilemmas via network reciprocity that allows for the emergence of cooperation via a spatial or social structure. Part II of the thesis focuses on direct reciprocity in repeated games. In this setting there is no network structure, in stead, cooperation can evolve through the expectation of repeated interactions with a fixed group of players. In the following, we will investigate how such probabilistic decision-making processes can be studied by their average behavior and show how strategic individuals can exert significant control in the long-run outcomes of repeated games.

Part II

STRATEGIC PLAY AND CONTROL IN REPEATED GAMES

Chapter 6

Exerting control in finitely repeated n-player social dilemmas

> The advantage of a bad memory is that one enjoys several times the same good things for the first time.
>
> *Friedrich Nietzsche*

The functionalities of many complex social systems rely on their composing individuals' willingness to set aside their interest for the benefit of the greater good [39]. In the previous chapters, we have studied how social influence and network structure can promote these selfless decisions. Another mechanism for the evolution of cooperation is known as *direct reciprocity*: even if in the short run it pays off to be selfish, mutual cooperation can be favored when the individuals encounter each other repeatedly. Direct reciprocity is often studied in the standard model of repeated games and it is only recently, inspired by the discovery of a novel class of strategies, called zero-determinant (ZD) strategies [64], that repeated games began to be examined from a new angle by investigating the level of *control* that a single player can exert on the average payoffs of its opponent. In [64] Press and Dyson showed that in *infinitely* repeated 2×2 prisoners dilemma games if a player can remember the actions in the previous round, this player can *unilaterally* impose some linear relation between his/her own expected payoff and that of his/her opponent. It is emphasized that this enforced linear relation cannot be avoided even if the opponent employs some intricate strategy with a large memory. Such strategies are called *zero-determinant* because they enforce a particular matrix, that depends on the player's strategy to have a determinant equal to zero. Later, ZD strategies were extended to games with more than two possible actions [114], continuous action spaces [115], and alternative moves [116]. Most of the literature has focused on two-player games; however, in [117] the existence of ZD strategies in infinitely repeated public goods games was shown by extending the arguments in [64] to a symmetric public goods game. Around the same time, characterization of the feasible ZD strategies in multiplayer social dilemmas and those strategies that maintain cooperation in such n-player games were reported in [118]. In this chapter, we study the existence of ZD strategies in n-player social dilemmas with a finite but undetermined number of rounds. That is, future payoffs are discounted using a fixed and common *discount factor* [53]. In doing so, we will unravel how an individual can exert a significant level of control under the "shadow of the future".

6.1 Symmetric n-player games

We consider n-player games in which players can repeatedly choose to either cooperate or defect. The set of actions for each player is denoted by $\mathcal{A} = \{C, D\}$. The actions chosen in the group in round t of the repeated game is described by an action profile $\sigma^t \in \boldsymbol{\mathcal{A}} = \{C, D\}^n$. A player's payoff in a given round depends on the player's action and the actions of the $n-1$ co-players. In a group in which z co-players cooperate, a cooperator receives payoff a_z, whereas a defector receives b_z. As in [117, 118] we assume the game is symmetric, such that the outcome of the game depends only on one's own decision and the number of cooperating co-players, and hence does not depend on which of the co-players have cooperated. Accordingly, the payoffs of all possible outcomes for a player can be conveniently summarized as in Table 6.1.

Table 6.1: Payoffs of the symmetric n-player games. A player's payoff depends on its own decision and the number of co-players who cooperate.

Number of cooperators among co-players	$n-1$	$n-2$	...	2	1	0
Cooperator's payoff	a_{n-1}	a_{n-2}	...	a_2	a_1	a_0
Defector's payoff	b_{n-1}	b_{n-2}	...	b_2	b_1	b_0

We have the following assumptions on the payoffs of the symmetric n-player game.

Assumption 7 (Social dilemma assumption [118,119])**.** *The payoffs of the symmetric n-player game satisfy the following conditions:*

a) For all $0 \leq z < n-1$, *it holds that* $a_{z+1} \geq a_z$ *and* $b_{z+1} \geq b_z$.

b) For all $0 \leq z < n-1$, *it holds that* $b_{z+1} > a_z$.

c) $a_{n-1} > b_0$.

Assumption 7 is standard in n-player social dilemma games and it ensures that there is a conflict between the interest of each individual and that of the group as a whole. Thus, those games whose payoffs satisfy Assumption 7 can model a social dilemma that results from selfish behaviors in a group. Besides the classic prisoner's dilemma game, examples of n-player games that satisfy these assumptions include the n-player public goods game [109], the volunteers dilemma [120], and the n-player snowdrift and stag-hunt games [109]. Detailed examples can be found in Section 6.5.

6.1.1 Strategies in repeated games

In repeated games, the players must choose how to update their actions as the game interactions are repeated over multiple rounds of plays. A *strategy* of a player determines the conditional probabilities with which actions are chosen by the player. To formalize this concept we introduce some additional notation. A history of plays up to round t is denoted by $h^t = (\sigma^0, \sigma^1, \ldots, \sigma^{t-1}) \in \boldsymbol{\mathcal{A}}^t$ such that each $\sigma^k \in \boldsymbol{\mathcal{A}}$ for all $k = 0 \ldots t-1$. The union of possible histories is denoted by $\mathcal{H} = \cup_{t=0}^{\infty} \boldsymbol{\mathcal{A}}^t$, with $\boldsymbol{\mathcal{A}}^0 = \emptyset$ being the empty set. Finally, let $\Delta(\mathcal{A})$ denote the probability distribution over the action set $\mathcal{A}$. As is standard in the theory of repeated games, a strategy of player i is then defined by a function $\rho : \mathcal{H} \to \Delta(\mathcal{A})$ that maps the history of play to the probability distribution over the action set. An interesting and important subclass is comprised of those strategies that only take into account the action profile in round $t-1$, (i.e. $\sigma^{t-1} \in h^t$) to determine the conditional probabilities to choose some action in round t. Correspondingly these strategies are called *memory-one strategies* and are formally defined as follows.

Definition 23 (Memory-one strategy [121])**.** *A strategy ρ is a memory-one strategy if* $\rho(h^t) = \rho(\hat{h}^{t'})$ *for all histories* $h^t = (\sigma^0, \ldots, \sigma^{t-1})$ *and* $\hat{h}^{t'} = (\hat{\sigma}^0, \ldots, \hat{\sigma}^{t'-1})$ *with* $t, t' \geq 1$ *and* $\sigma^{t-1} = \hat{\sigma}^{t'-1}$.

The theory of Press and Dyson showed that, for determining the best performing strategies in terms of expected payoffs in two-action repeated games, it is sufficient to consider only the space of memory-one strategies [64, 114].

6.2 Mean distributions and memory-one strategies

In this section we zoom in on a particular player that employs a memory-one strategy in the n-player game and refer to this player as the *key player*. In particular, we focus on the relation between the mean distribution of the action profile and the memory-one strategy of the key player. Let $p_\sigma \in [0,1]$ denote the probability that the key player cooperates in the next round given that the current action profile is $\sigma \in \boldsymbol{\mathcal{A}}$. By stacking these probabilities for all possible outcomes into a vector, we obtain the memory-one strategy $\mathbf{p} = (p_\sigma)_{\sigma \in \boldsymbol{\mathcal{A}}}$ whose elements are conditional probabilities for the key player to cooperate in next round. Accordingly, the memory-one strategy $\mathbf{p}_\sigma^{\text{rep}}$, gives the probability to cooperate when the current action is simply repeated. To be more precise, let $\sigma = (\sigma_i, \sigma_{-i}) \in \boldsymbol{\mathcal{A}}$, where $\sigma_i \in \{C, D\}$ and $\sigma_{-i} \in \{C, D\}^{n-1}$. Then for all σ_{-i}, the entries of the repeat strategy are given by $\mathbf{p}^{\text{rep}}_{C,\sigma_{-i}} = 1$ and $\mathbf{p}^{\text{rep}}_{D,\sigma_{-i}} = 0$. To describe the relation between the memory-one strategy and the mean distribution of the action profile we introduce some additional notation. Let $v_\sigma(t)$ denote the probability that the outcome of round t is $\sigma \in \boldsymbol{\mathcal{A}}$. And let $v(t) = (v_\sigma(t))_{\sigma \in \boldsymbol{\mathcal{A}}}$ be the vector of outcome probabilities in round t. As in [115, 116, 121, 122], we focus on repeated games with a finite but undetermined expected number of rounds.[1]. Given the current round, a fixed and common discount factor $0 < \delta < 1$ determines the probability that a next round takes place. By taking the limit of the geometric sum of δ, the expected number of rounds is $\frac{1}{1-\delta}$. As in [121], the mean distribution of $v(t)$ is:

$$\mathbf{v} = (1 - \delta) \sum_{t=0}^{\infty} \delta^t v(t). \tag{6.1}$$

As is common in the theory of repeated games, we are interested in the *average discounted payoffs* of the repeated n-player game. Alternatively, this can be interpreted as the expect payoff to the player at the end of the repeated game. Let g^i_σ denote the payoff in a given round that player i receives in the action profile $\sigma \in \mathcal{A}$. By stacking the possible payoffs we obtain the vector $g^i = (g^i_\sigma)_{\sigma \in \boldsymbol{\mathcal{A}}}$ that contains all possible payoffs in a given round of player i. The expected "one-shot" payoff of player i in round t is $\pi_i(t) = g^i \cdot v(t)$. And the average discounted payoff in repeated game

[1]There is some inconsistency in the repeated games literature and the ZD literature about the terminology "finitely repeated". Here we adopt the terminology of [122], in which "finite" refers to the *expected* number of rounds $\frac{1}{1-\delta}$ with $\delta < 1$. In the repeated games literature, this is referred to as an infinitely repeated game with a finite but undermined number of rounds, or simply as a infinite repeated game with discounting. In this case, "infinite" refers to the infinite horizon sum over which the expected number of rounds and the expected payoff (see Eq. (6.2)) are calculated.

for player i is

$$\pi_i = (1-\delta)\sum_{t=0}^{\infty}\delta^t\pi_i(t) = (1-\delta)\sum_{t=0}^{\infty}\delta^t g^i\cdot v(t) = g^i\cdot\mathbf{v}. \tag{6.2}$$

The following lemma relates the limit distribution $\mathbf{v}$ of the finitely repeated game to the memory-one strategy $\mathbf{p}$ of the key player. The presented lemma is a straightforward n-player extension of the 2-player case that is given in [121] and relies on the fundamental results from [123].

Lemma 9 (Limit distribution)**.** *Suppose the key player applies memory-one strategy* $\mathbf{p}$ *and the strategies of the other players are arbitrary, but fixed. For the finitely repeated n-player game, it holds that*

$$(\delta\mathbf{p}-\mathbf{p}^{\text{rep}})\cdot\mathbf{v} = -(1-\delta)p_0,$$

where p_0 is the key player's initial probability to cooperate.

Proof. The probability that i cooperated in round t is $q_C(t) = \mathbf{p}^{\text{rep}}\cdot v(t)$. And the probability that i cooperates in round $t+1$ is $q_C(t+1) = \mathbf{p}\cdot v(t)$. Now define,

$$u(t) := \delta q_C(t+1) - q_C(t) = (\delta\mathbf{p}-\mathbf{p}^{\text{rep}})\cdot v(t). \tag{6.3}$$

Multiplying Eq. (6.3) by $(1-\delta)\delta^t$ and summing up over $t = 0,\ldots,\tau$ we obtain

$$(1-\delta)\sum_{t=0}^{\tau}\delta^t u(t) = (1-\delta)(\delta q_c(1) - q_c(0) + \delta^2 q_c(2) - \delta q_c(1) +$$
$$\cdots + \delta^{\tau+1}q_c(\tau+1) - \delta^\tau q_c(\tau) = (1-\delta)\delta^{\tau+1}q_c(\tau+1) - (1-\delta)q_c(0).$$

Because $0<\delta<1$, it follows that

$$\lim_{\tau\to\infty}(1-\delta)\sum_{t=0}^{\tau}\delta^t u(t) = -(1-\delta)p_0.$$

And by the definition of $\mathbf{v}$ in Eq. (6.1):

$$\lim_{\tau\to\infty}(1-\delta)\sum_{t=0}^{\tau}\delta^t(\delta\mathbf{p}-\mathbf{p}^{\text{rep}})\cdot v(t) = (\delta\mathbf{p}-\mathbf{p}^{\text{rep}})\cdot\mathbf{v}.$$

By substituting $u(t)$ back into the equation we obtain

$$(\delta\mathbf{p}-\mathbf{p}^{\text{rep}})\cdot\mathbf{v} = -(1-\delta)p_0.$$

This completes the proof. □

Remark 8 (An infinite expected number of rounds)**.** *Note that in the limit $\delta \to 1$, the infinitely repeated game is recovered. In this setting, the expected number of rounds is infinite. And, if the limit exists, the average payoffs are given by*

$$\pi_i = \lim_{\tau \to \infty} \frac{1}{\tau + 1} \sum_{t=0}^{\tau} \pi_i(t).$$

By Akin's Lemma (see [118, 123]), for the infinitely repeated game without discounting, irrespective of the initial probability to cooperate, it holds that

$$(\mathbf{p} - \mathbf{p}^{\text{rep}}) \cdot \mathbf{v} = 0. \tag{6.4}$$

Hence, a key difference between the infinitely repeated and finitely repeated games is that p_0 is important for the relation between the memory-one strategy $\mathbf{p}$ and the mean distribution $\mathbf{v}$ when the game is repeated a finite number of expected rounds. When the game is infinitely repeated, i.e. $\delta \to 1$, the importance of the initial conditions on the relation between $\mathbf{p}$ and $\mathbf{v}$ disappears [118].

6.3 ZD strategies in finitely repeated n-player games

Based on Lemma 9 we now formally define a ZD strategy for a finitely repeated n-player game. To this end, let $\mathbb{1} = (1)_{\sigma \in \boldsymbol{\mathcal{A}}}$.

Definition 24 (ZD strategy)**.** *A memory-one strategy $\mathbf{p}$ is a ZD strategy for an n-player game if there exist constants α, β_j, γ, $1 \leq j \leq n$ with $\sum_{j \neq i}^{n} \beta_j \neq 0$ such that*

$$\delta \mathbf{p} = \mathbf{p}^{\text{rep}} + \alpha g^i + \sum_{j \neq i}^{n} \beta_j g^j + (\gamma - (1 - \delta) p_0) \mathbb{1}. \tag{6.5}$$

The following proposition shows how the ZD strategy can enforce a linear relation between the key player's expected payoff and that of her co-players.

Proposition 2 (Enforcing a linear payoff relation)**.** *Suppose the key player employs a fixed ZD strategy with parameters α, β_j and γ as in definition 24. Then, irrespective of the fixed strategies of the remaining $n-1$ co-players, the payoffs obey the equation*

$$\alpha \pi_i + \sum_{j \neq i}^{n} \beta_j \pi_j + \gamma = 0. \tag{6.6}$$

Proof.

$$
\begin{aligned}
(\delta\mathbf{p} - \mathbf{p}^{\text{rep}}) &= \alpha g^i + \sum_{j\neq i}^{n} \beta_j g^j + (\gamma - (1-\delta)p_0)\mathbb{1} \\
(\delta\mathbf{p} - \mathbf{p}^{\text{rep}}) \cdot \mathbf{v} &= \alpha\pi_i + \sum_{j\neq i}^{n} \beta_j \pi_j + \gamma - (1-\delta)p_0 \\
(\delta\mathbf{p} - \mathbf{p}^{\text{rep}}) \cdot \mathbf{v} + (1-\delta)p_0 &= \alpha\pi_i + \sum_{j\neq i}^{n} \beta_j \pi_j + \gamma \\
0 &= \alpha\pi_i + \sum_{j\neq i}^{n} \beta_j \pi_j + \gamma.
\end{aligned}
\tag{6.7}
$$

□

To be consistent with the earlier work on ZD strategies in infinitely repeated n-player games [118], we introduce the parameter transformations:

$$
l = \frac{-\gamma}{(\alpha + \sum_{k\neq i}^{n} \beta_k)}, \; s = \frac{-\alpha}{\sum_{k\neq i}^{n} \beta_k},
$$

$$
w_{j\neq i} = \frac{\beta_j}{\sum_{k\neq i}^{n} \beta_k}, \;\; \phi = -\sum_{k\neq i}^{n} \beta_k, \;\; w_i = 0.
$$

Using these parameter transformations, Eq. (6.5) can be written as

$$
\delta\mathbf{p} = \mathbf{p}^{\text{rep}} + \phi \left[s g^i - \sum_{j\neq i}^{n} w_j g^j + (1-s) l \mathbb{1} \right] - (1-\delta)p_0\mathbb{1}, \tag{6.8}
$$

under the conditions that $\phi \neq 0$, $w_i = 0$ and $\sum_{j=1}^{n} w_j = 1$. Moreover, the linear payoff relation in Eq. (6.6) becomes

$$
\pi^{-i} = s\pi_i + (1-s)l, \tag{6.9}
$$

where $\pi^{-i} = \sum_{j\neq i}^{n} w_j \pi_j$.

Remark 9. *When all weights are equal, i.e. $w_j = \frac{1}{n-1}$ for all $j \neq i$, the formulation of the ZD strategy for a symmetric multiplayer social dilemma can be simplified using only the number of cooperators in the social dilemma. To this end, let $g^{-i}_{\sigma_i,z}$ denote the average payoff of the $n-1$ co-players of i when player i selects action $\sigma_i \in \{C, D\}$ and $0 \leq z \leq n-1$ co-players cooperate. Using the payoffs in Table 6.1 this may be written as $g^{-i}_{C,z} = \frac{a_z z + (n-1-z)b_{z+1}}{n-1}$, and $g^{-i}_{D,z} = \frac{a_{z-1} z + (n-1-z)b_z}{n-1}$. We obtain $g^{-i} = (g^{-i}_{\sigma_i,z})$ by stacking these payoffs into a vector. Similarly, let $v_{\sigma_i,z}(t)$ denote the probability that* *at round t, player i chooses action σ_i and z co-players cooperate. By stacking these probabilities into a vector we obtain $v(t) = (v_{\sigma_i,z}(t))$. The expected payoff of player i at time t is again given by $\pi^i(t) = g^i \cdot v(t)$. Moreover, the average expected payoff of the co-players at time t can be conveniently written as $\pi^{-i}(t) = g^{-i} \cdot v(t)$. The mean distribution of $v(t)$ is again obtained by using* Eq. (6.1)*, but now the entries of $\mathbf{v}$ provide the fraction of rounds in the repeated game in which player i chooses σ_i and z players cooperate. Then $\pi^i = g^i \cdot \mathbf{v}$ and $\pi^{-i} = g^{-i} \cdot \mathbf{v}$ which leads to the ZD strategy*

$$\delta\mathbf{p} = \mathbf{p}^{\text{rep}} + \alpha g^i + g^{-i} + (\gamma - (1-\delta)p_0)\mathbb{1}_{2n}.$$

The four most widely studied ZD strategies are given in Table 6.2. When the mutual cooperation payoff a_{n-1} results in the highest possible *average payoff of the group*, the enforced payoff relation of generous ZD strategies ensure $\pi_{-i} \geq \pi_i$. On the other hand, when mutual defection gives the lowest possible average payoff of the group, extortionate ZD strategies ensure $\pi_{-i} \leq \pi_i$. However, in both cases, the positive slopes (s) of the linear payoff relation Eq. (6.9) ensures that the payoff of the of the ZD strategist and the average payoff of his/her co-players are positively related. Implying that the collective best response of the co-players is to maximize the payoff of the ZD strategist by cooperating.

Table 6.2: The four most widely studied ZD strategies. Depending on the parameter values s and l, players may be fair, generous, extortionate or equalizers.

ZD strategy	Parameter values	Enforced relation	Typical relation
Fair	$s = 1$	$\pi_{-i} = \pi_i$	$\pi_{-i} = \pi_i$
Generous	$l = a_{n-1}$, $0 < s < 1$	$\pi_{-i} = s\pi_i + (1-s)a_{n-1}$	$\pi_{-i} \geq \pi_i$
Extortionate	$l = b_0$, $0 < s < 1$	$\pi_{-i} = s\pi_i + (1-s)b_0$	$\pi_{-i} \leq \pi_i$
Equalizer	$s = 0$	$\pi_{-i} = l$	$\pi_{-i} = l$

Because the entries of the ZD strategy correspond to conditional probabilities, they are required to belong to the unit interval. Hence, not every linear payoff relation with parameters s, l is valid. Let $w = (w_i) \in \mathbb{R}^{n-1}$ denote the vector of weights that the ZD strategist assigns to her co-players. Consider the following definition that was given in [121] for two-player games.

Definition 25 (Enforceable payoff relations)**.** *Given a discount factor $0 < \delta < 1$, a payoff relation $(s, l) \in \mathbb{R}^2$ with weights w is enforceable if there are $\phi \in \mathbb{R}$ and $p_0 \in [0, 1]$, such that each entry in $\mathbf{p}$ according to* Eq. (6.5) *is in $[0, 1]$. We indicate the set of enforceable payoff relations by $\mathcal{E}_\delta$.*

An intuitive implication of decreasing the expected number of rounds in the repeated game (e.g. by decreasing δ) is that the set of enforceable payoff relations will decrease as well. This monotone effect is formalized in the following proposition that extends a result from [121] to the n-player case.

Proposition 3 (Monotonicity of $\mathcal{E}_\delta$)**.** *If* $\delta' \leq \delta''$*, then* $\mathcal{E}_{\delta'} \leq \mathcal{E}_{\delta''}$*.*

Proof. Albeit with different formulations of $\mathbf{p}$, the proof follows from the same argument used in the two-player case [118]. It is presented here to make the chapter self-contained. From Definition 25, $(s, l) \in \mathcal{E}_\delta$ if and only if one can find $\phi \in \mathbb{R}$ and $p_0 \in [0, 1]$ such that the entries of $\mathbf{p}$ are in the closed unit interval. Let $\mathbb{0} = (0)_{\sigma \in \boldsymbol{\mathcal{A}}}$, we have

$$\mathbb{0} \leq \mathbf{p} \leq \mathbb{1} \Rightarrow \mathbb{0} \leq \delta \mathbf{p} \leq \delta \mathbb{1}. \tag{6.10}$$

Then by substituting Eq. (6.8) into the above inequality we obtain,

$$p_0(1-\delta)\mathbb{1} \leq \mathbf{p}^\infty \leq \delta \mathbb{1} + (1-\delta)p_0\mathbb{1}, \tag{6.11}$$

with

$$\mathbf{p}^\infty = \mathbf{p}^{\text{rep}} + \phi \left[s g^i - \sum_{j \neq i}^{n} w_j g^j + (1-s) l \mathbb{1} \right].$$

Now observe that $p_0(1-\delta)\mathbb{1}$ on the left-hand side of the inequality Eq. (6.11) is decreasing for increasing δ. Moreover, $\delta\mathbb{1} + (1-\delta)p_0\mathbb{1}$ on the right-hand side of the inequality is increasing for increasing δ. The middle part of the inequality, which is exactly the definition of a ZD strategy for the infinitely repeated game in [118], is independent of δ. It follows that by increasing δ the range of possible ZD parameters (s, l, ϕ) and p_0 increases and hence if $\mathbb{0} \leq \mathbf{p} \leq \mathbb{1}$ is satisfied for some δ', then it is also satisfied for some $\delta'' \geq \delta'$. □

Now we are ready to state the existence problem studied in this chapter.

Problem 1 (The existence problem in n-player social dilemmas)**.** *For the class of n-player games with payoffs as in Table 6.1 that satisfy Assumption 7, what are the enforceable payoff relations when the expected number of rounds is finite, i.e.,* $\delta \in (0, 1)$*?*

6.4 Existence of ZD strategies

In this section, the main results on the existence problem are given. We begin by formulating conditions on the parameters of the ZD strategy that are necessary for the payoff relation to be enforceable in the finitely repeated n-player game.

Proposition 4. *The enforceable payoff relations (l, s, w) for the finitely repeated n-player game with $0 < \delta < 1$, with payoffs as in Table 6.1 that satisfy Assumption 7, require the following necessary conditions:*

$$-\frac{1}{n-1} \leq -\underset{j \neq i}{min}\ w_j < s < 1,$$

$$\phi > 0,$$

$$b_0 \leq l \leq a_{n-1}, \tag{6.12}$$

with at least one strict inequality in Eq. (6.12).

Proof. Suppose all players are cooperating e.g. $\sigma = (C, C, \ldots, C)$. Then from the definition of $\delta\mathbf{p}$ in Eq. (6.8) and the payoffs given in Table 6.1, it follows that

$$\delta\mathbf{p}_{(C,C,\ldots,C)} = 1 + \phi(1-s)(l - a_{n-1}) - (1-\delta)p_0. \tag{6.13}$$

Now suppose that all players are defecting. Similarly, we have

$$\delta\mathbf{p}_{(D,D,\ldots,D)} = \phi(1-s)(l - b_0) - (1-\delta)p_0. \tag{6.14}$$

In order for these payoff relations to be enforceable, it needs to hold that both entries in Eq. (6.13) and Eq. (6.14) are in the interval $[0, \delta]$. Equivalently,

$$(1-\delta)(1-p_0) \leq \phi(1-s)(a_{n-1} - l) \leq 1 - (1-\delta)p_0, \tag{6.15}$$

and

$$0 \leq p_0(1-\delta) \leq \phi(1-s)(l - b_0) \leq \delta + (1-\delta)p_0 \tag{6.16}$$

Combining Eq. (6.15) and Eq. (6.16) it follows that $0 < (1-\delta) \leq \phi(1-s)(a_{n-1} - b_0)$. From the assumption that $a_{n-1} > b_0$ listed in Assumption 7, it follows that

$$0 < \phi(1-s). \tag{6.17}$$

Now suppose there is a single defecting player, i.e., $\sigma = (C, C, \ldots, D)$ or any of its permutations. In this case, the entries of the memory-one strategy are

$$\delta\mathbf{p}_\sigma = \begin{cases} 1 + \phi[sa_{n-2} - (1-w_j)a_{n-2} - w_j b_{n-1} + (1-s)l] - (1-\delta)p_0, & \text{if } x_i = C; \\ \phi[sb_{n-1} - a_{n-2} + (1-s)l] - (1-\delta)p_0, & \text{if } x_i = D. \end{cases} \tag{6.18}$$

Again, for both cases we require $\delta\mathbf{p}_\sigma$ to be in the interval $[0, \delta]$. This results in the inequalities given in Eq. (6.19) and Eq. (6.20).

$$0 \leq p_0(1-\delta) \leq \phi[sb_{n-1} - a_{n-2} + (1-s)l] \leq \delta + (1-\delta)p_0 \tag{6.19}$$

$$(1-\delta)(1-p_0) \leq \phi[-sa_{n-2} + (1-w_j)a_{n-2} + w_j b_{n-1} - (1-s)l] \leq 1 - p_0(1-\delta) \tag{6.20}$$

By combining these equations we obtain

$$0 < (1-\delta) \leq \phi(s + w_j)(b_{n-1} - a_{n-2}). \tag{6.21}$$

Because of the assumption $b_{z+1} > a_z$ it follows that

$$0 < \phi(s + w_j), \ \forall j \neq i. \tag{6.22}$$

Then, Eq. (6.22) and Eq. (6.17) together imply that

$$0 < \phi(1 + w_j), \forall j \neq i. \tag{6.23}$$

Because at least one $w_j > 0$, it follows that

$$\phi > 0. \tag{6.24}$$

Combining with Eq. (6.17) we obtain

$$s < 1. \tag{6.25}$$

In combination with Eq. (6.22) it follows that

$$\forall j \neq i : s + w_j > 0 \Leftrightarrow \forall j \neq i : w_j > -s \Leftrightarrow \min_{j \neq i} w_j > -s. \tag{6.26}$$

The inequalities in Eq. (6.25) and Eq. (6.26) finally produce the bounds on slope:

$$-\min_{j \neq i} w_j < s < 1. \tag{6.27}$$

Moreover, because it is required that $\sum_{j=1}^{n} w_j = 1$, it follows that $\min_{j \neq i} w_j \leq \frac{1}{n-1}$. Hence the necessary condition turns into:

$$-\frac{1}{n-1} \leq -\min_{j \neq i} w_j < s < 1. \tag{6.28}$$

We continue to show the necessary upper and lower bound on l. From Eq. (6.15) we obtain:

$$\phi(1-s)(l - a_{n-1}) \leq (1-p_0)(\delta - 1) \leq 0. \tag{6.29}$$

From Eq. (6.17) we know $\phi(1-s)>0$. Together with Eq. (6.29) this implies the necessary condition

$$l - a_{n-1} \leq 0 \Leftrightarrow l \leq a_{n-1}. \tag{6.30}$$

We continue with investigating the lower-bound on l, from Eq. (6.16)

$$0 \leq p_0(1-\delta) \leq \phi(1-s)(l-b_0) \leq \delta + (1-\delta)p_0. \tag{6.31}$$

Because $\phi(1-s)>0$ (see Eq. (6.17) it follows that

$$l \geq b_0.$$

Naturally, when $l = a_{n-1}$ by assumption 7 it holds that $l > b_0$ and when $l = b_0$ then $l < a_{n-1}$.

□

Because fair strategies are defined with the slope $s = 1$ (see, Table 6.2), an immediate consequence of Proposition 4 is stated in the following corollary.

Corollary 5. *For repeated n-player social dilemma with a finite number of expected rounds and payoffs that satisfy Assumption 7, there do not exist fair ZD strategies.*

In the following result, we extend the theory for infinitely repeated n-player games from [118] to repeated games with a finite number of expected interactions. To write the statement compactly, let $a_{-1} = b_n = 0$. Moreover, let $\hat{w}_z = \min_{w_h \in w} (\sum_{h=1}^{z} w_h)$ denote the sum of the z smallest weights and let $\hat{w}_0 = 0$.

***Theorem* 8** (Characterizing the enforceable set)**.** *For the finitely repeated n-player game with payoffs as in Table 6.1 that satisfy Assumption 7, the payoff relation $(s,l) \in \mathbb{R}^2$ with weights $w \in \mathbb{R}^{n-1}$ is enforceable if and only if $-\frac{1}{n-1} < s < 1$ and*

$$\begin{aligned} \max_{0 \leq z \leq n-1} \left\{ b_z - \frac{\hat{w}_z (b_z - a_{z-1})}{(1-s)} \right\} \leq l, \\ \min_{0 \leq z \leq n-1} \left\{ a_z + \frac{\hat{w}_{n-z-1} (b_{z+1} - a_z)}{(1-s)} \right\} \geq l, \end{aligned} \tag{6.32}$$

moreover, at least one inequality in Eq. (6.32) *is strict.*

Proof. In the following we refer to the key player, who is employing the ZD strategy, as player i. Let $\sigma = (x_1, \ldots, x_n)$ such that $x_k \in \mathcal{A}$ and let σ^C be the number of $i's$ co-players that cooperate and let $\sigma^D = n - 1 - \sigma^C$, be the number of $i's$ co-players that defect. Also, let $|\sigma|$ be the total number of cooperators including player i. Using this notation, for some action profile σ we may write the ZD strategy as

$$\delta \mathbf{p}_\sigma = \mathbf{p}^{\text{rep}} + \phi[(1-s)(l - g_\sigma^i) + \sum_{j \neq i}^{n} w_j (g_\sigma^i - g_\sigma^j)] - (1-\delta)p_0. \tag{6.33}$$

Also, note that

$$\sum_{j\neq i}^{n} w_j g_\sigma^j = \sum_{k\in\sigma^D} w_k g_\sigma^k + \sum_{h\in\sigma^C} w_h g_\sigma^h, \tag{6.34}$$

and because $\sum_{j\neq i}^{n} w_j = 1$ it holds that

$$\sum_{l\in\sigma^C} w_l = 1 - \sum_{k\in\sigma^D} w_k.$$

Substituting this into Eq. (6.34) and using the payoffs as in Table 6.1 we obtain

$$\sum_{j\neq i}^{n} w_j g_\sigma^j = a_{|\sigma|-1} + \sum_{j\in\sigma^D} w_j (b_{|\sigma|} - a_{|\sigma|-1}).$$

Accordingly, the entries of the ZD strategy $\delta\mathbf{p}_\sigma$ are

$$\delta\mathbf{p}_\sigma = \begin{cases} 1+\phi\left[(1-s)(l-a_{|\sigma|-1}) - \sum_{j\in\sigma^D} w_j(b_{|\sigma|}-a_{|\sigma|-1})\right] - (1-\delta)p_0, & \text{if } x_i = C, \\ \phi\left[(1-s)(l-b_{|\sigma|}) + \sum_{j\in\sigma^C} w_j(b_{|\sigma|}-a_{|\sigma|-1})\right] - (1-\delta)p_0, & \text{if } x_i = D. \end{cases} \tag{6.35}$$

For all $\sigma \in \boldsymbol{\mathcal{A}}$ we require that

$$0 \leq \delta\mathbf{p}_\sigma \leq \delta. \tag{6.36}$$

This leads to the inequalities,

$$0 \leq (1-\delta)(1-p_0) \leq \phi\left[(1-s)(a_{|\sigma|-1}-l) + \sum_{j\in\sigma^D} w_j(b_{|\sigma|}-a_{|\sigma|-1})\right] \leq 1-(1-\delta)p_0, \tag{6.37}$$

$$0 \leq (1-\delta)p_0 \leq \phi\left[(1-s)(l-b_{|\sigma|}) + \sum_{j\in\sigma^C} w_j(b_{|\sigma|}-a_{|\sigma|-1})\right] \leq \delta+(1-\delta)p_0. \tag{6.38}$$

Because $\phi > 0$ can be chosen arbitrarily small, the inequalities in Eq. (6.37) can be satisfied for some $\delta \in (0,1)$ and $p_0 \in [0,1]$ if and only if for all σ such that $x_i = C$ the inequalities in Eq. (6.39) are satisfied.

$$0 \leq (1-s)(a_{|\sigma|-1}-l) + \sum_{j\in\sigma^D} w_j(b_{|\sigma|}-a_{|\sigma|-1}). \tag{6.39}$$

The inequality Eq. (6.39) together with the necessary condition $s < 1$ (see also Proposition 4) implies that

$$a_{|\sigma|-1} + \frac{\sum\limits_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \geq l, \tag{6.40}$$

and thus provides an upper-bound on the enforceable baseline payoff l. We now turn our attention to the inequalities in Eq. (6.38) that can be satisfied if and only if for all σ such that $x_i = D$ the following holds

$$\begin{aligned} 0 &\leq (1-s)(l - b_{|\sigma|}) + \sum_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1}) \\ &\xRightarrow{(1-s)>0} b_{|\sigma|} - \frac{\sum\limits_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \leq l. \end{aligned} \tag{6.41}$$

Combining Eq. (6.41) and Eq. (6.40) we obtain

$$\begin{aligned} \max_{|\sigma| s.t. x_i=D} &\left\{ b_{|\sigma|} - \frac{\sum\limits_{l\in\sigma^C} w_l(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \right\} \leq l, \\ l \leq \min_{|\sigma| s.t. x_i=C} &\left\{ a_{|\sigma|-1} + \frac{\sum\limits_{k\in\sigma^D} w_k(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \right\}. \end{aligned} \tag{6.42}$$

Because $b_{|\sigma|} - a_{|\sigma|-1} > 0$ and $(1-s) > 0$ the minima and maxima of the bounds in Eq. (6.42) are achieved by choosing the w_j as small as possible. That is, the extrema of the bounds on l are achieved for those states $\sigma|_{x_i=D}$ in which $\sum\limits_{l\in\sigma^C} w_l$ is minimum and those $\sigma|_{x_i=C}$ in which $\sum\limits_{k\in\sigma^D} w_k$ is minimum. Let $\hat{w}_z = \min\limits_{w_h\in w}(\sum_{h=1}^{z} w_h)$ denote the sum of the j smallest weights and let $\hat{w}_0 = 0$. By the above reasoning, Eq. (6.42) can be equivalently written as in the theorem in the main text. Now, suppose we have a *non-strict* upper-bound on the base-level payoff, i.e.,

$$l = a_{|\sigma|-1} + \frac{\sum\limits_{k\in\sigma^D} w_k(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)}.$$

Then from Eq. (6.37) it follows that $p_0 = 1$ is required. Then Eq. (6.38) implies

$$\begin{aligned} 0 &< (1-s)(l - b_{|\sigma|}) + \sum_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1}) \\ &\xRightarrow{(1-s)>0} b_{|\sigma|} - \frac{\sum\limits_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} < l. \end{aligned} \tag{6.43}$$

This is exactly the corresponding lower-bound of l, which is thus required to be strict when the upper-bound is non-strict.

Now suppose we have a non-strict lower bound, e.g.

$$l = b_{|\sigma|} - \frac{\sum_{l \in \sigma^C} w_l (b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)}.$$

From Eq. (6.38) it follows that $p_0 = 0$ is required. Then, the inequalities in Eq. (6.37) require that

$$\begin{aligned} 0 &< (1-s)(a_{|\sigma|-1} - l) + \sum_{j \in \sigma^D} w_j (b_{|\sigma|} - a_{|\sigma|-1}) \\ &\xRightarrow{(1-s)>0} a_{|\sigma|-1} + \frac{\sum_{j \in \sigma^D} w_j (b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} > l. \end{aligned} \tag{6.44}$$

This completes the proof. □

Remark 10 (The prisoner's dilemma)**.** *For $n = 2$ the full weight is placed on the single opponent i.e., $\hat{w}_j = 1$. When the payoff parameters are defined as $b_1 = T$, $b_0 = P$, $a_1 = R$, $a_0 = S$, the result in Theorem 8 recovers the result obtained for the finitely repeated 2-player game in [121].*

Theorem 8 does not stipulate any conditions on the key player's initial probability to cooperate other than $p_0 \in [0, 1]$. However, the existence of extortionate and generous strategies does depend on the value of p_0. This is formalized in the following proposition.

Proposition 5 (The influence of the initial probability to cooperate)**.** *For the existence of extortionate strategies it is necessary that $p_0 = 0$. Moreover, for the existence of generous ZD strategies it is necessary that $p_0 = 1$.*

Proof. For brevity, in the following proof we refer to equations that are found in the proof of Proposition 4. Assume the ZD strategy is extortionate, hence $l = b_0$. From the lower bound in Eq. (6.16) in order for l to be enforceable, it is necessary that $p_0 = 0$. This proves the first statement. Now assume the ZD strategy is generous, hence $l = a_{n-1}$. From the lower bound in Eq. (6.15) in order for l to be enforceable, it is necessary that $p_0 = 1$. This proves the second statement and completes the proof. □

These requirements on the key player's initial probability to cooperate make intuitive sense. In a finitely repeated game, if the key player aims to be an *extortioner* that profits from the cooperative actions of others, she cannot start to cooperate because she could be taken advantage of by defectors. On the other hand, if she aims to be *generous*, she cannot start as a defector because this will punish both cooperating and defecting co-players.

6.5 Applications

In the following, we will apply the theory developed in this chapter to three n-player social dilemmas: the n-player linear public goods game, the n-player snowdrift game, and the n-player stag-hunt game. For simplicity the following assumption is made.

Assumption 8 (Equal weights)**.** *The ZD strategists puts equal weight on each co-player, such that $w_j = \frac{1}{n-1}$ for all $j \neq i$.*

Under this assumption, we will derive explicit conditions on the group size n, and the payoff parameters of the n-player social dilemmas under which generous, extortionate, and equalizer strategies exist. Detailed proofs are provided to show how the results are obtained, and numerical examples are used to illustrate the implications of the theory under a variety of circumstances.

6.5.1 n-player linear public goods games

In the n-player linear public goods game, cooperators contribute an amount $c > 0$ to a publicly available good that grows linearly with the number of cooperators [109, 124–126]. The sum of the contributions is scaled by a public goods multiplier $1 < r < n$ and then distributed evenly among all players. For cooperators, this results in one-shot payoffs $a_z = \frac{rc(z+1)}{n} - c$ and defectors receive $b_z = \frac{rcz}{n}$. The following Lemma characterizes the bounds on the baseline payoffs.

Lemma 10. *For the public goods game the enforceable baseline payoffs are determined by*

$$\max\left\{0, \frac{rc(n-1)}{n} - \frac{c}{1-s}\right\} \leq l, \tag{6.45}$$

$$\min\left\{\frac{rc}{n} - c + \frac{c}{1-s}, rc - c\right\} \geq l, \tag{6.46}$$

with at least one strict inequality.

Proof. The bounds are obtained by substituting the single-round payoffs $a_z = \frac{rc(z+1)}{n} - c$ and $b_z = \frac{rcz}{n}$ into the inequalities of Theorem 8 and use the fact that the bounds are linear in z. The obvious details are omitted. □

Proposition 6 (Extortion in public goods games). *Suppose* $p_0 = 0$, $l = 0$ *and* $0 < s < 1$. *For a public goods game with* $r > 1$, *every slope* $s \geq \frac{r-1}{r}$ *can be enforced independent of* n. *If* $s < \frac{r-1}{r}$, *the slope can be enforced if and only if*

$$n \leq \frac{r(1-s)}{r(1-s)-1}.$$

Proof. For extortionate strategies $l = 0$ and $0 < s < 1$. The inequalities Eq. (6.45) and Eq. (6.46) in Lemma 10 become

$$\max\left\{0, \frac{rc(n-1)}{n} - \frac{c}{1-s}\right\} \leq 0 \tag{6.47}$$

$$\min\left\{\frac{rc}{n} - c + \frac{c}{1-s}, rc - c\right\} \geq 0 \tag{6.48}$$

Solving for s will yield the enforceable slopes in the extortionate ZD strategy. Observe that a necessary condition for Eq. (6.47) to hold is that the left-hand side is equal 0 and in order for this to hold it is required that

$$\frac{rc(n-1)}{n} - \frac{c}{1-s} \leq 0 \Leftrightarrow rc(n-1) - n\frac{c}{1-s} \leq 0. \tag{6.49}$$

Equivalently,

$$n(r - \frac{1}{1-s}) \leq r \Leftrightarrow n(r(1-s)-1) \leq r(1-s). \tag{6.50}$$

The conditions $-\frac{1}{n-1} < s < 1$ in Theorem 8 and the assumption that r is positive implies that $r(1-s)$ in the right-hand side of Eq. (6.50) is required to be strictly positive. It follows that if $r(1-s) - 1 \leq 0$ the inequalities in Eq. (6.49) are always satisfied. To obtain the criteria on the slope s we may write,

$$r(1-s) - 1 \leq 0 \Leftrightarrow -rs \leq 1 - r \Leftrightarrow s \geq \frac{r-1}{r}. \tag{6.51}$$

Note that if $s \geq \frac{r-1}{r}$ is satisfied, the left-hand side of the inequality Eq. (6.48) reads as $rc - c$. The requirement $0 \leq rc - c$ leads to $r \geq 1$, which is very natural and satisfied for the payoff of the public goods game. It follows that for every $r > 1$, every $s \geq \frac{r-1}{r}$ can be enforced independent of n. Due to the requirement that at least one of the inequalities needs to be strict it follows that for the special case $r = 1$ it must hold that $s > 0$.

On the other hand, when $s < \frac{r-1}{r}$ in order for Eq. (6.49) to be satisfied it must hold that

$$n \leq \frac{r(1-s)}{r(1-s)-1}. \tag{6.52}$$

Note that $s < \frac{r-1}{r}$ implies $r(1-s) - 1 \neq 0$ so the above inequality is well-defined. If Eq. (6.52) does not hold and $s < \frac{r-1}{r}$ than

$$\frac{rc(n-1)}{n} - \frac{c}{1-s} > 0, \tag{6.53}$$

thus the lower-bound in Eq. (6.47) is not satisfied and consequently there cannot exist extortionate strategies. We now investigate the inequality Eq. (6.48). We already know that when $s \geq \frac{r-1}{r}$ the upper-bound reads as $0 < rc - c$ and is satisfied for any $r > 1$. On the other hand, the left-hand side of Eq. (6.48) is equal to $\frac{rc}{n} - c + \frac{c}{1-s}$ if

$$\frac{rc}{n} - c + \frac{c}{1-s} \leq rc - c \Leftrightarrow \qquad n[(1-s)r - 1] \geq r(1-s).$$

Because $r(1-s) > 0$, these inequalities can only be satisfied if $s < \frac{r-1}{r}$ and

$$n \geq \frac{r(1-s)}{r(1-s)-1}. \tag{6.54}$$

Under these conditions, the only possibility for an enforceable payoff relation is the equality case in which $n = \frac{r(1-s)}{(1-s)r-1}$, otherwise the lower-bound is not satisfied and there cannot exist extortionate strategies.

Finally, we check the necessary condition for the existence of solutions of Eq. (6.47) and Eq. (6.48) that the lower-bound cannot exceed the upper-bound. We already know that when $s \geq \frac{r-1}{r}$ the lower and upper-bound read as $0 \leq 0 \leq rc - c$ and is satisfied for any $r > 1$. When $s < \frac{r-1}{r}$ for existence, n cannot exceed $\frac{r(1-s)}{r(1-s)-1}$. When equality holds note that we have

$$\text{if } n = \frac{r(1-s)}{r(1-s)-1} \text{ and } s < \frac{r-1}{r}:$$
$$0 = \frac{rc(n-1)}{n} - \frac{c}{1-s} \leq 0 \leq \frac{rc}{n} - c + \frac{c}{1-s} = rc - c,$$

which is satisfied with a strict upper-bound if $r > 1$. We conclude that the lower-bound never exceeds the upper-bound and this condition does not limit the existence of extortionate ZD strategies in the public goods game. This completes the proof. □

We now continue to characterize the generous strategies in linear public goods games.

Proposition 7 (Generosity in public goods games)**.** *Suppose* $p_0 = 1$, $l = rc - c$ *and* $0 < s < 1$. *For a public goods game with* $1 < r < n$, *the region of enforceable slopes of generous strategies is equivalent to the region of the enforceable slopes for extortionate strategies.*

Proof. For generous strategies $l = rc - c$ and $0 < s < 1$, the inequalities Eq. (6.45) and Eq. (6.46) in Lemma 10 become

$$\max\left\{0, \frac{rc(n-1)}{n} - \frac{c}{1-s}\right\} \leq rc - c, \tag{6.55}$$

$$\min\left\{\frac{rc}{n} - c + \frac{c}{1-s}, rc - c\right\} \geq rc - c. \tag{6.56}$$

Clearly in order for generous strategies to exist it is necessary that the left-hand side of Eq. (6.56) reads as $rc - c$. Therefor it is required that

$$\frac{rc}{n} - c + \frac{c}{1-s} \geq rc - c \Leftrightarrow n(r(1-s) - 1) \leq (1-s)r.$$

Hence, this condition is equivalent to the condition in Eq. (6.50) and thus this condition gives the same feasible region for the existence of extortionate strategies. Now suppose that, $s < \frac{r-1}{r}$ and $n \geq \frac{r(1-s)}{r(1-s)-1}$. Also in this case, only equality is possible i.e. $n = \frac{r(1-s)}{r(1-s)-1}$ because otherwise the upper-bound is not satisfied. Next to this, if $s < \frac{r-1}{r}$ and $n = \frac{r(1-s)}{r(1-s)-1}$ in order for the lower-bound to be satisfied it is required that

$$rc - c = \frac{rc}{n} - c + \frac{c}{1-s} \geq rc - c \geq 0,$$

which is satisfied with a strict lower-bound for any $r > 1$. We conclude that, in the linear public goods game, the region of feasible slopes for generous strategies is equivalent to the region of feasible sloped for extortionate strategies. This completes the proof. □

Proposition 8 (Equalizers in public goods games)**.** *Suppose $s = 0$. For a public goods game with $1 < r < n$, if $n \leq \frac{r}{r-1}$ an equalizer strategy can enforce any baseline payoff $0 \leq l \leq rc - c$. If $\frac{r}{r-1} < n < \frac{2r}{r-1}$ the equalizer strategy can enforce $\frac{rc(n-1)}{n} - c \leq l \leq \frac{rc}{n}$. If $n \geq \frac{2r}{r-1}$ no equalizer strategies exist.*

Proof. Suppose $s = 0$ such that the ZD strategy is an equalizer. Then equation Eq. (6.45) and Eq. (6.46) of Lemma 10 become

$$\max\left\{0, \frac{rc(n-1)}{n} - c\right\} \leq l \leq \min\left\{rc - c, \frac{rc}{n}\right\}. \tag{6.57}$$

Solving for l yield the baseline payoffs that an equalizer strategy can enforce. We first investigate the conditions under which the entire range of baseline payoffs can

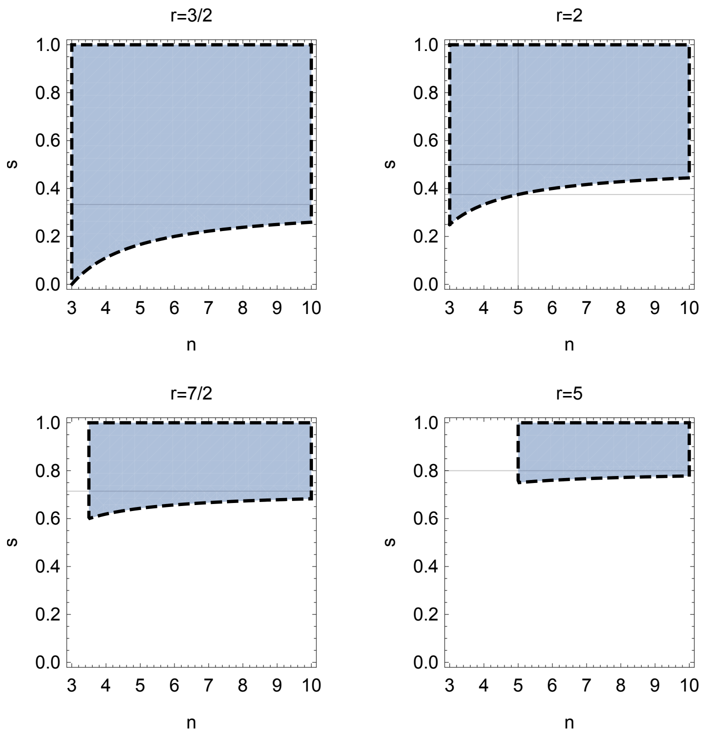


Figure 6.1: Numerical examples of enforceable slopes for extortionate and generous strategies in n-player linear public goods games. Observe that when n increases, the range of enforceable slopes decreases according to the condition on n in Proposition 6 that implies that for larger groups the slope must satisfy $s \geq 1 - \frac{n}{r(n-1)}$. Also, when r increases, the set of slopes that can be enforced independent of n decreases according to the condition $s \geq \frac{r}{r-1}$. One can also see that the requirement $r < n$, shifts the feasible region as r increases.

be enforced by the equalizer strategy. Note that the left-hand side of the inequality Eq. (6.57) is equal to zero if and only if

$$\frac{rc(n-1)}{n} - c \leq 0 \Leftrightarrow n \leq \frac{r}{r-1} \Leftrightarrow r \leq \frac{n}{n-1}.$$

In this case, the upper-bound of Eq. (6.57) is equal to $rc - c$. It follows that when $n \leq \frac{r}{r-1}$ or equivalently $r \leq \frac{n}{n-1}$, then Eq. (6.57) becomes

$$\text{if } n \leq \frac{r}{r-1}: \quad 0 \leq l \leq rc - c, \tag{6.58}$$

in other words, almost the entire range (remember one inequality is necessarily strict) of possible payoffs can be enforced by the equalizer strategy. In the case that $n > \frac{r}{r-1}$ Eq. (6.57) becomes

$$\text{if } n > \frac{r}{r-1}: \quad \frac{rc(n-1)}{n} - c \leq l \leq \frac{rc}{n}. \tag{6.59}$$

Thus when n increases and r is fixed, an equalizer strategy can enforce a smaller range of baseline payoffs. Finally, it must be noted that in the case of Eq. (6.59) it is possible that the lower-bound is equal or larger than the upper-bound. In this case no equalizer strategies can exist. To obtain a condition we may write

$$\frac{rc(n-1)}{n} - c \geq \frac{rc}{n} \Leftrightarrow n \geq \frac{2r}{r-1} \Leftrightarrow r \geq \frac{n}{n-2}.$$

It follows that for $n \geq \frac{2r}{r-1}$ no equalizer strategies exist. Finally, we can conclude that within the range $\frac{r}{r-1} < n < \frac{2r}{r-1}$ the enforceable baseline payoffs for the equalizer strategy are

$$\frac{rc(n-1)}{n} - c \leq l \leq \frac{rc}{n},$$

with at least one strict inequality that is implied by the strict bounds on n. This completes the proof. □

Remark 11 (Enforcing the mutual cooperation payoff in public goods games)**.** *A particularly interesting implication on the bounds of the equalizer strategy is that whenever $r > 1$ and $n \leq \frac{r}{r-1}$, then the equalizer ZD strategist can enforce the mutual cooperation payoff, e.g. $\pi^{-i} = rc - c$. This also holds in the extreme case in which all co-players of the ZD strategist employ the ALLD strategy and thus always defect. In this special case, only the outcomes $(C, 0)$ and $(D, 0)$ can occur with a positive probability. Because all co-players employ the same strategy and payoffs are symmetric, all co-players receive the same payoff that depends on the chosen action of the strategic player, namely: $b_1 = \frac{rc}{n}$ if the ZD strategist cooperates, and $b_0 = 0$ otherwise. Because $r > 1$, the condition $n \leq \frac{r}{r-1}$ may be written as $r \leq \frac{n}{n-1}$. In this case, the highest possible public goods multiplier is $r = \frac{n}{n-1}$. By substituting this into the payoffs we obtain*

$$a_{n-1} = rc - c = \frac{n}{n-1}c - c = \frac{c}{n-1},$$

and

$$b_1 = \frac{rc}{n} = \frac{n}{(n-1)n}c = a_{n-1}.$$

Thus, under these conditions, the ZD strategist will enforce the mutual cooperation payoff to the ALLD co-players by always cooperating. Indeed, one can define the

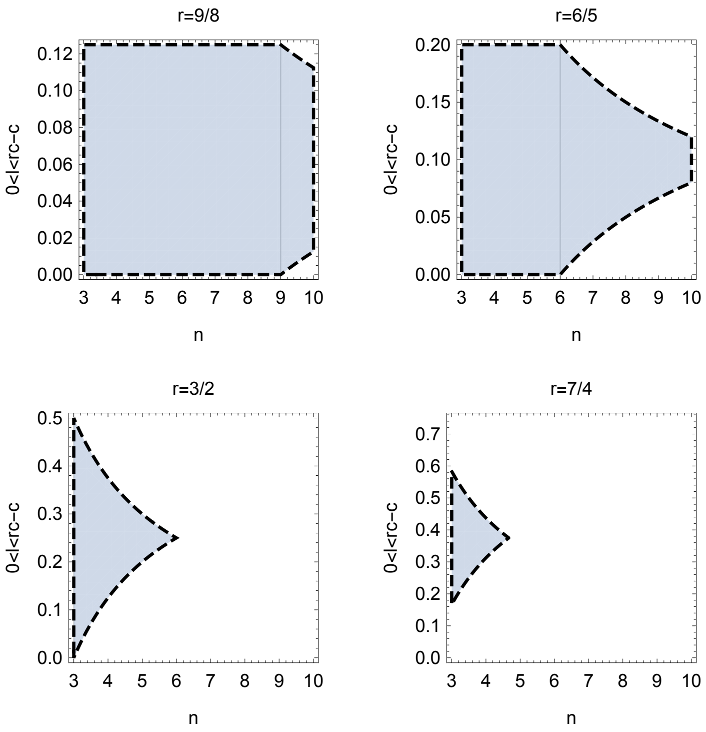


Figure 6.2: Numerical examples of the bounds on the baseline payoff for equalizer strategies in n-player linear public goods games. When n or r increases the feasible region becomes smaller. It can be observed that the entire range of baseline payoffs can be enforced if the group size is sufficiently small, e.g. $n \leq \frac{r}{r-1}$, see Proposition 8. Once this inequality is not satisfied anymore, the region of enforceable baseline payoffs shrinks according to $\frac{rc(n-1)}{n} - c \leq l \leq \frac{rc}{n}$. Note that the payoffs are obtained for $c = 1$, the payoffs can be scaled for higher values of c without affecting the result.

equalizer ZD strategy by setting $l = a_{n-1}$, $s = 0$ *and* $\phi = \frac{\delta}{(1-s)\left(\frac{rc}{n}-c\right)+c}$. *Then,* Eq. (6.8) *implies*

$$\delta \mathbf{p} = \delta \mathbb{1} \Rightarrow \mathbf{p} = \mathbb{1}.$$

How to exactly choose the parameter ϕ depending on δ and the payoff parameters is one of the topics in the next chapter (see Remark 13 as well).

6.5.2 n-player snowdrift games

The n-player snowdrift game traditionally describes a situation in which cooperators need to clear out a snowdrift so that everyone can go on their merry way. By clearing out the snowdrift together, cooperators share a cost c required to create a fixed benefit b [109, 127–129]. If a player cooperates together with z group members, their one-shot payoff is $a_z = b - \frac{c}{z+1}$. If there is at least one cooperator ($z > 0$) who clears out the snowdrift, then defectors obtain a benefit $b_z = b$. If no one cooperates, the snowdrift will not be cleared and everyone's payoff is $b_0 = 0$.

Lemma 11. *For the n-player snowdrift game the enforceable baseline payoffs l are determined as*

$$\max\left\{0, b - \frac{c}{(n-1)(1-s)}\right\} \leq l \leq b - \frac{c}{n}, \tag{6.60}$$

with at least one strict inequality.

Proof. Suppose $z = 0$, then the inequalities in Theorem 8 on the baseline payoff become

$$0 \leq l \leq b - c + \frac{c}{1-s}. \tag{6.61}$$

And if $1 \leq z \leq n-1$, the bounds on the enforceable baseline payoffs read as

$$l \geq b - \frac{c}{(n-1)(1-s)}, \tag{6.62}$$

$$l \leq \min_{1 \leq z \leq n-1} \left\{ b - \frac{c}{z+1} + \frac{n-z-1}{n-1} \frac{c}{(z+1)(1-s)} \right\}. \tag{6.63}$$

We continue to investigate the minimum upper-bound of l. After some basic manipulation we find that upper-bound in Eq. (6.63) can be written as

$$l \leq \min_{1 \leq z \leq n-1} b + \underbrace{\frac{((n-1)s+1)\,c}{(n-1)(z+1)(1-s)}}_{:=\xi(z)} - \frac{c}{(n-1)(1-s)}. \tag{6.64}$$

From Theorem 8, in order for a ZD strategy to exist it is necessary that $s < 1$ and because in a multiplayer game $n > 1$, the denominator of the fraction $\xi(z)$ is positive for any $0 \leq z \leq n-1$. Thus, if the numerator of $\xi(z)$ is positive as well, then the minimum of the upper-bound occurs when z is maximum. Now because $c > 0$ we have,

$$[(n-1)s + 1]c > 0 \Leftrightarrow (n-1)s + 1 > 0 \Leftrightarrow s > -\frac{1}{n-1}.$$

It follows from the bounds of enforceable slopes s in Theorem 8, that $\xi(z)$ is required to be positive, otherwise no ZD strategies can exist. Hence, for $1 \leq z \leq n-1$ and enforceable slope $-\frac{1}{n-1} < s < 1$, the minimum of the upper-bound occurs when all co-players are cooperating, i.e., $z = n-1$. In combination with the upper-bound in Eq. (6.61) for the case $z = 0$ we obtain $l \leq \min\{b - \frac{c}{n}, b - c + \frac{c}{1-s}\}$. Note that

$$b - \frac{c}{n} < b - c + \frac{c}{(1-s)} \Leftrightarrow s > \frac{1}{1-n} \Leftrightarrow s > -\frac{1}{n-1}.$$

Hence, for any enforceable slope $-\frac{1}{n-1} < s < 1$ we obtain

$$l \leq b - \frac{c}{n}.$$

In summary, for the n-player snowdrift game the enforceable base level payoffs l are determined as

$$\max\left\{0, b - \frac{c}{(n-1)(1-s)}\right\} \leq l \leq b - \frac{c}{n},$$

with at least one strict inequality. This completes the proof. □

Proposition 9 (Extortion in n-player snowdrift games)**.** *Suppose* $p_0 = 0$, $l = 0$ *and* $0 < s < 1$. *For the n-player snowdrift game with* $b > c > 0$, *extortionate strategies can enforce any* $s \geq 1 - \frac{c}{b(n-1)}$. *For high benefit-to-cost* $\frac{b}{c} > \frac{1}{(n-1)(1-s)}$ *no extortionate strategies exist.*

Proof. Suppose $l = 0$ and $0 < s < 1$, such that the strategy is extortionate. In this case, Eq. (6.60) in Lemma 11 becomes

$$\max\left\{0, b - \frac{c}{(n-1)(1-s)}\right\} \leq 0 \leq b - \frac{c}{n}. \tag{6.65}$$

For any $b > c$ the upper-bound is satisfied. In order for the lower-bound to be satisfied it is required to hold that

$$b - \frac{c}{(n-1)(1-s)} \leq 0 \Leftrightarrow \frac{b}{c} \leq \frac{1}{(n-1)(1-s)}, \tag{6.66}$$

or equivalently, $s \geq 1 - \frac{c}{b(n-1)}$. Clearly, for smaller slopes $s < 1 - \frac{c}{b(n-1)}$ no extortionate strategies can exist. Finally, because the lower-bound cannot exceed the upper-bound as long as $s > -\frac{1}{n-1}$ we conclude that extortionate strategies with slopes $s \geq 1 - \frac{c}{b(n-1)}$ can exist in the n-player snowdrift game. □

Proposition 10 (Generosity in n-player snowdrift games)**.** *Suppose* $p_0 = 1$, $l = b - \frac{c}{n}$ *and* $0 < s < 1$. *For the n-player snowdrift game with* $b > c$, *generous strategies can enforce any* $0 < s < 1$ *independent of* n.

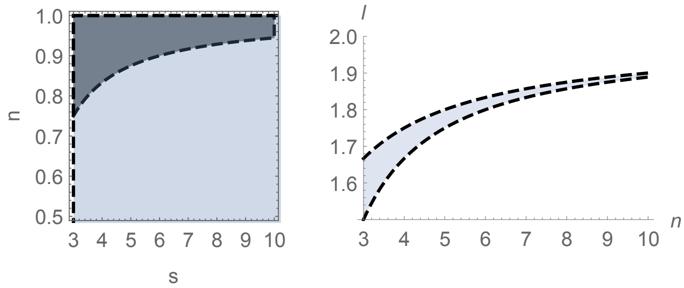


Figure 6.3: Enforceable slopes in the n-player snowdrift game for extortionate (left, dark area) and generous ZD strategies (left, light area) with $b = 2$ and $c = 1$. In this game, extortionate strategies only exist when the slope is sufficiently high, see Proposition 9, in this numerical example $s \geq 1 - \frac{1}{2(n-1)}$. In contrast, generous strategies can enforce any slope $0 < s < 1$, see Proposition 10. However, the desired slope will affect the minimum number of rounds necessary to enforce the linear payoff relation. As in Proposition 11 equalizer strategies can enforce a limited range of baseline payoffs that becomes smaller when the group size increases (right).

Proof. Now suppose $l = b - \frac{c}{n}$ and $0 < s < 1$. In this case, Eq. (6.60) in Lemma 11 becomes

$$\max\left\{0, b - \frac{c}{(n-1)(1-s)}\right\} \leq b - \frac{c}{n} \leq b - \frac{c}{n}. \tag{6.67}$$

Clearly, for any $b > c > 0$, $n > 0$ these inequalities are satisfied for any $s > -\frac{1}{n-1}$. And hence, generous strategies always exist in the n-player snowdrift game. This completes the proof. □

Proposition 11 (Equalizers in n-player snowdrift games)**.** *Suppose $s = 0$. For the n-player snowdrift game with $b > c > 0$ the enforceable baseline payoffs for equalizer strategies are $b - \frac{c}{n-1} \leq l \leq b - \frac{c}{n}$.*

Proof. Suppose $s = 0$. We solve for the range of enforceable base-level payoffs. In this case, Eq. (6.60) in Lemma 11 becomes

$$\max\left\{0, b - \frac{c}{(n-1)}\right\} \leq l \leq b - \frac{c}{n}, \tag{6.68}$$

Clearly, for $b > c > 0$ and $n > 1$ the lower-bounds reads as $b - \frac{c}{(n-1)}$ and the lower-bound cannot exceed the upper-bound. This completes the proof. □

6.5.3 n-player stag hunt games

In the public goods and the n-player snowdrift game, a single player can create a benefit. In some social dilemmas a single cooperator is not sufficient to create a benefit. In the n-player stag hunt game players obtain a benefit b if only if all players cooperate [109]. This results in the payoffs,

$$b_z = 0, \text{ for all } 0 \leq z \leq n-1,$$

$$a_z = \begin{cases} b-c, & \text{if } z = n-1; \\ -c, & \text{otherwise.} \end{cases}$$

Lemma 12. *For the n-player stag hunt game the enforceable baseline payoffs are determined by*

$$0 \leq l \leq \min\left\{\frac{c}{1-s}\frac{1}{n-1} - c, b-c\right\}.$$

Proof. By substituting the expressions for the single round payoff of the n-player stag hunt game into the lower bound on l in Theorem 8 we obtain,

$$\max_{0 \leq z \leq n-1} \left\{-\frac{z}{n-1}\frac{c}{1-s}\right\} \leq l \tag{6.69}$$

Because $c > 0$ and $1-s > 0$, it follows that the maximum lower bound is 0. Now assume $0 \leq z \leq n-2$, the upper bound on the baseline payoff in Theorem 8 reads as

$$\min_{0 \leq z \leq n-1} \left\{-c + \frac{n-z-1}{n-1}\frac{c}{1-s}\right\} = -c + \frac{1}{n-1}\frac{c}{1-s} \tag{6.70}$$

Now suppose $z = n-1$, then the upper bound reads as $b - c > 0$. This completes the proof. □

From Lemma 12, we can immediately observe that there do not exist equalizer strategies in the n-player stag hunt game. Namely, by substituting $s = 0$ into the bounds of Lemma 10 one arrives at a contradiction because $b > c$ and $n > 1$. However, the following propositions show that extortionate and generous strategies do exist.

Proposition 12 (Extortion in n-player stag-hunt games)**.** *Suppose $p_0 = l = 0$, and $0 < s < 1$. For the n-player stag hunt game with $b > c$, extortionate strategies can enforce any slope $s \geq 1 - \frac{c}{(n-1)b}$ independent of the group size $n > 2$. For smaller slopes $s < 1 - \frac{c}{b(n-1)}$ it needs to hold that $n < \frac{2-s}{1-s}$.*

Proof. Assume $l = b_0 = 0$, the bounds on the baseline payoff in Lemma 12 become

$$0 \leq 0 \leq \min\left\{\frac{c}{1-s}\frac{1}{n-1} - c, b-c\right\}. \tag{6.71}$$

By assumption $b > c > 0$, hence if

$$\frac{c}{1-s}\frac{1}{n-1} - c \geq b - c > 0 \Rightarrow s \geq 1 - \frac{c}{(n-1)b},$$

then the bounds in Eq. (6.71) are satisfied with a strict lower bound. Alternatively, if $s < 1 - \frac{c}{(n-1)b}$, then the bounds are satisfied with one strict inequality if and only if

$$\frac{c}{1-s}\frac{1}{n-1} - c > 0 \Rightarrow n < \frac{2-s}{1-s}.$$

This completes the proof. □

Proposition 13 (Generosity in n-player stag-hunt games)**.** *Suppose $p_0 = 1$, $l = b - c$ and $0 < s < 1$. For the n-player stag hunt game with $b > c$, generous strategies can enforce any slope $s \geq 1 - \frac{c}{b(n-1)}$. Smaller slopes $s < 1 - \frac{c}{b(n-1)}$ cannot be enforced.*

Proof. Assume $l = a_{n-1} = b - c$, the bounds on the baseline payoff in Lemma 12 become

$$0 \leq b - c \leq \min\left\{\frac{c}{1-s}\frac{1}{n-1} - c, b - c\right\}, \tag{6.72}$$

clearly for this upper bound to hold it is required that

$$b - c \leq \frac{c}{1-s}\frac{1}{n-1} - c \Rightarrow s \geq 1 - \frac{c}{(n-1)b}.$$

This completes the proof. □

Remark 12. *Interestingly, the enforceable slopes of generous strategies in the n-player stag hunt game coincide with the enforceable slopes of extortionate strategies in n-player snowdrift games.*

6.6 Final Remarks

We have characterized the enforceable payoff relation in finitely repeated n-player social dilemma games. Even though the single-round payoffs of the players are symmetric, it turns out that a single player can exert a significant level of control on their co-players in a variety of social dilemmas. Naturally, exerting this control requires repeated interactions. In the next chapter we will investigate how “fast” a ZD strategist can enforce some desired payoff relation.

CHAPTER 7

The efficiency of exerting control in multi-player social dilemmas

However beautiful the strategy, you should occasionally look at the result.

Winston Churchill

In the previous chapter we characterized the enforceable payoff relations of ZD-strategies in repeated n-player social dilemma games with a finite but undetermined number of rounds. The obtained conditions generalize those for two-player games and illustrate how a single player can exert control over the outcome of an n-player repeated game with discounted payoffs. However, the conditions that result from the existence problem do not specify requirements on the discount factor other than $\delta \in (0, 1)$. One could be interested in how many expected rounds a ZD strategists would require to enforce some desired payoff relation. In this chapter, we will address exactly this "efficiency" problem.

Problem 2 (The minimum threshold problem)**.** *Suppose the desired payoff relation* $(s, l) \in \mathbb{R}^2$ *satisfies the conditions in Theorem 8. What is the minimum* $\delta \in (0, 1)$ *under which the linear relation* (s, l) *with weights* w *can be enforced by the ZD strategist?*

Because δ determines the expected number of rounds, solutions to this problem also provide insight into one's possibilities for exerting control given a constraint on the expected number of interactions. We will consider the three enforceable classes of ZD-strategies in n-player social dilemmas separately. Before giving the main results it is necessary to introduce some additional notation. Define $\tilde{w}_z = \max_{w_h \in w} \sum_{h=1}^{z} w_h$ to be the maximum sum of weights for some permutation of $\sigma \in \boldsymbol{\mathcal{A}}$ with z cooperating co-players. Additionally, for some given payoff relation $(s, l) \in \mathbb{R}^2$ and $w \in \mathbb{R}^{n-1}$ define

$$\begin{aligned}
\overline{\rho}^C &:= \max_{0 \le z \le n-1} (1-s)(a_z - l) + \tilde{w}_{n-z-1}(b_{z+1} - a_z), \\
\underline{\rho}^C &:= \min_{0 \le z \le n-1} (1-s)(a_z - l) + \hat{w}_{n-z-1}(b_{z+1} - a_z), \\
\overline{\rho}^D &:= \max_{0 \le z \le n-1} (1-s)(l - b_z) + \tilde{w}_z(b_z - a_{z-1}), \\
\underline{\rho}^D &:= \min_{0 \le z \le n-1} (1-s)(l - b_z) + \hat{w}_z(b_z - a_{z-1}).
\end{aligned} \tag{7.1}$$

In the following, we will use these extrema to derive threshold discount factors for extortionate, generous and equalizer strategies in symmetric n-player social dilemma games.

7.0.1 Extortionate ZD-strategies

We first consider the case in which $l = b_0$ and $0 < s < 1$, such that the ZD-strategy is extortionate.

Theorem 9 (Thresholds for extortion)**.** *Assume that $p_0 = 0$ and $(s, b_0) \in \mathbb{R}^2$ satisfy the conditions in Theorem 8, then $\overline{\rho}^C > 0$ and $\overline{\rho}^D + \underline{\rho}^C > 0$. Moreover, the threshold discount factor above which extortionate ZD-strategies exist is determined by*

$$\delta_\tau = max\left\{\frac{\overline{\rho}^C - \underline{\rho}^C}{\overline{\rho}^C}, \frac{\overline{\rho}^D}{\overline{\rho}^D + \underline{\rho}^C}\right\}.$$

Proof. For brevity in the following proof we refer to equations that can be found in the proof of Theorem 8, in Chapter 6. From Proposition 5 we know that in order for the extortionate payoff relation to be enforceable it is necessary that $p_0 = 0$. By substituting this into Eq. (6.37) it follows that in order for the payoff relation to be enforceable it is required that for all σ such that $x_i = C$ the following holds:

$$\rho^C(\sigma) := (1-s)(a_{|\sigma|-1} - l) + \sum_{j \in \sigma^D} w_j (b_{|\sigma| - a_{|\sigma|-1}}) > 0. \tag{7.2}$$

Hence, Eq. (6.37) with $p_0 = 0$ implies that for all σ such that $x_i = C$ it holds that

$$\frac{1-\delta}{\rho^C(\sigma)} \le \phi \le \frac{1}{\rho^C(\sigma)} \Rightarrow \frac{1-\delta}{\underline{\rho}^C(z, \hat{w}_z)} \le \phi \le \frac{1}{\overline{\rho}^C(z, \tilde{w}_z)}. \tag{7.3}$$

Naturally, it holds that $\overline{\rho}^C \geq \underline{\rho}^C$. In the special case in which equality holds, it follows from equation Eq. (7.3) that $\delta \geq 0$, which is true by definition of δ. We continue to investigate the case $\overline{\rho}^C > \underline{\rho}^C$. In this case, a solution to Eq. (7.3) for some $\phi > 0$ exists if and only if

$$\frac{1-\delta}{\underline{\rho}^C(z,\hat{w}_z)} \leq \frac{1}{\overline{\rho}^C(z,\tilde{w}_z)} \Rightarrow \delta \geq \frac{\overline{\rho}^C - \underline{\rho}^C}{\overline{\rho}^C}, \tag{7.4}$$

which leads to the first expression in the theorem. Now, from Eq. (6.38) with $p_0 = 0$, it follows that in order for the payoff relation to be enforceable it is necessary that

$$\forall \sigma \text{ s.t. } x_i = D: \quad 0 \leq \phi \rho^D(\sigma) \leq \delta \Rightarrow 0 \leq \phi \overline{\rho}^D(z,\tilde{w}_z) \leq \delta. \tag{7.5}$$

Because $\phi > 0$ is necessary for the payoff relation to be enforceable, it follows that $\rho^D(\sigma) \geq 0$ for all σ such that $x_i = D$. Let us first investigate the special case in which $\overline{\rho}^D(z,\tilde{w}_z) = 0$. Then Eq. (7.5) is satisfied for any $\phi > 0$ and $\delta \in (0,1)$. Now, assume $\overline{\rho}^D(z,\tilde{w}_z) > 0$. Then, Eq. (7.5) and Eq. (7.3) imply

$$\frac{1-\delta}{\underline{\rho}^C(z,\hat{w}_z)} \leq \phi \leq \frac{\delta}{\overline{\rho}^D(z,\tilde{w}_z)}. \tag{7.6}$$

In order for such a ϕ to exist it needs to hold that

$$\frac{1-\delta}{\underline{\rho}^C(z,\hat{w}_z)} \leq \frac{\delta}{\overline{\rho}^D(z,\tilde{w}_z)} \xrightarrow{\overline{\rho}^D,\ \underline{\rho}^C > 0} \delta \geq \frac{\overline{\rho}^D}{\overline{\rho}^D + \underline{\rho}^C}. \tag{7.7}$$

This completes the proof. □

7.0.2 Generous ZD-strategies

If a player instead aims to be generous, in general, different thresholds will apply. Thus, let us now consider the case in which $l = a_{n-1}$ and $0 < s < 1$ such that the ZD-strategy is generous.

***Theorem* 10** (Thresholds for generosity)**.** *Assume that $p_0 = 1$ and $(s, a_{n-1}) \in \mathbb{R}^2$ satisfy the conditions in Theorem 8. Then $\overline{\rho}^D > 0$ and $\overline{\rho}^C + \underline{\rho}^D > 0$. Moreover, the threshold discount factor above which generous ZD-strategies exist is determined by*

$$\delta_\tau = max\left\{\frac{\overline{\rho}^D - \underline{\rho}^D}{\overline{\rho}^D}, \frac{\overline{\rho}^C}{\overline{\rho}^C + \underline{\rho}^D}\right\}.$$

Proof. The proof is similar to the extortionate case in the proof of Theorem 9. From Proposition 5 we know that in order for the generous payoff relation to be enforceable it is necessary that $p_0 = 1$. By substituting this into Eq. (6.38) it follows that in order for the payoff relation to be enforceable it is required that for all σ such that $x_i = D$ the following holds:

$$\rho^D(\sigma) = (1-s)(l - b_{|\sigma|}) + \sum_{j \in \sigma^C} w_j (b_{|\sigma| - a_{|\sigma|-1}}) > 0. \tag{7.8}$$

Hence, Eq. (6.38) with $p_0 = 1$ implies that for all σ such that $x_i = D$ it holds that

$$\frac{1-\delta}{\rho^D(\sigma)} \leq \phi \leq \frac{1}{\rho^D(\sigma)} \Rightarrow \frac{1-\delta}{\underline{\rho}^D(z, \hat{w}_z)} \leq \phi \leq \frac{1}{\overline{\rho}^D(z, \tilde{w}_z)}. \tag{7.9}$$

If $\overline{\rho}^D = \underline{\rho}^D > 0$ this implies $\delta \geq 0$. Otherwise Eq. (7.9) implies that

$$\frac{1-\delta}{\underline{\rho}^D(z, \hat{w}_z)} \leq \frac{1}{\overline{\rho}^D(z, \tilde{w}_z)} \Rightarrow \delta \geq \frac{\overline{\rho}^D - \underline{\rho}^D}{\overline{\rho}^D}, \tag{7.10}$$

which leads to the first expression in the theorem. Moreover, from Eq. (6.37) we know that the following must hold:

$$\forall \sigma \text{ s.t. } x_i = C: \quad 0 \leq \phi \rho^C(\sigma) \leq \delta \Rightarrow 0 \leq \phi \overline{\rho}^C(z, \tilde{w}_z) \leq \delta. \tag{7.11}$$

Because $\phi > 0$ it follows that $\rho^C(\sigma) \geq 0$ for all σ such that $x_i = C$. Let us now consider the special case in which $\overline{\phi} \rho^C(z, \tilde{w}_z) = 0$. Then, Eq. (7.11) is satisfied for any $\phi > 0$ and $\delta \in (0,1)$. Now suppose $\overline{\rho}^C(z, \tilde{w}_z) > 0$. Then, Eq. (7.11) and Eq. (7.9) imply that in order for the generous strategy to be enforceable it is necessary that

$$\frac{1-\delta}{\underline{\rho}^D(z, \hat{w}_z)} \leq \phi \leq \frac{\delta}{\overline{\rho}^C(z, \tilde{w}_z)}. \tag{7.12}$$

Such a ϕ exists if and only if

$$\frac{1-\delta}{\underline{\rho}^D(z, \hat{w}_z)} \leq \frac{\delta}{\overline{\rho}^C(z, \tilde{w}_z)} \xRightarrow{\underline{\rho}^D,\, \overline{\rho}^C > 0} \delta \geq \frac{\overline{\rho}^C}{\underline{\rho}^D + \overline{\rho}^C}. \tag{7.13}$$

This completes the proof. □

Remark 13. *The proofs of the threshold discount factors rely on the existence of solutions of the parameter $\phi > 0$ that make the ZD strategy well-defined. In Remark 11 (Chapter 6), ϕ was chosen as the upper bound in* Eq. (7.6). *In the public goods game this is a valid choice of ϕ for both generous and extortionate strategies, see* Eq. (7.35).

7.0.3 Equalizer ZD-strategies

The existence of equalizer strategies with $s = 0$ does not impose any requirement on the initial probability to cooperate. In general, one can identify different regions of the unit interval for p_0 in which different threshold discount factors exist. For instance, the boundary cases can be examined in a similar manner as was done for extortionate and generous strategies and, in general, will lead to different requirements on the discount factor. In this section, we derive an expression for the threshold discount factor such that the equalizer payoff relation can be enforced for a variable initial probability to cooperate that is within the open unit interval, i.e. $p_0 \in (0, 1)$.

***Theorem* 11** (p_0 and δ conditions for equalizers)**.** *Suppose $s = 0$ and l satisfies the bounds in Theorem 8. Then, the equalizer payoff relation can be enforced for $p_0 \in (0, 1)$ if and only if the following inequalities hold*

$$\delta \geq 1 - \frac{\underline{\rho}^D}{\underline{\rho}^D + (\overline{\rho}^D - \underline{\rho}^D)p_0}, \tag{7.14}$$

$$\delta \geq 1 - \frac{\underline{\rho}^C}{(1 - p_0)(\underline{\rho}^C + \overline{\rho}^D)}, \tag{7.15}$$

$$\delta \geq 1 - \frac{\underline{\rho}^C}{(1 - p_0)(\overline{\rho}^C - \underline{\rho}^C) + \underline{\rho}^C}, \tag{7.16}$$

$$\delta \geq 1 - \frac{\underline{\rho}^D}{\left(\overline{\rho}^C + \underline{\rho}^D\right) p_0}. \tag{7.17}$$

Proof. For brevity, we refer to equations found in the proof of Theorem 8. From Eq. (6.37) and Eq. (6.38) it follows that in order for the payoff relation to be enforceable for any $p_0 \in (0, 1)$ it must hold that for all σ such that $x_i = C$, $\rho^C(\sigma) > 0$, and for all σ such that $x_i = D$, $\rho^D(\sigma) > 0$. For the existence of equalizer strategies this must also hold for the special case in which $s = 0$. Hence, we can rewrite Eq. (6.37) and Eq. (6.38) to obtain the following set of inequalities

$$\frac{(1 - \delta)(1 - p_0)}{\underline{\rho}^C(z, \hat{w}_z)} \leq \phi \leq \frac{1 - (1 - \delta)p_0}{\overline{\rho}^C(z, \tilde{w}_z)}, \tag{7.18}$$

$$\frac{(1 - \delta)p_0}{\underline{\rho}^D(z, \hat{w}_z)} \leq \phi \leq \frac{\delta + (1 - \delta)p_0}{\overline{\rho}^D(z, \tilde{w}_z)}. \tag{7.19}$$

There exists such a $\phi > 0$ if and only if the following inequalities are satisfied

$$\frac{(1-\delta)p_0}{\underline{\rho}^D(z,\hat{w}_z)} \leq \frac{\delta + (1-\delta)p_0}{\overline{\rho}^D(z,\tilde{w}_z)}, \tag{7.20}$$

$$\frac{(1-\delta)p_0}{\underline{\rho}^D(z,\hat{w}_z)} \leq \frac{1 - (1-\delta)p_0}{\overline{\rho}^C(z,\tilde{w}_z)}, \tag{7.21}$$

$$\frac{(1-\delta)(1-p_0)}{\underline{\rho}^C(z,\hat{w}_z)} \leq \frac{1 - (1-\delta)p_0}{\overline{\rho}^C(z,\tilde{w}_z)}, \tag{7.22}$$

$$\frac{(1-\delta)(1-p_0)}{\underline{\rho}^C(z,\hat{w}_z)} \leq \frac{\delta + (1-\delta)p_0}{\overline{\rho}^D(z,\tilde{w}_z)}. \tag{7.23}$$

By collecting the terms in p_0 and δ for Eq. (7.20)-Eq. (7.23) the conditions can be derived as follows. Eq. (7.20) can be satisfied if and only if

$$p_0(1-\delta)\left(\overline{\rho}^D(z,\tilde{w}_z) - \underline{\rho}^D(z,\hat{w}_z)\right) \leq \underline{\rho}^D(z,\hat{w}_z)\delta.$$

In the special case that $\overline{\rho}^D(z,\tilde{w}_z) - \underline{\rho}^D(z,\hat{w}_z) = 0$, this is satisfied for every $p_0 \in (0,1)$ and $\delta \in (0,1)$. On the other hand, if $\overline{\rho}^D(z,\tilde{w}_z) - \underline{\rho}^D(z,\hat{w}_z) > 0$, then the inequality can be satisfied for every $p_0 \in (0,1)$ if and only if Eq. (7.14) holds. Likewise, Eq. (7.22) can be satisfied if and only if

$$-p_0(1-\delta)\left(\overline{\rho}^C - \underline{\rho}^C\right) \leq \underline{\rho}^C - (1-\delta)\overline{\rho}^C.$$

If $\overline{\rho}^C - \underline{\rho}^C = 0$, this inequality is satisfied for every $p_0 \in (0,1)$. On the other hand, if $\overline{\rho}^C - \underline{\rho}^C > 0$, the inequality is satisfied if and only if the condition in Eq. (7.16) holds. Eq. (7.21) holds if and only if the condition in Eq. (7.17) holds. Finally, Eq. (7.23) holds if and only if the condition in Eq. (7.15) holds. □

Based on Lemma 11, the following corollary provides relatively easy to check sufficient conditions that allow an equalizer strategy to enforce a desired linear relation for every initial probability to cooperate in the open unit interval. These sufficient conditions link thresholds for generous and extortionate strategies to those of equalizer strategies.

Corollary 6 (Sufficient conditions for equalizer thresholds)**.** *Suppose $s = 0$ and l satisfies the bounds in Theorem 8. Then, the equalizer payoff relation can be enforced for any $p_0 \in (0,1)$ if*

$$\delta \geq \delta_\tau = \max\left\{\frac{\overline{\rho}^C - \underline{\rho}^C}{\overline{\rho}^C}, \frac{\overline{\rho}^D - \underline{\rho}^D}{\overline{\rho}^D}, \frac{\overline{\rho}^D}{\underline{\rho}^C + \overline{\rho}^D}, \frac{\overline{\rho}^C}{\overline{\rho}^C + \underline{\rho}^D}\right\}.$$

Proof. It follows from the proof of Theorem 11 that for all $p_0 \in (0,1)$ it holds that $\underline{\rho}^D > 0$. Because $\overline{\rho}^D - \underline{\rho}^D \geq 0$ and Eq. (7.14) is linear in p_0 it follows that the condition Eq. (7.14) is satisfied for all $p_0 \in (0,1)$ if it holds in particular for the extreme case $p_0 = 1$, that is

$$\delta \geq \frac{\overline{\rho}^D - \underline{\rho}^D}{\overline{\rho}^D}.$$

Likewise, the conditions in Eq. (7.15), Eq. (7.16) and Eq. (7.17) are linear in p_0 and in their most stringent cases imply the fractions $\frac{\overline{\rho}^D}{\underline{\rho}^C + \overline{\rho}^D}$, $\frac{\overline{\rho}^C - \underline{\rho}^C}{\overline{\rho}^C}$, and $\frac{\overline{\rho}^C}{\overline{\rho}^C + \underline{\rho}^D}$ respectively. □

7.1 Applications

Under Assumption 8 the ZD strategist puts equal weight on each co-player and thus enforces a linear payoff relation between her own average discounted payoff and the mean of the average discounted payoffs of all her co-players. In this case, the functions that determine the threshold discount factors in Eq. (7.1) simplify into

$$\begin{aligned}
\overline{\rho}^C &= \max_{0 \leq z \leq n-1} (1-s)(a_z - l) + \frac{n-z-1}{n-1}(b_{z+1} - a_z), \\
\underline{\rho}^C &= \min_{0 \leq z \leq n-1} (1-s)(a_z - l) + \frac{n-z-1}{n-1}(b_{z+1} - a_z), \\
\overline{\rho}^D &= \max_{0 \leq z \leq n-1} (1-s)(l - b_z) + \frac{z}{n-1}(b_z - a_{z-1}), \\
\underline{\rho}^D &= \min_{0 \leq z \leq n-1} (1-s)(l - b_z) + \frac{z}{n-1}(b_z - a_{z-1}).
\end{aligned} \tag{7.24}$$

In the following, these functions will be used to derive threshold discount factors in the three social dilemma games that we have studied in Chapter 6.

7.1.1 Thresholds for n-player linear public goods games

Let us first examine the threshold discount factors of extortionate strategies and thus, $l = 0$ and $0 < s < 1$. In this case the parameters in Eq. (7.24) result from the extreme points of the functions

$$\rho_e^C(z) := (1-s)\left(\frac{rc(z+1)}{n} - c\right) + \frac{n-z-1}{n-1}c, \tag{7.25}$$

$$\rho_e^D(z) := -(1-s)\left(\frac{rcz}{n}\right) + \frac{z}{n-1}c. \tag{7.26}$$

From Proposition 6 we know that if $-\frac{1}{n-1} < s \leq 1 - \frac{n}{r(n-1)}$ no extortionate strategies can exist. Therefore, suppose that the slope is sufficiently large, i.e. $s \geq 1 - \frac{n}{r(n-1)}$. Then, the extreme points of $\rho_e^C(z)$ and $\rho_e^D(z)$ are determined as

$$\begin{aligned} \overline{\rho}_e^C &= \rho_e^C(0), & \underline{\rho}_e^C &= \rho_e^C(n-1), \\ \overline{\rho}_e^D &= \rho_e^D(n-1), & \underline{\rho}_e^D &= \rho_e^D(0). \end{aligned} \tag{7.27}$$

In the public goods game, next to the region of enforceable slopes, also the threshold discount factors for generous and extortionate strategies are equivalent, as highlighted in the following proposition.

Proposition 14 (Thresholds for extortion and generosity in public goods games)**.** *For the enforceable slopes $s \geq 1 - \frac{n}{r(n-1)}$, in the public goods game the threshold discount factor for extortionate and generous strategies is determined as*

$$\delta_\tau = \frac{1 - (1-s)(r - \frac{r}{n})}{1 - (1-s)(1 - \frac{r}{n})}. \tag{7.28}$$

Proof. For the linear public goods game the functions in Eq. (7.24) can be obtained from the extrema of the following functions

$$\begin{aligned} \rho^C(z) &= (1-s)\left(\frac{rc(z+1)}{n} - c - l\right) + \frac{n-z-1}{n-1}c, \\ \rho^D(z) &= (1-s)\left(l - \frac{rcz}{n}\right) + \frac{z}{n-1}c \end{aligned} \tag{7.29}$$

We focus first on the case in which $l = 0$ and $0 < s < 1$, and thus the strategy is extortionate. In this case Eq. (7.29) become

$$\rho_e^C(z) := (1-s)\left(\frac{rc(z+1)}{n} - c\right) + \frac{n-z-1}{n-1}c \tag{7.30}$$

$$\rho_e^D(z) := -(1-s)\left(\frac{rcz}{n}\right) + \frac{z}{n-1}c \tag{7.31}$$

We continue to obtain the maximizers and minimizers of Eq. (7.25), that because of linearity in z can only occur at the extreme points $z = 0$ and $z = n-1$. When $n > r$ and $r > 1$, as is the case when the linear public goods game is a social dilemma, we have the following simple conditions on the slope of the extortionate strategy. If $-\frac{1}{n-1} < s \leq 1 - \frac{n}{r(n-1)}$ no extortionate or generous strategies can exist. Hence assume $s \geq 1 - \frac{n}{r(n-1)}$. Then,

$$\begin{aligned}
\overline{\rho}_e^C &= \rho_e^C(0) = (1-s)\left(\frac{rc}{n} - c\right) + c, \\
\underline{\rho}_e^C &= \rho_e^C(n-1) = (1-s)(rc-c) > 0, \\
\overline{\rho}_e^D &= \rho_e^D(n-1) = -(1-s)\left(\frac{rc(n-1)}{n}\right) + c, \\
\underline{\rho}_e^D &= \rho_e^D(0) = 0.
\end{aligned} \tag{7.32}$$

The fractions in Theorem 9 become

$$\frac{\overline{\rho}_e^D}{\overline{\rho}_e^D + \underline{\rho}_e^C} = \frac{\overline{\rho}_e^C - \underline{\rho}_e^C}{\overline{\rho}_e^C} = \frac{(1-s)(\frac{r}{n} - r) + 1}{(1-s)(\frac{r}{n} - 1) + 1}. \tag{7.33}$$

We focus now on the case in which $l = rc - c$ and $0 < s < 1$, and hence the strategy is generous. If $l = rc - c$, Eq. (7.29) becomes

$$\begin{aligned}
\rho_g^C(z) &:= (1-s)\left(\frac{rc(z+1)}{n} - rc\right) + \frac{n-z-1}{n-1}c \\
\rho_g^D(z) &:= (1-s)\left(rc - c - \frac{rcz}{n}\right) + \frac{z}{n-1}c
\end{aligned} \tag{7.34}$$

The extreme points of these functions read as

$$\begin{aligned}
\overline{\rho}_g^C &= \rho_g^C(0) = \overline{\rho}_e^D, \\
\underline{\rho}_g^C &= \rho_g^C(n-1) = \underline{\rho}_e^D, \\
\overline{\rho}_g^D &= \rho_g^D(n-1) = \overline{\rho}_e^C, \\
\underline{\rho}_g^D &= \rho_g^D(0) = \underline{\rho}_e^C.
\end{aligned} \tag{7.35}$$

It follows that the fractions in Theorem 10 are equivalent to those in Theorem 9. This completes the proof. □

Remark 14 (Efficiency of enforcing the mutual cooperation payoff)**.** *Assume $r \leq \frac{n}{n-1}$; then the range of enforceable slopes in Proposition 14 includes $s = 0$ and the strategy can thus be an equalizer and the extreme points of the threshold functions ρ^C and ρ^D remain the same. Now assume $s = 0$; from Proposition 8 we know that $l = rc - c$ is enforceable by the equalizer strategy and from Proposition 14 we know that the threshold discount factor for the equalizer ZD-strategist to enforce the mutual cooperation payoff to her co-players is given by*

$$1 - n\left(1 - \frac{1}{r}\right). \tag{7.36}$$

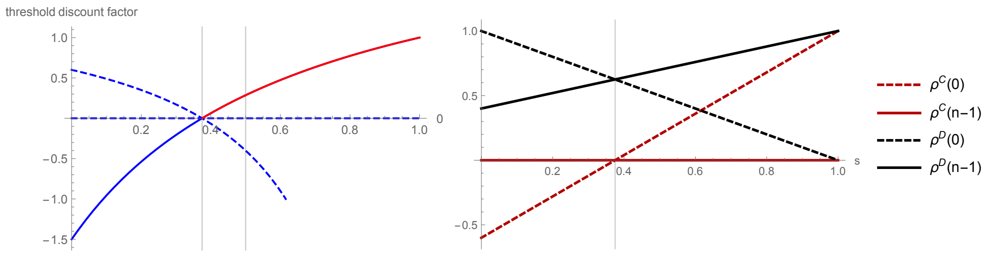


Figure 7.1: The left figure shows a numerical example of threshold discount factors for extortionate and generous strategies in the linear public goods game. The parameter values are $c = 1, r = 2, n = 5$. The lines represent the values of the fractions in the expression for δ_τ in Theorem 9 and Theorem 10 using the extreme points of the functions in Eq. (7.25) and Eq. (7.26). The threshold discount factor for an enforceable s can be determined from the left figure by the red curve. In the right figure, one can see how the extreme points in Eq. (7.27) change over s. For the existence of generous and extortionate strategies the point $s = 1 - \frac{n}{r(n-1)} = 3/8$ is a crucial point, namely, beyond this point up to $s < 1$ all the functions $\rho_e^C(z)$ and $\rho_e^D(z)$ in Eq. (7.25) and Eq. (7.26) (and those of generous strategies) are non-negative, which is necessary for existence. An equivalent condition is formulated in Proposition 6 in which for any slope $s < \frac{r-1}{r}$ for existence of extortionate and generous strategies it is necessary that $n \leq \frac{r(1-s)}{r(1-s)-1}$. Before this point, no generous or extortionate strategies can exist. The second vertical line indicates the point $s = \frac{r-1}{r} = 1/2$, after which any slope can be enforced independent of n, see Proposition 6.

7.1.2 Thresholds for n-player snowdrift games

The values of $\overline{\rho}^C(z)$ and $\underline{\rho}^C(z)$ from Eq. (7.24) for the n-player snowdrift game are obtained from the extreme points of the following expression, for $0 \leq z \leq n-1$:

$$\rho^C(z) = (1-s)\left(b - \frac{c}{z+1} - l\right) + \frac{n-z-1}{n-1}\frac{c}{z+1}. \tag{7.37}$$

For any enforceable slope $-\frac{1}{n-1} < s < 1$ the extreme points read as

$$\begin{aligned} \overline{\rho}^C(z) &= \rho^C(0) = (1-s)(b-c-l) + c, \\ \underline{\rho}^C(z) &= \rho^C(n-1) = (1-s)(b - \frac{c}{n} - l). \end{aligned} \tag{7.38}$$

The values of $\overline{\rho}^D(z)$ and $\underline{\rho}^D(z)$ are obtained from the extreme points of the function

$$\rho^D(z) = \begin{cases} (1-s)l, & \text{if } z = 0 \\ (1-s)(l-b) + \frac{c}{n-1}, & \text{if } z = 1 \ldots n-1. \end{cases} \tag{7.39}$$

We first focus on the extortionate case. Suppose $l = 0$, we obtain

$$\rho_e^D(z) := \begin{cases} 0, & \text{if } z = 0 \\ -(1-s)b + \frac{c}{n-1}, & \text{if } z = 1 \ldots n-1. \end{cases} \tag{7.40}$$

From Proposition 9, we know that for $s < 1 - \frac{c}{b(n-1)}$ no extortionate strategies can exist, therefore assume $s \geq 1 - \frac{c}{b(n-1)}$. Then, the extreme points of $\rho_e^D(z)$ read as

$$\underline{\rho}_e^D = \rho_e^D(0) = 0 \text{ and } \overline{\rho}_e^D = -b(1-s)\frac{c}{n-1}. \tag{7.41}$$

Proposition 15 (Thresholds for extortion in n-player snowdrift games)**.** *For the n-player snowdrift game with $b > c$, the threshold discount factor for the enforceable slopes $s \geq 1 - \frac{c}{b(n-1)}$ of an extortionate strategy is given by*

$$\delta_\tau = \frac{(1-s)(\frac{c}{n} - c) + c}{(1-s)(b-c) + c} \tag{7.42}$$

Proof. Assume $l = 0$ and $0 < s < 1$ such that the ZD-strategy is extortionate, from Eq. (7.39) we obtain

$$\rho_e^D(z) := \begin{cases} 0, & \text{if } z = 0 \\ \frac{c}{n-1} - (1-s)b, & \text{if } z = 1 \ldots n-1. \end{cases} \tag{7.43}$$

For $\frac{b}{c} > \frac{1}{(1-s)(n-1)}$ or equivalently, $s < 1 - \frac{c}{b(n-1)}$ it follows that $\overline{\rho}^D(z) = 0$. From Proposition 9 we know that in this case no extortionate strategies can exist. Therefore assume $s \geq 1 - \frac{c}{b(n-1)}$. Then, the extreme points of $\rho_e^D(z)$ are

$$\begin{aligned} \overline{\rho}_e^D &= \frac{c}{n-1} - (1-s)b \geq 0, \\ \underline{\rho}_e^D &= \rho_e^D(0) = 0. \end{aligned}$$

Using the expressions in Eq. (7.38) and substituting $l = 0$ we also have

$$\begin{aligned} \overline{\rho}_e^C &= \rho_e^C(0) = (1-s)(b-c) + c, \\ \underline{\rho}_e^C &= \rho_e^C(n-1) = (1-s)(b - \frac{c}{n}) > 0. \end{aligned} \tag{7.44}$$

The fractions in Theorem 9 become

$$\frac{\overline{\rho}_e^C - \underline{\rho}_e^C}{\overline{\rho}_e^C} = \frac{(1-s)(\frac{c}{n} - c) + c}{(1-s)(b-c) + c} \tag{7.45}$$

$$\frac{\overline{\rho}_e^D}{\overline{\rho}_e^D + \underline{\rho}_e^C} = \frac{\frac{c}{n-1} - b(1-s)}{\frac{c}{n-1} - \frac{c}{n}(1-s)} \tag{7.46}$$

Note that the denominator in Eq. (7.46) is strictly positive for any enforceable slope $s > -\frac{n}{n-1}$. Furthermore, for any $0 < s < 1$ and $b > c > 0$ it holds that

$$\frac{(1-s)(\frac{c}{n}-c)+c}{(1-s)(b-c)+c} > \frac{\frac{c}{n-1}-b(1-s)}{\frac{c}{n-1}-\frac{c}{n}(1-s)}.$$

This completes the proof. □

Now let us look at the threshold discount factors of generous strategies. Suppose $l = b - \frac{c}{n}$ and $0 < s < 1$.

$$\rho_g^D(z) := \begin{cases} (1-s)(b-\frac{c}{n}), & \text{if } z = 0 \\ -(1-s)\frac{c}{n} + \frac{c}{n-1}, & \text{if } z = 1 \dots n-1. \end{cases} \tag{7.47}$$

For $s \leq 1 - \frac{c}{b(n-1)}$ we have

$$\overline{\rho}_g^D = \rho_g^D(0), \quad \underline{\rho}_g^D = \rho_g^D(n-1). \tag{7.48}$$

And for $s > 1 - \frac{c}{b(n-1)}$ the extreme points become

$$\overline{\rho}_g^D = \rho_g^D(n-1), \quad \underline{\rho}_g^D = \rho_g^D(0). \tag{7.49}$$

Proposition 16 (Thresholds for generosity in n-player snowdrift games)**.** *For the n-player snowdrift game with $b > c$ and $n \geq 2$, for slopes $s \leq 1 - \frac{c}{b(n-1)}$ the threshold discount factor is determined by*

$$\delta_\tau = \max\left\{\frac{n-1}{n}, \frac{(1-s)b - \frac{c}{n-1}}{(1-s)(b-\frac{c}{n})}\right\} \tag{7.50}$$

For higher slopes $s > 1 - \frac{c}{b(n-1)}$,

$$\delta_\tau = \frac{(1-s)(\frac{c}{n}-c)+c}{(1-s)(b-c)+c} \tag{7.51}$$

Proof. Assume $l = b - \frac{c}{n}$ and $0 < s < 1$ such that the ZD-strategy is generous, from Eq. (7.39) we obtain

$$\rho_g^D(z) := \begin{cases} (1-s)(b-\frac{c}{n}), & \text{if } z = 0 \\ \frac{c}{n-1} - (1-s)\frac{c}{n}, & \text{if } z = 1 \dots n-1. \end{cases} \tag{7.52}$$

For $s \leq 1 - \frac{c}{b(n-1)}$ we have

$$\begin{aligned} \overline{\rho}_g^D &= \rho_g^D(0) = (1-s)(b-\frac{c}{n}) > 0, \\ \underline{\rho}_g^D &= \frac{c}{n-1} - (1-s)\frac{c}{n}. \end{aligned} \tag{7.53}$$

Note that $\underline{\rho}_g^D > 0$ for all $s > -\frac{1}{n-1}$. Using the expressions in Eq. (7.38) and substituting $l = 0$ we also have

$$\begin{aligned} \overline{\rho}_g^C &= (1-s)(\frac{c}{n} - c) + c > 0, \\ \underline{\rho}_g^C &= 0. \end{aligned} \tag{7.54}$$

Note that $\overline{\rho}_g^C > 0$ for any $s > -\frac{n}{n-1}$. The fractions in Theorem 10 become

$$\frac{\overline{\rho}_g^D - \underline{\rho}_g^D}{\overline{\rho}_g^D} = \frac{(1-s)b - \frac{c}{n-1}}{(1-s)(b - \frac{c}{n})} \tag{7.55}$$

$$\frac{\overline{\rho}_g^C}{\overline{\rho}_g^C + \underline{\rho}_g^D} = \frac{n-1}{n} \tag{7.56}$$

This completes the proof of the first statement. We now continue to the case in which $1 - \frac{c}{b(n-1)} < s < 1$. Then, the extreme points become

$$\overline{\rho}_g^D = \frac{c}{n-1} - (1-s)\frac{c}{n}, \quad \underline{\rho}_g^D = \rho_g^D(0) = (1-s)(b - \frac{c}{n}), \tag{7.57}$$

and the fractions in Theorem 10 become

$$\frac{\overline{\rho}_g^D - \underline{\rho}_g^D}{\overline{\rho}_g^D} = \frac{\frac{c}{n-1} - b(1-s)}{\frac{c}{n-1} - \frac{c}{n}(1-s)}, \tag{7.58}$$

$$\frac{\overline{\rho}_g^C}{\overline{\rho}_g^C + \underline{\rho}_g^D} = \frac{(1-s)(\frac{c}{n} - c) + c}{(1-s)(b-c) + c}, \tag{7.59}$$

it follows that in this region, the threshold discount factors for extortion and generosity in the n-player snowdrift game are equivalent. This completes the proof. □

Remark 15 (Efficiency of mutual cooperation in n-player snowdrift games)**.** *Assume* $s = 0 < 1 - \frac{c}{b(n-1)}$*. In this case,* Eq. (7.49) *and* Eq. (7.38) *are still satisfied. From Proposition 11 we know that* $l = b - \frac{c}{n}$ *is enforceable. From Proposition 16 we know the threshold discount factor for the equalizer strategist to enforce the mutual cooperation payoff on all its co-players is*

$$\max\left\{\frac{n-1}{n}, \frac{b(n-1) - c}{(b - \frac{c}{n})(n-1)}\right\}. \tag{7.60}$$

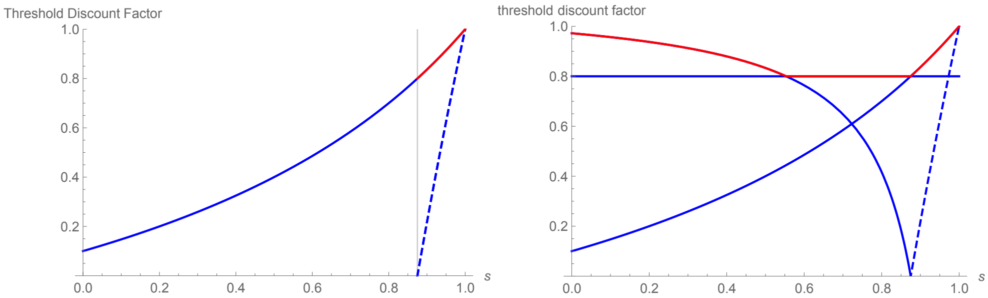


Figure 7.2: A numerical example of threshold discount factors for extortionate (left) and generous (right) strategies in the n-player snowdrift game with $b = 2$, $c = 1$ and $n = 5$. For extortionate strategies, the threshold is determined by the expression in Proposition 15. The red part of the line shows the threshold discount factor for enforceable slopes. As one can see, only relatively large slopes that satisfy $s \geq 1 - \frac{c}{b(n-1)} = \frac{7}{8}$ can be enforced by the extortioner. This critical point is indicated by the vertical line in the figure to the left and coincides with value of the slope for $n = 5$ in Figure 6.3. In the figure on the right, the threshold discount factors for generous strategies are shown as determined by Proposition 10. One can see that even though any slope can be enforced by a generous ZD strategist, the threshold discount factor depends on the value of the slope, and is illustrated by the red line. The blue lines in the plots indicate the several expressions for the threshold discount factor as formulated in the main text.

7.1.3 Thresholds for n-player stag-hunt games

The thresholds for extortionate and generous strategies can be determined by the extreme points of the functions

$$\begin{aligned} \rho^C(z) &= \begin{cases} \frac{(n-z-1)c}{n-1} - (1-s)(c+l), & \text{if } 0 \leq z < n-1; \\ (1-s)(b-c-l), & \text{if } z = n-1. \end{cases} \\ \rho^D(z) &= (1-s)l + \frac{zc}{n-1}. \end{aligned} \tag{7.61}$$

Now suppose $l = 0$; then the functions in Eq. (7.61) become,

$$\begin{aligned} \rho_e^C(z) &= \begin{cases} \frac{(n-z-1)c}{n-1} - (1-s)c, & \text{if } 0 \leq z < n-1; \\ (1-s)(b-c), & \text{if } z = n-1. \end{cases} \\ \rho_e^D(z) &= \frac{zc}{n-1}. \end{aligned} \tag{7.62}$$

Proposition 17 (Thresholds for extortion in n-player stag-hunt games)**.** *For the n-player stag hunt game with $b > c$, for any slope $s \geq 1 - \frac{c}{(n-1)b}$ the threshold discount factor for extortionate strategies is determined by*

$$\delta_\tau = \frac{c}{c + (1-s)(b-c)}.$$

Assume $n < \frac{2-s}{1-s}$ holds. Then, enforceable extortionate slopes in the region $1 - \frac{c}{b} \leq s < 1 - \frac{c}{b(n-1)}$ have a threshold discount factor determined by

$$\delta_\tau = \frac{1}{\frac{1}{n-1} + s}.$$

Assume $n < \frac{2-s}{1-s}$ holds. For enforceable extortionate slopes in the region $\frac{n-2}{n-1} < s \leq 1 - \frac{c}{b}$, the threshold discount is determined by

$$\delta_\tau = \max\left\{\frac{(1-s)b - \frac{c}{n-1}}{(1-s)(b-c)}, \frac{1}{\frac{1}{n-1} + s}\right\}.$$

Proof. Suppose $l = 0$ and $0 < s < 1$. Then, the extreme points of Eq. (7.62) become

$$\overline{\rho}_e^C = \max\{(1-s)(b-c), sc\} \tag{7.63}$$

$$\underline{\rho}_e^C = \min\left\{\frac{c}{n-1} - (1-s)c, (1-s)(b-c)\right\} \tag{7.64}$$

$$\overline{\rho}_e^D = c \tag{7.65}$$

$$\underline{\rho}_e^D = 0 \tag{7.66}$$

For different regions of the slope s the extreme points of ρ_e^C are different. For $s \geq 1 - \frac{c}{b(n-1)}$ we have

$$\overline{\rho}_e^C = \rho_e^C(0) = sc, \quad \underline{\rho}_e^C = \rho_e^C(n-1) = (1-s)(b-c). \tag{7.67}$$

And the thresholds are

$$\frac{\overline{\rho}_e^C - \underline{\rho}_e^C}{\overline{\rho}_e^C} = \frac{c - (1-s)b}{sc}, \tag{7.68}$$

$$\frac{\overline{\rho}_e^D}{\overline{\rho}_e^D + \underline{\rho}_e^C} = \frac{c}{c + (1-s)(b-c)}. \tag{7.69}$$

For $b > c > 0$ and $0 < s < 1$, the right-hand-side of Eq. (7.69) is larger than or equal to the right-hand-side of Eq. (7.68). This completes the first statement. From Proposition 12 we know that for slopes $s < 1 - \frac{c}{b(n-1)}$ in order for extortionate strategies to exist it needs to hold that $n < \frac{s-2}{s-1}$. Hence, assume that $n < \frac{s-2}{s-1}$. For the region of slopes $1 - \frac{c}{b} \leq s < 1 - \frac{c}{b(n-1)}$ we have

$$\overline{\rho}_e^C = \rho_e^C(0) = sc, \quad \underline{\rho}_e^C = \rho_e^C(n-2) = \frac{c}{n-1} - (1-s)c. \tag{7.70}$$

The thresholds become

$$\frac{\overline{\rho}_e^C - \underline{\rho}_e^C}{\overline{\rho}_e^C} = \frac{1 - \frac{1}{n-1}}{s}, \tag{7.71}$$

$$\frac{\overline{\rho}_e^D}{\overline{\rho}_e^D + \underline{\rho}_e^C} = \frac{1}{\frac{1}{n-1} + s}. \tag{7.72}$$

For $n < \frac{s-2}{s-1}$, the right-hand-side of Eq. (7.72) is larger than or equal to the right-hand-side of Eq. (7.71). This completes the second statement.

We again assume $n < \frac{s-2}{s-1}$, for smaller slopes in the region $\frac{n-2}{n-1} < s \leq 1 - \frac{c}{b}$ we obtain

$$\overline{\rho}_e^C = (1-s)(b-c), \quad \underline{\rho}_e^C = \frac{c}{n-1} - (1-s)c. \tag{7.73}$$

The thresholds become

$$\frac{\overline{\rho}_e^C - \underline{\rho}_e^C}{\overline{\rho}_e^C} = \frac{(1-s)b - \frac{c}{n-1}}{(1-s)(b-c)}, \tag{7.74}$$

$$\frac{\overline{\rho}_e^D}{\overline{\rho}_e^D + \underline{\rho}_e^C} = \frac{1}{\frac{1}{n-1} + s}. \tag{7.75}$$

It is worth noting that in the case of a non-strict upper bound $s = 1 - \frac{c}{b}$, it holds that $\overline{\rho}_e^C = (1-s)(b-c) = sc$ and the right-hand-side of Eq. (7.71) is equal to the right-hand-side of Eq. (7.74). This completes the proof. □

Now suppose $l = b - c$; then the functions in Eq. (7.61) become

$$\rho_g^C(z) = \begin{cases} \frac{(n-z-1)c}{n-1} - (1-s)b, & \text{if } 0 \leq z < n-1; \\ 0, & \text{if } z = n-1. \end{cases}$$
$$\rho_g^D(z) = (1-s)(b-c) + \frac{zc}{n-1}. \tag{7.76}$$

Proposition 18 (Thresholds for generosity in n-player snowdrift games)**.** *For the n-player stag hunt game with $b > c$, for any slope $s \geq 1 - \frac{c}{(n-1)b}$ the threshold discount factor for generous strategies is determined by*

$$\delta_\tau = \frac{c}{c + (1-s)(b-c)}$$

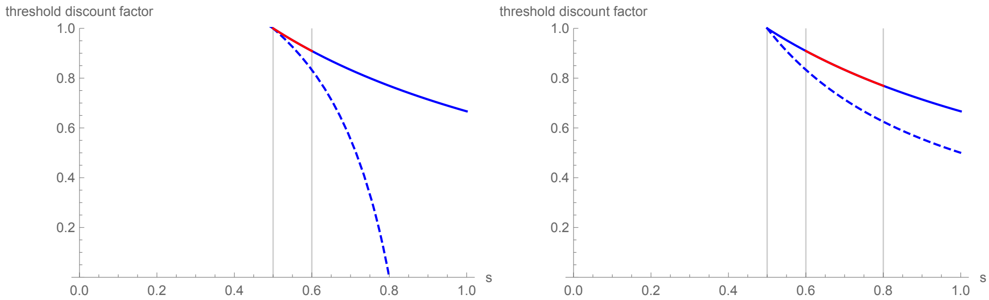


Figure 7.3: Numerical example of threshold discount factors for extortionate strategies in the n-player stag-hunt game with $n = 3$, $b = 5/2$ and $c = 1$. The left figure shows the threshold discount factors for slopes $\frac{n-2}{n-1} < s \leq 1 - \frac{c}{b}$. The figure on the right shows threshold discount factors for slopes $1 - \frac{c}{b} \leq s < 1 - \frac{c}{b(n-1)}$. The figures are obtained by the expressions in Proposition 17.

Proof. Suppose $l = b - c$. Then, the threshold functions read as

$$\overline{\rho}_g^C = \max\{0, c - (1-s)b\} \tag{7.77}$$

$$\underline{\rho}_g^C = \min\left\{\frac{c}{n-1} - (1-s)b, 0\right\} \tag{7.78}$$

$$\overline{\rho}_g^D = (1-s)(b-c) + c \tag{7.79}$$

$$\underline{\rho}_g^D = (1-s)(b-c) \tag{7.80}$$

From Proposition 13 we know that only slopes $s \geq 1 - \frac{c}{b(n-1)}$ can be enforced. For this region we obtain

$$\frac{\overline{\rho}_g^D - \underline{\rho}_g^D}{\overline{\rho}_g^D} = \frac{c}{(1-s)(b-c)+c} \tag{7.81}$$

$$\frac{\overline{\rho}_g^C}{\overline{\rho}_g^C + \underline{\rho}_g^D} = \frac{c-(1-s)b}{c(1+\frac{1}{n-1})-(1-s)b} \tag{7.82}$$

Because the denominator of Eq. (7.82) is strictly larger than 0 for $s > 1 - \frac{c}{b}$, the threshold is well-defined for any $s \geq 1 - \frac{c}{b(n-1)}$. Moreover, for $s > 1 - \frac{c}{b}$, the right-hand-side of Eq. (7.82) is larger than the right-hand-side of Eq. (7.81). This completes the proof. $\square$

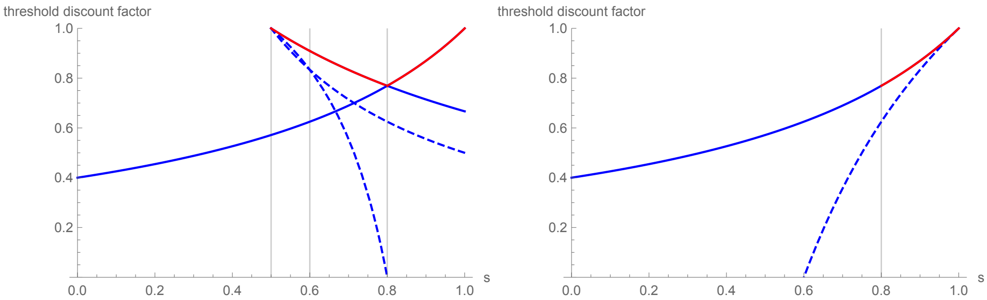


Figure 7.4: Numerical example of threshold discount factors in the n-player stag-hunt game with $n = 3$, $b = 5/2$ and $c = 1$. The red line in the left figure shows the threshold discount factors for the complete range of enforceable extortionate slopes. In this figure, one can also observe how the different regions of enforceable slopes, indicated by the vertical lines, are determined by the intersections of the blue lines that represent the ratios in Theorem 9 evaluated at the different extreme points of Eq. (7.62). The figure on the right shows the the threshold discount factors for generous strategies, as detailed in Proposition 18.

7.2 Final Remarks

With Theorems 9, 10 and 6, we have provided expressions for deriving the minimum discount factor for some desired linear relation. Because the expressions depend on the one-shot payoff of the n-player game, in general, they will differ between social dilemmas. To determine these expressions, one needs to find the global extrema of a function over z that can be efficiently done for a large class of social dilemma games. The derived thresholds can, for example, be used as an indicator for a minimum number of rounds in experiments on extortion and generosity in repeated games, or simply as an indicator for how many expected interactions a single ZD strategists requires to enforce some desired payoff relation in a group of decision-makers. Of particular interest to the emergence of cooperation in social dilemmas are the thresholds for equalizer strategies that enforcer the full cooperation payoff to all co-players. In the linear public goods game, this threshold depends non-linearly on n and r, see Eq. (7.36). When $n = 2$, this requirement turns into the simple condition $\delta_\tau = \frac{2-r}{r}$. For $1 < r < 2$ this is a decreasing function in r, which is to be expected. In the n-player snowdrift game it is also possible to enforce full cooperation, see Eq. (7.60). In the classic two-player snowdrift game, a simple threshold can be formulated that is the maximum between a half and $\frac{\frac{b}{c}-1}{\frac{b}{c}-\frac{1}{2}}$.

CHAPTER 8

Evolutionary stability of ZD strategy

> Mutation is random; natural selection is the very opposite of random.
>
> *Richard Dawkins*

IN the previous two chapters we have seen how a single memory-one player can exert a level of control on the payoffs of opponents by employing a ZD strategy in a repeated game. In this chapter we explore the stability of such strategies in an *evolutionary* setting. In [130], stochastic evolutionary imitation dynamics were studied in a finite population of ZD strategists playing an iterated prisoner's dilemma game. In these dynamics, strategies that receive higher expected payoffs are typically preferred by forces of selection [131]. It was shown that in reasonably large populations, extortionate strategies can act as a catalyst for the evolution of cooperation but they are not a stable outcome of natural selection. Through numerical simulation, it was also argued that the *population size* has a considerable impact on the dynamics and long-run outcomes in a well-mixed population, see also [132]. This effect was attributed to the probability of one being paired to themselves, which decreases as the population size increases. In [133], it was analytically shown that, in the limit of weak selection and pairwise interactions in a finite population, only generous strategies are "evolutionarily robust" against any other strategy in the repeated prisoner's dilemma game. That is, under the stochastic evolutionary imitation dynamics proposed in [134], only generous strategies cannot be selectively replaced by a mutant strategy. The

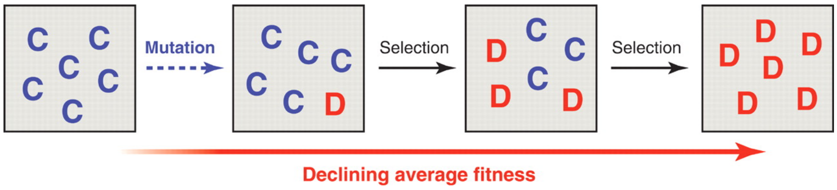


Figure 8.1: As a result of a random mutation, a population of C players is invaded by a single D player. If the mutant D player has an advantage in terms of payoffs over resident C players, selection will tend to favor D and the mutant can eventually take over the population. Figure from [8], reprinted with permission from AAAS.

evolutionary performance of ZD strategy in *multiplayer* games under similar stochastic evolutionary imitation dynamics were studied in [135]. Their simulations showed that for small group sizes extortionate ZD strategy are critical for the *emergence* of cooperation, and generous strategies are important to *maintain* cooperation in large finite populations. However, as group sizes become larger, the generous strategies become unstable and selfish behavior dominates in the long-run.

In this chapter, we study conditions for ZD strategy to be *evolutionarily stable* in a finite population of size N, when players interact in randomly formed groupwise contests of size $2 \leq n \leq N$. When the number of interactions is large, the composition of these groups can be described by a hypergeometric distribution. To obtain neat analytical results in this setting, we will focus on a finite population that is invaded by a single mutant (Fig. 8.1). Selection prefers the mutant strategy if the single mutant obtains a higher expected payoff than the resident players. On the other hand, selection will favor the resident strategy if residents obtain a higher expect payoff than the mutant. An illustration of this process in a social dilemma game is shown in Fig. 8.1, were a defecting mutant can take over the population of cooperators. In the following, we will derive explicit relations between the population size, the group size, and the ZD strategy parameters that allow us to characterize necessary and sufficient conditions under which a resident ZD strategy is evolutionarily stable with respect to a single mutant strategy. Moreover, a computationally convenient method is provided to evaluate the evolutionary stability of a resident strategy that is invaded by any number of mutant ZD strategy. The theoretical conditions shed light on how population sizes and group sizes influence the emergence and maintenance of cooperation in large finite populations and are consistent with simulation results in existing literature [130, 135].

The results in this chapter may appropriate only for the situation in which all players apply ZD strategies, however, when mutual defection leads to the lowest possible average group payoff, extortionate ZD strategies can ensure to do at least as good as any other strategy. In this case, the result presented in this chapter can be generalized to an arbitrary mutant strategy and therefore provide a rather general stability condition for extortionate strategies. The same holds for generous strategies when mutual cooperation leads to the highest possible average group payoff.

8.1 The standard ESS conditions

Consider a large well-mixed population of players of which the majority employs a resident strategy indicated by $\mathbf{p}^R$, and a small number of mutants employ a mutant strategy $\mathbf{p}^M$. The players interact in random pairwise contests. If a resident is matched with another resident they obtain $\pi(\mathbf{p}^R, \mathbf{p}^R)$, likewise if a resident is matched with a mutant the resident receives $\pi(\mathbf{p}^R, \mathbf{p}^M)$ and the mutant receives $\pi(\mathbf{p}^M, \mathbf{p}^R)$. If two mutants are matched they receive $\pi(\mathbf{p}^M, \mathbf{p}^M)$. Let π^R (resp. π^M) denote the *expected* payoff of a resident (resp. mutant) player in the evolutionary game. Then, the resident strategy is an *Evolutionarily Stable Strategy* (ESS) if it satisfies the following two conditions [29, 136].

a) Equilibrium condition: the resident strategy $\mathbf{p}^R$ is a best response against itself

$$\forall \mathbf{p}^M \neq \mathbf{p}^R : \quad \pi\left(\mathbf{p}^M, \mathbf{p}^R\right) \leq \pi\left(\mathbf{p}^R, \mathbf{p}^R\right). \tag{8.1}$$

b) Stability conditions: if $\pi(\mathbf{p}^M, \mathbf{p}^R) = \pi\left(\mathbf{p}^R, \mathbf{p}^R\right)$, then

$$\pi\left(\mathbf{p}^M, \mathbf{p}^M\right) < \pi\left(\mathbf{p}^R, \mathbf{p}^M\right). \tag{8.2}$$

These conditions are valid in infinite populations. To see this, suppose that a small fraction $\epsilon > 0$ of the population are mutants, then

$$\pi^R = (1-\epsilon)\pi\left(\mathbf{p}^R, \mathbf{p}^R\right) + \epsilon\pi\left(\mathbf{p}^R, \mathbf{p}^M\right), \tag{8.3}$$

and

$$\pi^M = (1-\epsilon)\pi\left(\mathbf{p}^M, \mathbf{p}^R\right) + \epsilon\pi\left(\mathbf{p}^M, \mathbf{p}^M\right). \tag{8.4}$$

The conditions in Eq. (8.1) and Eq. (8.2) simply imply that

$$\exists \epsilon > 0 : \quad \pi^M < \pi^R,$$

and thus the mutation will be selected against. However, if the population size is *finite* a problem arises with this definition. Because a player cannot be matched in a contest against herself, the probability of a mutant being matched to another mutant is not equal to the probability of a resident being matched to a mutant. Naturally, this discrepancy becomes larger when N becomes smaller and smaller.

8.2 Generalized ESS equilibrium condition

Earlier work that studied ESS ZD strategy in finite populations used the evolutionary stability conditions presented in [132] that rely on the computation of fixation probabilities. In contrast, this chapter employs an approach proposed in [136] that is suitable for analysis and has an intuitive connection to the standard Maynard-Smith conditions for evolutionarily stable strategies described in the previous section. Let us now consider a finite population of size N. At each discrete time step, that may represent a generation, groups of size n are randomly formed by sampling from the population without replacement and engage in contests of size $n \leq N$. Let $\pi(\mathbf{p}^R|n-1-j,j)$ denote the payoff that a resident player obtains from a contest with $n-1-j$ other residents and j mutants and let $\pi(\mathbf{p}^M|n-1-j,j)$ denote the payoff that a mutant obtains from a contest against $n-1-j$ residents and j mutants. Starting from a homogeneous population of resident players, let us assume one resident player is replaced by a mutant. With this single mutant in the population, the average payoff of a resident ZD player is

$$\pi^R = \left(1 - \frac{n-1}{N-1}\right) \pi\left(\mathbf{p}^R|n-1,0\right) + \frac{n-1}{N-1}\pi\left(\mathbf{p}^R|n-2,1\right). \tag{8.5}$$

And the average payoff of the mutant is simply:

$$\pi^M = \pi\left(\mathbf{p}^M|n-1,0\right). \tag{8.6}$$

Using the expected payoffs in Eq. (8.5) and Eq. (8.6), for the resident strategy to be evolutionarily stable the following equilibrium condition is necessary [136],

$$\forall \mathbf{p}^M \neq \mathbf{p}^R: \quad \pi^M \leq \pi^R. \tag{8.7}$$

By substituting the expected payoffs, we obtain

$$\pi\left(\mathbf{p}^M|n-1,0\right) \leq \left(1 - \frac{n-1}{N-1}\right) \pi\left(\mathbf{p}^R|n-1,0\right) + \frac{n-1}{N-1}\pi\left(\mathbf{p}^R|n-2,1\right).$$

Observe that for $n=2$ and $N \to \infty$, this equilibrium condition turns into the standard Maynard-Smith equilibrium condition in Eq. (8.1) in which $\mathbf{p}^R$ needs to be a best response against itself.

Remark 16 (Playing the field). *The following observations were made in [136], well before ZD strategy were invented. A strategy satisfying the ESS equilibrium condition* Eq. (8.7) *may be written as a solution of*

$$\underset{\mathbf{p^M}}{\operatorname{argmax}}\ \pi^M - \pi^R. \tag{8.8}$$

Now suppose that $n = N$, which is a situation Maynard-Smith refers to as "playing the field". Then, $\pi^R = \pi(\mathbf{p}^R|n-2,1)$ and an ESS strategy is a solution to

$$\underset{\mathbf{p^M}}{\operatorname{argmax}}\ \pi\left(\mathbf{p}^M|n-1,0\right) - \pi\left(\mathbf{p}^R|n-2,1\right) \tag{8.9}$$

Because every player in the population is in the contest, we can use the conventional notation of n-player games. To this end, label the mutant as player m and the ESS players as $r \neq m$. Furthermore, denote the strategy of player j by s_j, $j = 1 \dots N$. Because all ESS players employ the same strategy, for symmetric games, we may write

$$\pi^{ESS} = \pi^R = \pi\left(\mathbf{p}^R|n-2,1\right) = \frac{1}{n-1}\sum_{i\neq m}^{N} \pi\left(s_i, s_{-i}\right),$$

by substituting this into the equilibrium condition Eq. (8.8) *one can see that any strategy satisfying the equilibrium ESS condition is a solution to*

$$\underset{\mathbf{p^M}}{\operatorname{argmax}}\ \pi^M - \frac{1}{n-1}\sum_{i\neq m}^{N} \pi(s_i, s_{-i}). \tag{8.10}$$

This, in turn, is a best response for Shubik's zero sum "beat-the-average" game [137]. The relation in Eq. (8.10) *is a natural connection between the ESS equilibrium condition and extortionate strategies that ensure the ZD strategist's expected payoff is larger than the average of her opponents. Thus, when $n = N$, extortionate ZD strategy satisfy the ESS equilibrium condition. This property is further examined in Theorem 12.*

8.3 Equilibrium conditions for ZD strategies

To investigate the ESS conditions of ZD strategy in finite populations and groupwise contests, we assume that at each time step, groups of size n are randomly formed by sampling from the population without replacement and engage in contest in the form of a finitely repeated n-player game. The following assumption ensures that the ZD strategists enforce a linear relation in each groupwise contest.

Assumption 9 (Repeated contests). *In each contest, the players in the evolutionary game participate in a sufficient number of expected rounds to enforce a payoff relation. This can be determined by the threshold discount factors in Chapter 7.*

Assumption 9 requires multiple interactions to occur during one time step or generation in the evolutionary game. This differs from the traditional setup but is common in evolutionary dynamics of direct *reciprocity* [131].

To obtain an equilibrium condition for some resident ZD strategy, we assume all weights are equal such that Assumption 8 is satisfied. Let us consider a finite population that is initially composed entirely of players that employ a ZD strategy such that each player enforces a payoff relation with parameters (s, l). Now, say that one random player is replaced by a mutant that employs a ZD strategy that enforces the payoff relation $(\hat{s}, \hat{l})$. We indicate this mutant strategy by $\mathbf{p}^M \neq \mathbf{p}^R$.

***Theorem* 12** (Equilibrium conditions)**.** *Assume all weights are equal and payoffs satisfy Assumption 7. When the $N-1$ resident players employ an enforceable ZD strategy $\mathbf{p}^R$ and the single mutant employs some enforceable ZD strategy $\mathbf{p}^M \neq \mathbf{p}^R$ in a population of size N with contests of size $n \leq N$, then the resident ZD strategy $\mathbf{p}^R$ satisfies the equilibrium condition* Eq. (8.7) *if and only if one of the following conditions hold:*

(i) $s > \frac{1}{N-1} + \frac{n-2}{n-1}$ *and* $\hat{l} \leq l$*;*

(ii) $s < \frac{1}{N-1} + \frac{n-2}{n-1}$ *and* $\hat{l} \geq l$*;*

(iii) $s = \frac{1}{N-1} + \frac{n-2}{n-1}$*.*

Proof. We begin by finding an expression for $\pi(\mathbf{p}^R|n-1, 0)$. Because the payoff relation (s, l) is assumed to be enforceable it must hold that $s < 1$. Then each resident enforces the linear relation:

$$\pi\left(\mathbf{p}^R|n-1, 0\right) = \pi\left(\mathbf{p}^R|n-1, 0\right)s + (1-s)l.$$

It follows that $\pi\left(\mathbf{p}^R|n-1, 0\right) = l$. We continue to find expressions for the payoffs when the mutant is present in the contest. For ease of notation, let $\pi^* = \pi\left(\mathbf{p}^R|n-2, 1\right)$. Because the residents apply an enforceable ZD strategy we obtain the payoff relation:

$$\frac{n-2}{n-1}\pi^* + \frac{1}{n-1}\pi^M = \pi^* s + (1-s)l. \tag{8.11}$$

Define $s^* := s(n-1) - (n-2)$. Then the above equation can be written as

$$\pi^M = s^*\pi^* + (1-s^*)l. \tag{8.12}$$

Because the mutant plays an enforceable ZD strategy with parameters $(\hat{l}, \hat{s})$ the following linear relation is enforced by the mutant:

$$\pi^* = \pi^M\hat{s} + (1-\hat{s})\hat{l}. \tag{8.13}$$

Substituting this into Eq. (8.12) we obtain an expression for the payoff of the mutant in terms of the ZD parameters $s^*, l, \hat{s}, \hat{l}$:

$$\pi^M = \frac{l(1-s^*) + \hat{l}s^*(1-\hat{s})}{1-\hat{s}s^*}. \tag{8.14}$$

Because $s < 1$, it holds that $s^* < 1$ and because the mutant payoff relation is also assumed to be enforceable it also holds that $\hat{s} < 1$. Hence the above equation is well defined for any enforceable ZD strategy in the repeated contest. By substituting the expressions for π^M, π^* and $\pi(\mathbf{p}^R|n-1, 0)$ into the equilibrium condition Eq. (8.7), we obtain

$$s^*\pi^* + (1-s^*)\, l \leq \left(1 - \frac{n-1}{N-1}\right) l + \frac{n-1}{N-1}\pi^*. \tag{8.15}$$

Collecting the terms in π^* and bringing them to the left-hand side we obtain

$$\left(s^* - \frac{n-1}{N-1}\right)\pi^* \leq \left(s^* - \frac{n-1}{N-1}\right) l. \tag{8.16}$$

From Eq. (8.16) we can distinguish three cases: first, $s^* = \frac{n-1}{N-1}$; second, $s^* > \frac{n-1}{N-1}$; third, $s^* < \frac{n-1}{N-1}$. We consider these cases separately. First suppose $s^* = \frac{n-1}{N-1}$ then Eq. (8.16) implies $0 = 0$ and the equilibrium condition is always satisfied and condition (iii) follows. Now suppose $s^* > \frac{n-1}{N-1}$, then Eq. (8.16) implies:

$$\pi^* \leq l \xRightarrow{\text{Eq. (8.13)}} \pi^M\hat{s} + (1-\hat{s})\hat{l} \leq l.$$

By substituting Eq. (8.14) into this equation we obtain

$$\left[\frac{l(1-s^*) + \hat{l}s^*(1-\hat{s})}{1-\hat{s}s^*}\right]\hat{s} + (1-\hat{s})\hat{l} \leq l. \tag{8.17}$$

Because $1 - \hat{s}s^* > 0$, Eq. (8.17) is satisfied if and only if

$$(\hat{l} - l)(1-\hat{s}) \leq 0.$$

And because $(1-\hat{s}) > 0$, this inequality can be satisfied if and only if

$$\hat{l} \leq l.$$

Hence, when $s^* > \frac{n-1}{N-1}$, then $\hat{l} \leq l$ is a necessary and sufficient condition for the equilibrium condition to be satisfied for enforceable payoff relations (s, l) and $(\hat{s}, \hat{l})$. Condition (i) is obtained by substituting the definition of s^* into this requirement.

Now suppose $s^* < \frac{n-1}{N-1}$; then Eq. (8.16) implies:

$$\pi^* \geq l \Rightarrow \hat{l} \geq l.$$

By flipping the inequality sign in Eq. (8.17), condition (ii) follows. This completes the proof. □

Corollary 7 (Playing the field equilibrium conditions)**.** *Under the conditions of Theorem 12, when the group size of the contests is equal to the population size, that is $n = N$, in the finitely repeated n-player game only extortionate ZD strategy can satisfy the equilibrium condition* Eq. (8.7)*. In particular, by substituting $N = n$ into condition (ii) it follows that any extortionate ZD strategy with a slope $s < 1$ satisfies the equilibrium condition.*

Corollary 8 (Maynard-Smith conditions)**.** *Let $n = 2$ such that the interactions are pairwise. From condition (ii), it follows that an extortionate strategy satisfies the equilibrium condition if and only if $s < \frac{1}{N-1}$. Now, observe that in the limit of an infinite population size $N \to \infty$ this condition becomes $s < 0$, which is a contradiction with the definition of an extortionate strategy that have positive slopes $0 < s < 1$. Hence, in the standard Maynard-Smith ESS equilibrium condition, extortionate strategies cannot be ESS. In fact, only equalizer strategies with $s = 0$ and generous strategies with $s > 0$ and $l = a_{n-1}$ satisfy the equilibrium condition under the classic conditions.*

Remark 17. *When $N \to \infty$, the equilibrium condition (i) in Theorem 12 is in line with the analytical condition for generous strategies being able to withstand an invasion of an ALLD mutant given in [135]. Our result however shows that when the population size N is finite, in order for generous strategies to withstand an invasion from an extortionate mutant, the generosity of the residents players needs to decrease according to the size of the population.*

Remark 18 (Instability of generous equalizers)**.** *Condition (i) in Theorem 12 indicates that for a resident ZD strategy with $l = a_{n-1}$ to satisfy the equilibrium condition it must hold that $s > \frac{1}{N-1} + \frac{n-2}{n-1}$. Because this equilibrium condition on the slope implies $s > 0$, any equalizer strategy with $s = 0$ and $l = a_n - 1$ cannot be evolutionarily stable. Hence, strategies that enforce the full cooperation payoff to all co-players exist in the public goods game and the n-player snowdrift game but cannot be sustained in an evolutionary selection process within a finite population.*

8.4 Stability conditions for ZD strategies

Let us now consider the case in which there exist k identical mutants in the population. Assuming a large number of interactions, the expected payoffs of the residents and mutants are [136]

$$\pi^R = \sum_{j=0}^{k} \frac{\binom{k}{j}\binom{N-1-k}{n-1-j}}{\binom{N-1}{n-1}} \pi(\mathbf{p}^R \mid n-1-j, j), \tag{8.18}$$

$$\pi^M = \sum_{j=0}^{k-1} \frac{\binom{k-1}{j}\binom{N-k}{n-1-j}}{\binom{N-1}{n-1}} \pi(\mathbf{p}^M \mid n-j, j-1). \tag{8.19}$$

To obtain convenient expressions for these payoffs we use the following Lemma from which the payoffs of a group of ZD strategists can be obtained in terms of the slopes and baseline payoffs of their strategies.

Lemma 13 (payoffs of groups of ZD strategists, [135])**.** *Suppose in a group of n players, every player applies some ZD strategy with parameters s_i and l_i. Then the expected payoff of player i can be written as*

$$\pi_i = (\kappa_i + 1)\frac{\sum_{k=1}^{n} \kappa_k \cdot l_k}{\sum_{k=1}^{n} \kappa_k} - \kappa_i \cdot l_i, \ \textit{with } \kappa_k := \frac{(n-1)(1-\mathrm{s}_k)}{1+(n-1)\mathrm{s}_k}. \tag{8.20}$$

Proof. The proof follows from rewriting the enforced linear payoff relations by including one's own expected payoff in the average of the group. Further details can be found in [135]. □

In order to obtain an expression for the expected payoffs of residents and mutants, let us define

$$\kappa_R := \frac{(n-1)(1-s)}{1+(n-1)s}, \kappa_M := \frac{(n-1)(1-\hat{s})}{1+(n-1)\hat{s}}, \psi_j := \frac{j \cdot \kappa_M \cdot \hat{l} + (n-j)\kappa_R \cdot l}{j \cdot \kappa_M + (n-j)\kappa_R}.$$

From Lemma 13 it follows that the payoffs of the residents and mutants in a group with j identical mutants and $n-j$ identical residents are given by

$$\pi(\mathbf{p}^R \mid n-1-j, j) = (\kappa^R + 1) \cdot \rho_j - \kappa_R \cdot l,$$
$$\pi(\mathbf{p}^M \mid n-j, j) = (\kappa^M + 1)\rho_j - \kappa_M \cdot \hat{l}.$$

By plugging these payoffs into Eq. (8.18), the stability condition in a finite population of ZD strategy can be evaluated solely on the basis of the hypergeometric distribution and the ZD strategy parameters in the population.

8.5 Applications

In the remainder of this chapter, we apply the result in Theorem 12 to the n-player linear public goods games, n-player snowdrift games and n-player stag hunt games. In doing so, we combine the existence conditions for generous and extortionate strategies in the n-player social dilemma games with the equilibrium condition Eq. (8.7). The following results give insight under which conditions on the population size N and the group size n, generous and extortionate strategies exist and when they are able to withstand an invasion by a single mutant strategist with an arbitrary but fixed strategy. Throughout this section it is assumed that Assumptions 9 and 8 are satisfied such that within each generation, the expected number of rounds is sufficient for the ZD strategist to enforce the linear payoff relation. Depending on the game being played, this minimum number of expected rounds in each generation can be determined by the threshold discount factor for generosity in Propositions 14, 16 and 18.

8.5.1 n-player linear public goods games

Proposition 19 (Stable generosity in public goods games). *Suppose Assumptions 9 and 8 are satisfied. Then in the n-player linear public goods game, all generous strategies with slopes $s \geq \frac{r-1}{r}$ satisfy the equilibrium condition if and only if $r < n < 1 + r$ and*

$$N \geq \frac{n(r-1)-1}{r-n+1},$$

and for smaller enforceable slopes $s < \frac{r-1}{r}$ this inequality needs to hold strictly.

Proof. From Proposition 7 we know that in order for generous strategies to enforce a payoff relation independent of the group size n it must hold that $s \geq \frac{r-1}{r}$. Moreover, from the conditions (i) and (iii) in Theorem 12, we know that the slope must satisfy

$$s \geq \frac{1}{N-1} + \frac{n-2}{n-1}. \tag{8.21}$$

This equilibrium condition on the slope can be satisfied for all enforceable slopes $s \geq \frac{r-1}{r}$ if and only if

$$\frac{1}{N-1} + \frac{n-2}{n-1} \leq \frac{r-1}{r}.$$

This inequality is satisfied for the conditions in the statement. For smaller slopes $s < \frac{r-1}{r}$ to be enforceable, it needs to hold that $n \leq \frac{r(1-s)}{r(1-s)-1}$ or equivalently, $s \geq 1 - \frac{n}{r(n-1)}$. Hence for $s < \frac{r-1}{r}$, the equilibrium condition can be satisfied if and only if

$$s \geq \max\left\{\frac{1}{N-1} + \frac{n-2}{n-1}, 1 - \frac{n}{r(n-1)}\right\} \overset{r<n}{\Longrightarrow} s \geq \frac{1}{N-1} + \frac{n-2}{n-1}.$$

It follows that in order for the equilibrium condition to be satisfied it is required that

$$\frac{1}{N-1} + \frac{n-2}{n-1} \leq s < \frac{r-1}{r},$$

such an s exist if and only if $\frac{1}{N-1} + \frac{n-2}{n-1} < \frac{r-1}{r}$, which is satisfied when the condition in the statement is strict. □

Proposition 20 (Stable extortion in public goods games)**.** *Suppose Assumptions 9 and 8 are satisfied. Then in the n-player linear public goods game, any enforceable extortionate strategy with a slope $s \geq \frac{r-1}{r}$ satisfies the equilibrium condition if $n > r+1$ or $n < r+1$ and $N \leq \frac{n(r-1)-1}{r-n+1}$. Smaller slopes $s < \frac{r}{r-1}$ satisfy the equilibrium condition regardless of n and N.*

Proof. From Theorem 12 and Proposition 6 it follows that in order for enforceable slopes $s \geq \frac{r}{r-1}$ to satisfy the equilibrium condition it must hold that

$$\frac{r-1}{r} \leq s \leq \frac{1}{N-1} + \frac{n-2}{n-1}.$$

These inequalities are satisfied for the conditions in the statement. Smaller slopes $s < \frac{r}{r-1}$ are only enforceable if the group size satisfies $n \leq \frac{r(1-s)}{r(1-s)-1}$. Together with the requirements in Theorem 12, it follows that in order for the enforceable slope to satisfy the equilibrium condition it needs to hold that

$$1 - \frac{n}{r(n-1)} \leq s \leq \frac{1}{N-1} + \frac{n-2}{n-1}.$$

This condition is satisfied for any $N \geq n > r$. □

Figure 8.2 shows numerical examples of the influence of N and n on the equilibrium conditions for generous and extortionate strategies in the n-player linear public goods game. The blue regions indicate the slopes and population sizes for which generous strategies satisfy the equilibrium condition. Extortionate strategies satisfy the equilibrium condition in the region between the vertical line and the border of the blue region, this region is indicated in the figure by the dot and text "stable extortion". Thus, for relatively small slopes, extortionate strategies satisfy the equilibrium condition. On the contrary, for higher slopes, generous strategy satisfy the equilibrium condition. At the border of the equilibrium regions for extortionate and generous strategies at $s = \frac{1}{N-1} + \frac{n-2}{n-1}$, both strategies satisfy the equilibrium condition. This is an immediate result of condition (*iii*) in Theorem 12. One can observe in the figures that as N increases more generous slopes satisfy the equilibrium condition. Thus, as the population size increases, players can become more generous without risking an invasion of defecting strategies. However, as n increases, the total region in which generous strategies satisfy the equilibrium condition shrinks. Indicating that in larger groups, it becomes more difficult for generous strategies to resist an invasion of an extortionate strategy. The bottom two figures show an example of

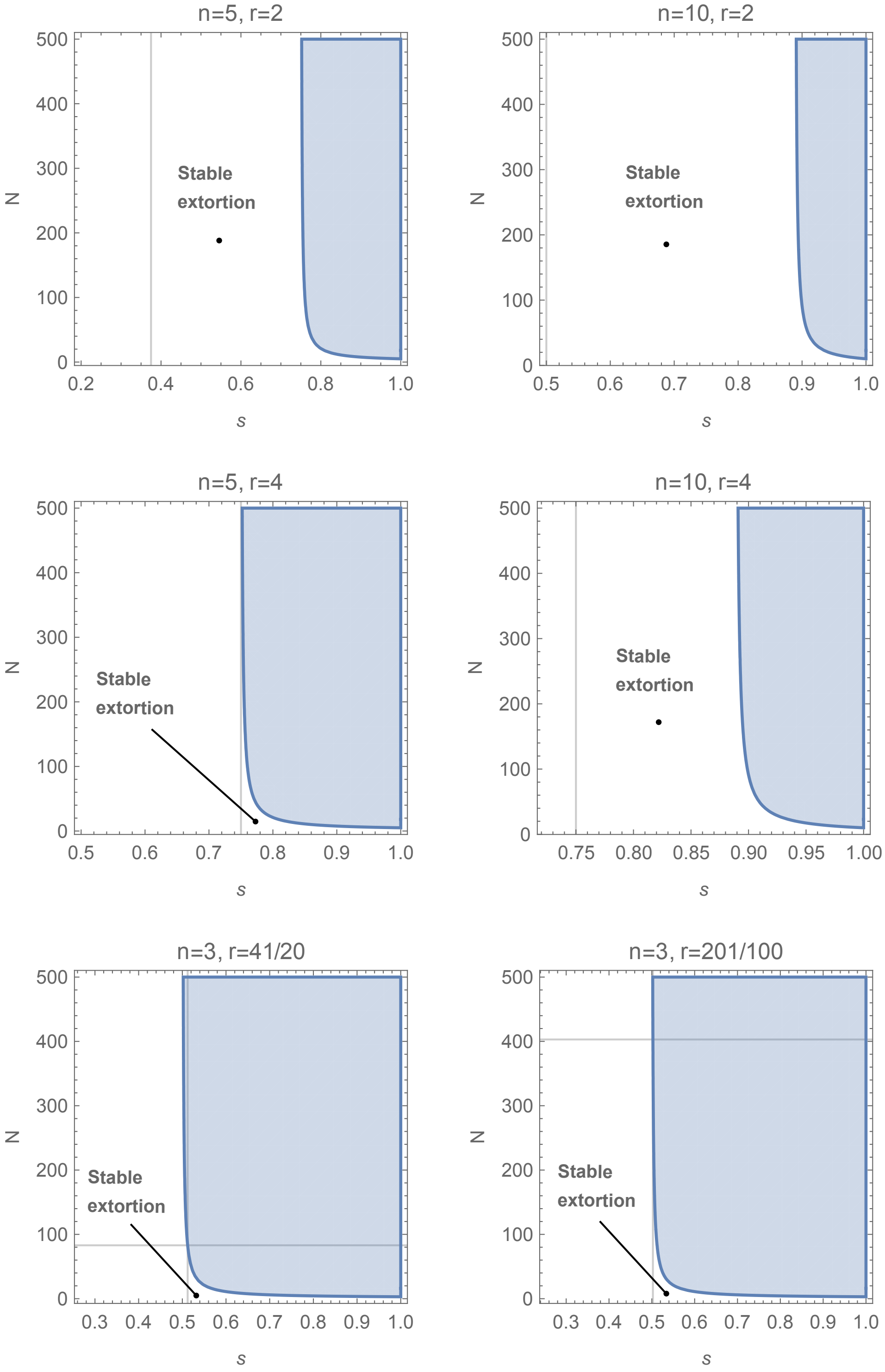


Figure 8.2: Equilibrium condition for generous and extortionate strategies in n-player linear public goods game played in a finite population of size N. In the blue area, generous strategies satisfy the equilibrium condition. In the area to the right of the vertical line and below the blue curve, extortionate strategies satisfy the equilibrium condition.

the result in Proposition 19. The vertical line indicates the point $s = \frac{r-1}{r}$, after which an extortionate and generous strategy can enforce a linear relation independent of n (see Propositions 6 and 7). For generous strategies, every such slope satisfies the equilibrium condition if and only if the population size is sufficiently large, see Proposition 19. For particular values of n and r, this population size is indicated by the horizontal line in the bottom two figures. One can see that the equilibrium condition is very sensitive to slight changes in the public goods multiplier r. When $r = \frac{41}{20}$, in order for all generous slopes $s \geq \frac{r-1}{r}$ to satisfy the equilibrium condition, the population size needs to be just above 90, but when r decreases to $r = 201/100$ (a difference of 0.04), the population size needs to grow beyond 400.

8.5.2 n-player snowdrift games

Proposition 21 (Stable generosity in n-player snowdrift games)**.** *Suppose Assumptions 9 and 8 are satisfied. In the n-player snowdrift game, any slope $0 < s < 1$ of a generous strategies is enforceable. Moreover, the generous strategy satisfies the equilibrium condition if and only if*

$$s \geq \frac{1}{N-1} + \frac{n-2}{n-1}.$$

Proof. The proof follows immediately from Proposition 10 and Theorem 12. □

Proposition 22 (Stable extortion in n-player snowdrift games)**.** *Suppose Assumptions 9 and 8 are satisfied. For the n-player snowdrift game, extortionate strategies satisfy the equilibrium condition if and only if $n = N$ or*

$$N \leq \frac{bn-c}{b-c}$$

Proof. From Proposition 9 we know that enforceable slopes satisfy

$$s \geq 1 - \frac{c}{b(n-1)}.$$

From Theorem 12 it follows that in order for the extortionate strategy to satisfy the equilibrium condition it needs to hold that

$$1 - \frac{c}{b(n-1)} \leq s \leq \frac{1}{N-1} + \frac{n-2}{n-1}.$$

Such an s exists if and only if $n = N$ or b satisfies the bound in the statement. □

Figure 8.3 shows numerical examples of generous and extortionate slopes in the n-player snowdrift game that satisfy the equilibrium condition. As can be seen in

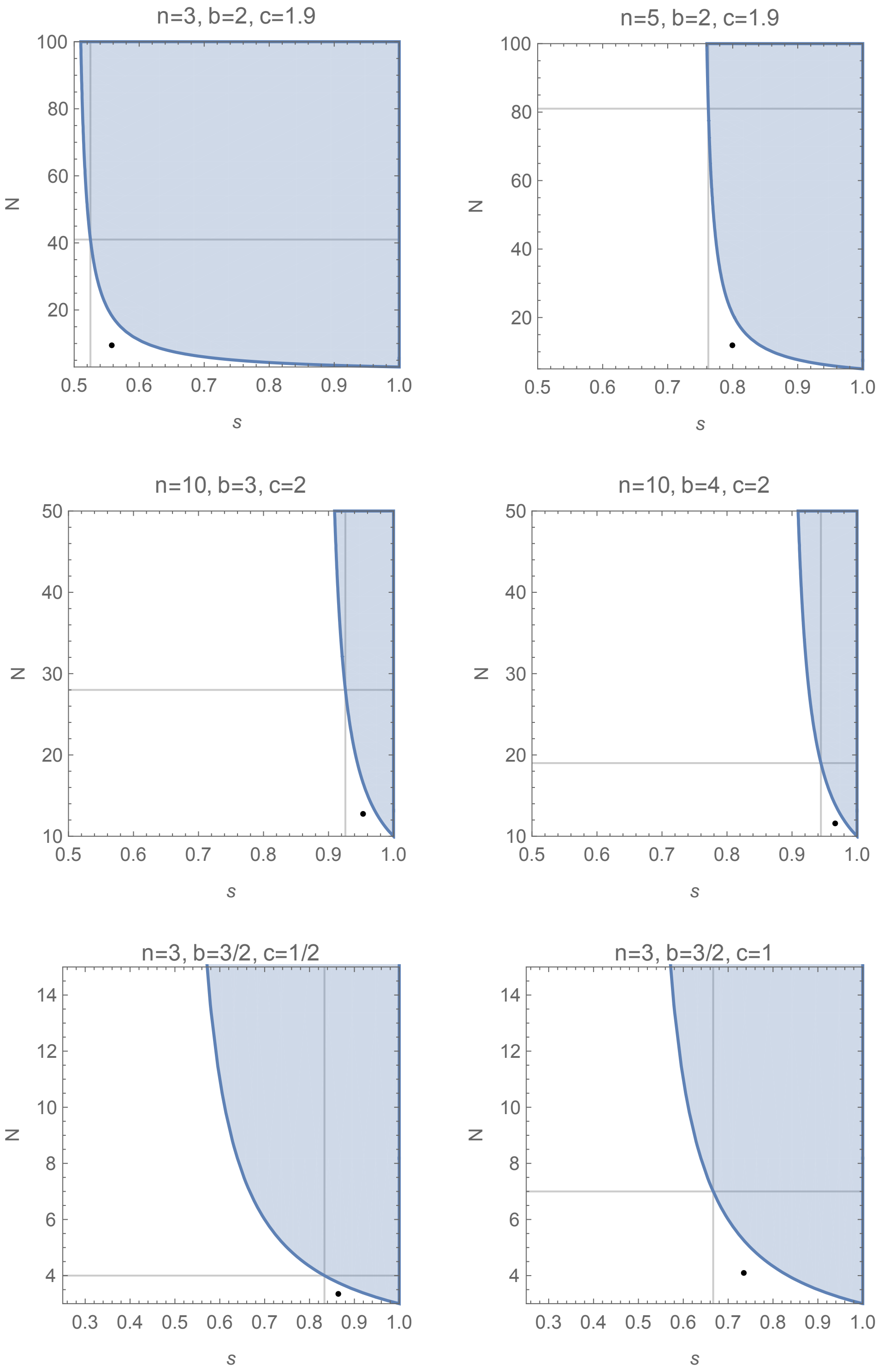


Figure 8.3: Equilibrium condition for generous and extortionate strategies in n-player snowdrift games played in a finite population of size N. In the blue area, generous strategies satisfy the equilibrium condition. In the small regions below the blue curve and to the right of the vertical line, extortionate strategies satisfy the equilibrium condition. A dot is used to indicate the region. On the border, the slope is $s = \frac{1}{N-1} + \frac{n-2}{n-1}$, and both classes of strategies satisfy the equilibrium condition.

the top two figures, when n increases, extortionate slopes can satisfy the equilibrium condition for larger population sizes and less generous strategies satisfy the equilibrium condition. As a result, when the size of groupwise contests increases, in order for generous strategies to satisfy the equilibrium condition, the population size must increase as well. This is also shown by the vertical lines that indicate the size of N, below which extortionate strategies can satisfy the equilibrium condition, see also Proposition 22. As can be seen in the middle two figures, as the benefit b increases, in order for extortionate slopes to satisfy the equilibrium condition, the population size must become smaller. This indicates that, as the benefit increases, extortionate strategies can only invade a generous population if the population size is sufficiently smaller.

8.5.3 n-player stag-hunt games

Proposition 23 (Stable generosity in n-player stag-hunt games)**.** *Suppose Assumptions 9 and 8 are satisfied. For the n-player stag hunt game with $b > c$, every generous strategy satisfies the equilibrium condition if and only if $N \geq \frac{bn-c}{b-c}$.*

Proof. From Proposition 13 we know that in order for the generous strategy to be enforceable it is required that $s \geq 1 - \frac{c}{b(n-1)}$. In this case, the conditions in Theorem 12 can be satisfied for every enforceable slope if and only if $\frac{1}{N-1} \leq 1 - \frac{c}{b(n-1)} - \frac{n-2}{n-1} = \frac{b-c}{b(n-1)}$, note that because $b > c > 0$ the right hand side is strictly larger than 0. This leads to the requirement in the main statement. □

Proposition 24 (Stable extortion in n-player stag-hunt games)**.** *For the n-player stag hunt game with $b > c$, extortionate strategies with a slope $s \geq 1 - \frac{c}{(n-1)b}$ satisfy the equilibrium condition if and only if $N \leq \frac{bn-c}{b-c}$. Smaller slopes $s < 1 - \frac{c}{b(n-1)}$ satisfy the equilibrium condition independent of $N \geq n > 1$.*

Proof. From Proposition 12 we know that an extortionate strategy can enforce any slope $s \geq 1 - \frac{c}{(n-1)b}$ independent of n. In order to satisfy the equilibrium condition it must hold that $1 - \frac{c}{(n-1)b} \leq s \leq \frac{1}{N-1} + \frac{n-2}{n-1}$; Such an s exists if and only if the condition on N in the main statement holds. For smaller enforceable slopes $s < 1 - \frac{c}{(n-1)b}$ it is required that $n < \frac{2-s}{1-s}$, equivalently, $s > \frac{n-2}{n-1}$; Such slopes satisfy the equilibrium condition if and only if $\frac{n-2}{n-1} < s \leq \frac{1}{N-1} + \frac{n-2}{n-1}$, which is clearly satisfied independent of $1 < N < \infty$. □

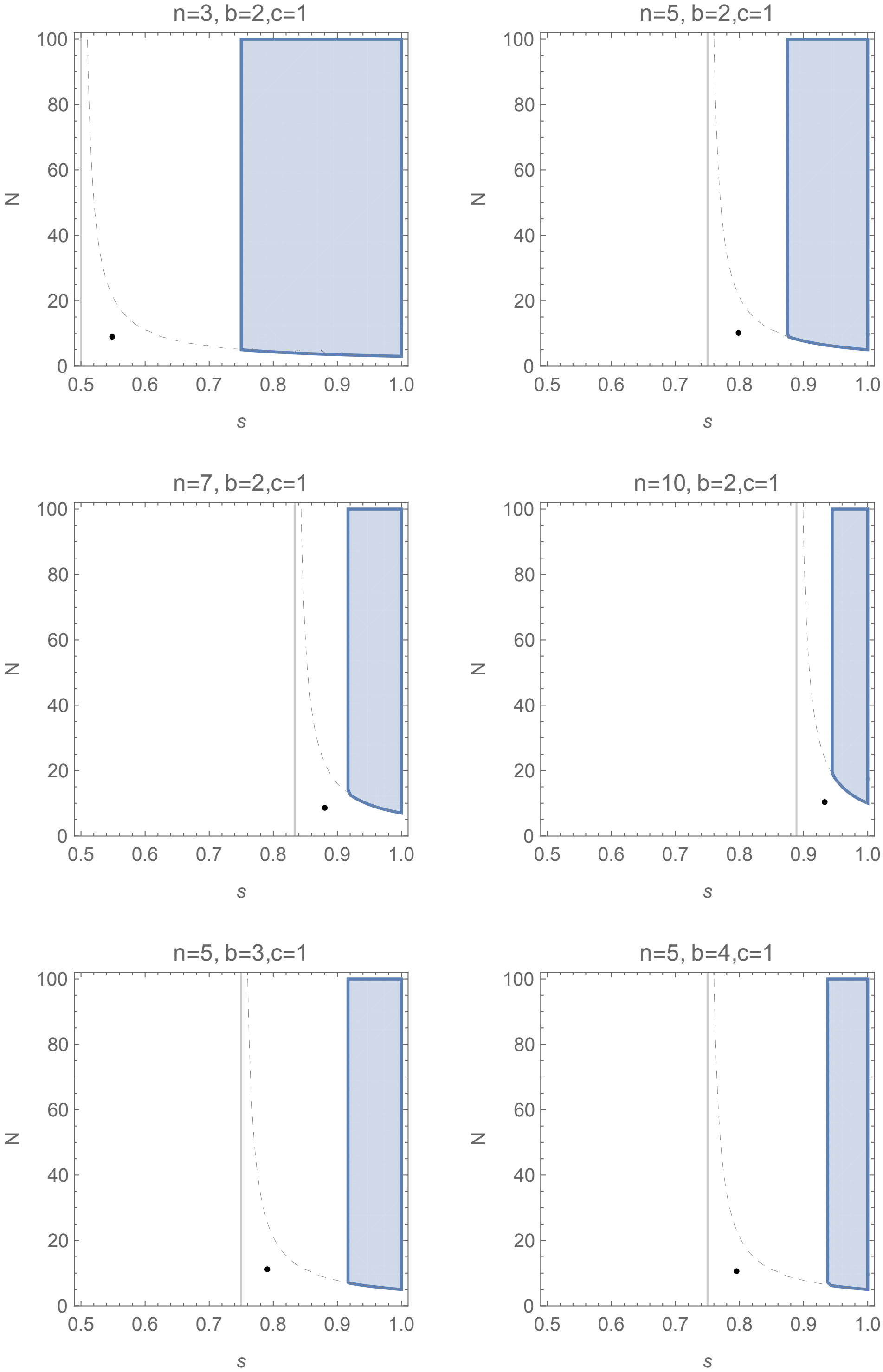


Figure 8.4: Equilibrium condition for generous and extortionate strategies in n-player stag-hunt games played in a finite population of size N.

8.6 Final Remarks

We have shown how the population size, the group size of the contests, and the ZD strategy parameters affect the evolutionary stability of a resident ZD strategy with respect to a single mutant. Extortionate strategies cannot be evolutionarily stable under the classic Maynard-Smith conditions. In sharp contrast, when the population size is equal to the group size, only extortionate strategies can be evolutionarily stable. In finite populations in which the group size is smaller than the population size both generosity and extortion can be stable; however, this highly depends on the benefit to cost ratio, the population size and group size of the contests. In the previous chapter, we have identified under which conditions equalizers strategies can enforce the full cooperation payoff in a group of players. However, here we have shown that these strategies cannot be evolutionarily stable. To show the utility of the results we have applied them to three n-player social dilemmas and provided explicit conditions under which generous and extortionate strategies are evolutionarily stable with respect to a single mutant ZD strategist in a finite population.

In the next chapter we return to the traditional repeated game setting in which a fixed group of players interact repeatedly. In particular, we will investigate the level of control that an individual can exert if the probability for continuation δ is *uncertain*.

CHAPTER 9

Exerting control under uncertain discounting of future outcomes

> I believe that we do not know anything for certain, but everything probably.
>
> *Christiaan Huygens*

If individuals choose between rewards that differ *only* in amount, timing, or certainty, decisions are relatively predictable because general principles of choice apply. For example, individuals tend to choose higher rewards over lower ones, sooner rewards over later rewards, and secure rewards over risky rewards. Indeed, such decisions make sense from both an economic and evolutionary perspective and are observed in both humans and animals [138, 139]. Predicting decisions becomes more challenging when the choice options differ in a combination of these factors. For example, it can be difficult to predict how an individual chooses between a small but immediate reward and a large but distant one. Although such combinations of different features usually require trade-offs in decision making, saliently temporal features can be studied from the perspective of *discounting* on the basis of the expected time or likelihood of their occurrence [139, 140]. Theoretical models of both temporal and probabilistic discounting often use hyperbola-like functions in which discount rates decrease monotonically over time [141–143]. And indeed, these hyperbolic discounting functions prove to be a better fit to empirical data than exponential functions [139, and the references therein]. From the very nature of *games*, outcomes depend on not only one's own decision but also the decisions of others. This interdependence inherently causes some level of uncertainty in the (probabilistic) outcome and its associated payoff. After all, in most real-world scenarios a decision-maker cannot force others to behave *exactly* to their liking. It becomes more complicated in *repeated* games, where a series of interactions occur over time, and individuals need to deal with the possibilities of how their past and current decisions can influence *future* payoffs under altruism, antagonism, punishment or reward [144–146]. Clearly, this is a rather complex setting in which both temporal and probability discounting based on hyperbolic discounting functions are likely to play a role. Consequently, traditional exponential discounting methods, which are commonly applied to repeated games, seem less suitable for describing how individuals would make trade-offs in real-world decisions. Indeed, discrepancies between economic and evolutionary models of cooperation and observed experimental behaviors motivated researchers to investigate how an individual's *uncertain* beliefs about the number of game interactions affects the possibilities to cooperate in one-shot prisoner's dilemma game [147]. It was found that this source of uncertainty can indeed explain the "overly" generous behavior that experimentalists observed [20, 20, 148, 149]. Interestingly, economics research on discounting shows that there is an immediate connection between uncertainty and the hyperbolic discounting functions observed in subjects. Namely, if one's belief of the discount rate is distributed according to a gamma or exponential distribution, then discounting will be hyperbolic [138, 150]. The exponential discount rates that are commonly applied to repeated games have two equivalent interpretations. First, it can be seen as a source of probabilistic discounting in which the constant discount rate represents a *continuation probability.* Second, it can be interpreted as a source of temporal discounting in which the present values of future payoffs are determined according to a fixed interest rate. However, independent of the interpretation of the discount rate, *uncertainty* in its value seems to be coupled to real-world discounting in such repeated interactions.

Inspired by the work of Press and Dyson [64], recent developments in the theory of repeated games suggest that a single, or small group of strategic individuals, can have a much larger influence on the other players' performances than previously anticipated [115–118, 121]. It is, however, not yet known how these intricate strategies hold up under the influence of uncertainty. By incorporating a common uncertain belief about the discount factor into these manipulative strategies, we generalize existing theories on zero-determinant strategies and show how a witty strategic player can unilaterally exert control in repeated games with probabilistic discounting. The proposed theoretical framework of discounting supports the hyperbolic form observed by experimentalists and can recover infinitely repeated games without discounting *and* exponentially discounted repeated games that were studied in Chapters 6 and 7. We postulate that this theoretical framework is more appropriate for describing real-world decision-making procedures in which judgments on the number of repeated interactions is made under uncertainty [147]. To obtain the specific results, we again consider the general class of symmetric repeated n-player social dilemma games introduced in Chapter 6.

9.1 Uncertain repeated games

In repeated games with finite but undetermined time horizons, the expected number of rounds is determined by a *fixed* and common discount factor $\delta \in (0, 1)$ that, given the current round of interactions, determines the probability of a next round and is therefore referred to as a *continuation probability.* Consequently, *expected* payoffs are calculated using a discounting function δ^t that corresponds to deterministic discrete-time exponential discounting with a constant discount rate [53]. If, however, the continuation probability or discount rate is *uncertain*, then the payoffs relying on these future interactions are uncertain as well and it is not immediate that a fixed parameter can be used to represent expected payoffs. In the spirit of *gamma discounting* [150], let us instead assume that discounting takes the form $d_k(t) = x_k^t$, where the $\{x_k\}$ are distributed according to the realization of a *random variable* x, whose probability density function $f(x, \alpha, \beta)$, defined for all $x \in [0, 1]$, is of the beta form[1]

$$f(x, \alpha, \beta) := \frac{x^{(\alpha-1)}(1-x)^{(\beta-1)}}{\mathrm{B}(\alpha, \beta)}, \quad \alpha, \beta \in \mathbb{R}_+,$$

where $\mathrm{B}(\alpha, \beta)$ is the beta function of the form

$$\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

The mean and variance of the beta distribution are, as usual, defined as

$$\mu = \frac{\alpha}{\alpha+\beta}, \quad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Indeed, the beta distribution is often used to describe the distribution of a *probability* and is thus a suitable choice [151]. Examples of such a distribution are given in Figure 9.1. By using the relation between beta and gamma functions, the obtained *effective discounting function* [150] becomes

$$d(t) := \int_0^1 x^t f(x, \alpha, \beta) dx = \frac{\Gamma(t+\alpha)\Gamma(\alpha+\beta)}{\Gamma(t+\alpha+\beta)\Gamma(\alpha)}, \tag{9.1}$$

[1] The notation of the shape parameters α, β of the Beta distribution should not be confused with the linear payoff relation parameters used in Chapter 6.

where $\Gamma(\cdot)$ indicates the gamma function. This effective discounting function indicates how in *beta discounting*, delayed or uncertain future payoffs are discounted when the probability for future interactions is uncertain. As one would expect, the payoffs that are received *now* are not subject to this uncertainty and are 'discounted' by the factor $d(0) = 1$. In contrast to the deterministic case with a fixed discount rate, the *rate of change* of the effective discount function in Eq. (9.1) is

$$\frac{d(t+1) - d(t)}{d(t)} = -\frac{\beta}{t + \alpha + \beta}, \quad t \geq 0, \tag{9.2}$$

and thus is in line with the empirically well-supported feature of hyperbolic discounting in which the effective discount rate *decreases* monotonically over time [138,139,150, and the references therein]. Evaluations of the effective discounting function and an example of deterministic exponential discounting are given in Figure 9.2. Going back to the main formulation of the effective discount function in Eq. (9.1), if one denotes the expected payoff of player i in round t by $\pi_i(t)$, then the *average discounted payoff* of player i in the repeated game with beta discounting is

$$\pi_i := \frac{\sum_{t=0}^{\infty} d(t)\pi_i(t)}{\sum_{t=0}^{\infty} d(t)}. \tag{9.3}$$

To evaluate this payoff, we note that the series of the effective discounting function *converges* for $\beta > 1$ to

$$\frac{\alpha + \beta - 1}{\beta - 1}. \tag{9.4}$$

Thus, the shape parameters of the discount factor's distribution analytically determine the normalization factor of the average discounted payoffs. It is worth pointing out that the requirement $\beta > 1$ excludes the possibility of a uniform or U-shaped distribution of the discount rate, but does not limit the skewness of the distribution as shown in Figure 9.1.

Remark 19 (Determistic limits). *In the deterministic limit $\alpha \to \infty$ and $\beta < \infty$, an infinitely repeated game without discounting is recovered. Moreover, if one sets $\beta = \frac{\alpha(1-\delta)}{\delta}$ with $\delta \in (0, 1)$, then in the deterministic limit $\alpha \to \infty$, arbitrarily high probability density is put on δ and exponential discounting with a fixed discount factor δ is recovered.*

9.1.1 Time-dependent memory-one strategies and mean distributions

Zero-determinant strategies determine the probability to select an action based only on the outcome of the previous round, and therefore belong to the class of

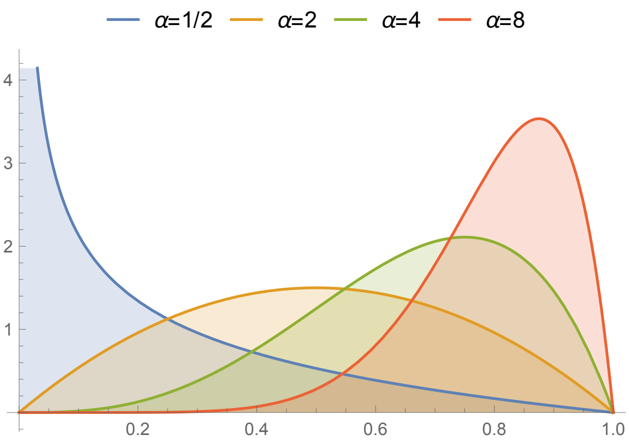


Figure 9.1: Examples of distributions of the discount factor for $\beta = 2$ and variable $\alpha \in \{\frac{1}{2}, 2, 4, 8\}$. If $\alpha = \beta$ the distribution is symmetric. If $\alpha > \beta$ (resp. $\alpha < \beta$), the distribution is negatively skewed (resp. positively skewed). The requirement $\beta > 1$ excludes the possibility for a uniform distribution with $\alpha = \beta = 1$ or a u-shaped distribution with $\alpha, \beta < 1$.

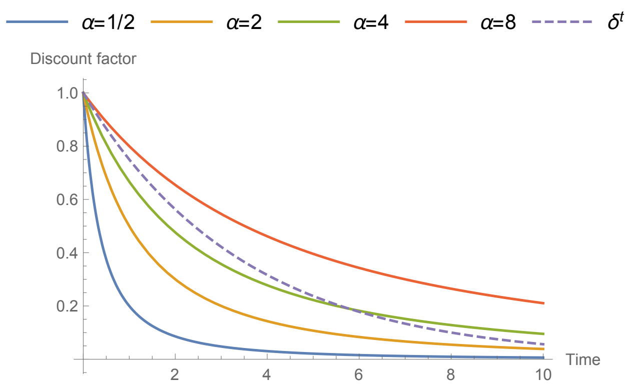


Figure 9.2: Evaluations of the effective discount function in Eq. (9.1) for $\beta = 2$ and $\alpha = \{\frac{1}{2}, 2, 4, 8\}$. The purple line corresponds to deterministic exponential discounting with $\delta = \frac{3}{4}$. One can see that with a constant discount rate, the curve bends faster towards zero. This deterministic behavior can be approximated by the effective discounting function in Eq. (9.1) using $\beta = \frac{\alpha}{3}$ for sufficiently large values of α.

memory-one strategies. When the rate of change of the discount factor is *fixed*, these strategies can be written as a vector whose elements are time-independent conditional probabilities [64, 118, 152]. However, when the discount factor is uncertain, its rate of change given in Eq. (9.2), varies with time and a strategic player should take this time dependency into account when deciding whether to cooperate (C) or defect (D). Now, let $p^t_\sigma \in [0,1]$ denote the probability that the strategic player cooperates at round $t+1$ given that, at round t, the action profile is σ. By stacking the conditional probabilities for all possible outcomes into a vector, we obtain a *time-dependent* memory-one strategy that determines the probability for the strategic player to cooperate at round $t+1$:

$$\mathbf{p}^t = (p^t_\sigma)_{\sigma\in\{C,D\}^n}.$$

Accordingly, the repeated memory-one strategy given by $\mathbf{p}^{\text{rep}}$ determines the probability to cooperate when the current decision is simply repeated and is defined as in Chapter 6. Let $v(t) = (v_\sigma(t))_{\sigma\in\{C,D\}^n}$ be the vector of outcome probabilities at round t as in Chapter 6. Using the limit of the series of the effective discounting function in Eq. (9.4), the mean distribution of the action profiles is

$$\mathbf{v} = \frac{\sum_{t=0}^{\infty} d(t)v(t)}{\sum_{t=0}^{\infty} d(t)} = \frac{\beta-1}{\alpha+\beta-1}\sum_{t=0}^{\infty} d(t)v(t). \tag{9.5}$$

In order to relate the average discounted payoff to the mean distribution we, introduce some additional notation adopted from [118]. Remember that g^i_σ denotes the one-shot payoff that the strategic player i receives in the action profile $\sigma \in \{C,D\}^n$. By stacking the possible payoffs into a vector, one obtains g^i, which contains all possible payoffs of player i in a given round of play. As in Chapter 6, the expected payoff of player i at round t can then be expressed by multiplying this one-shot payoff vector by the probability of the outcome. That is, $\pi_i(t) = g^i \cdot v(t)$. Consequently, the average discounted payoff is $\pi_i = g^i \cdot \mathbf{v}$, and the expected payoffs in the repeated game follow from the mean distribution $\mathbf{v}$ and the payoff functions of the social dilemma.

9.2 Risk-adjusted zero-determinant strategies

Let us now investigate under which conditions a single strategic player, say player i, can unilaterally enforce a linear relation in the average discounted payoff calculated according to Eq. (9.3). Towards this end, one would need to know the relation between the time dependent memory-one strategy $\mathbf{p}^t$ and the mean distribution $\mathbf{v}$. As in the deterministic case, we use the fact that the probability that player i cooperated at round t is $q_C(t) = \mathbf{p}^{\text{rep}} \cdot v(t)$. And the probability that i cooperates at the next round $t+1$ is $q_C(t+1) = \mathbf{p}^t \cdot v(t)$. Now, let us define the function

$$u(t) := \frac{d(t+1)}{d(t)} q_C(t+1) - q_C(t) = \frac{t+\alpha}{t+\alpha+\beta} q_C(t+1) - q_C(t), \tag{9.6}$$

and observe that its discounted *telescoping* sum evaluates as

$$\sum_{t=0}^{T} d(t)u(t) = d(0)\left(\frac{\alpha}{\alpha+\beta}q_C(1) - q_C(0)\right) + \\ d(1)\left(\frac{1+\alpha}{1+\alpha+\beta}q_C(2) - q_C(1)\right) + d(2)\left(\frac{2+\alpha}{2+\alpha+\beta}q_C(3) - q_C(2)\right) + \dots \\ + d(t)\left(\frac{T+\alpha}{T+\alpha+\beta}q_C(T+1) - q_C(T)\right) = d(t)\frac{T+\alpha}{T+\alpha+\beta}q_C(T+1) - d(0)q_C(0). \tag{9.7}$$

For real $\beta > 1$ and $\alpha > 0$, we have

$$\lim_{t\to\infty} d(t) = 0. \tag{9.8}$$

Thus,

$$\lim_{T\to\infty}\sum_{t=0}^{T} d(t)u(t) = \lim_{T\to\infty}\sum_{t=0}^{T} d(t)\left(\frac{t+\alpha}{t+\alpha+\beta}\mathbf{p}^t - \mathbf{p}^{\text{rep}}\right)\cdot v(t) = -q_C(0).$$

Furthermore, dividing by the series in Eq. (9.4), we obtain

$$\frac{\beta-1}{\alpha+\beta-1}\sum_{t=0}^{\infty} d(t)u(t) = \frac{\beta-1}{\alpha+\beta-1}\sum_{t=0}^{\infty} d(t)\left(\frac{t+\alpha}{t+\alpha+\beta}\mathbf{p}^t - \mathbf{p}^{\text{rep}}\right)\cdot v(t) \tag{9.9}$$

$$= \frac{\beta-1}{\alpha+\beta-1}\sum_{t=0}^{\infty} d(t)v(t)\cdot\left(\frac{t+\alpha}{t+\alpha+\beta}\mathbf{p}^t - \mathbf{p}^{\text{rep}}\right) \tag{9.10}$$

$$= -\frac{\beta-1}{\alpha+\beta-1}q_C(0) = -\frac{\beta-1}{\alpha+\beta-1}p_0, \tag{9.11}$$

where p_0 is player i's initial probability to cooperate, i.e. $p_0 := q_C(0)$.

Remark 20 (Relation to deterministic discounting). *The relation in* Eq. (9.9) *can be seen as a probabilistic discounting extension of Akin's result on the relation between a memory-one strategy and the mean distribution of an infinitely repeated game without discounting, see [123, Theorem 1.3] and* Eq. (6.4)*. Indeed, in the deterministic limit $\alpha \to \infty$ and $\beta < \infty$, the influence of p_0 on the relation between $\mathbf{p}^t$ and $\mathbf{v}$ in* Eq. (9.9) *disappears. In the deterministic limit $\alpha \to \infty$ and $\beta = \frac{\alpha(1-\delta)}{\delta}$, one recovers a relation as in Lemma 9.*

The relation in Eq. (9.9) links the mean distribution of the action profiles $\mathbf{v}$ to the time dependent memory-one strategy $\mathbf{p}^t$ and is the starting point for defining strategies that allow a single player to exert significant influence on the outcome of the uncertain repeated game. We are now ready to formulate a risk-adjusted ZD strategy for repeated games with beta discounting.

Definition 26 (**risk-adjusted ZD strategy**). *A time dependent memory one strategy* $\mathbf{p}^t$ *with entries in the closed unit interval is a risk-adjusted ZD strategy for a symmetric n-player game if there exist shape parameters* $\alpha > 0, \beta > 1$, *constants* $(s, l) \in \mathbb{R}^2$, *weights* w_j *for* $1 \leq j \leq n$ *and* ϕ *such that*

$$\mathbf{p}^t = \frac{t+\alpha+\beta}{t+\alpha}\left(\mathbf{p}^{\text{rep}} + \phi\left[sg^i - \sum_{j\neq i}^{n} w_j g^j + (1-s)l\mathbb{1}\right] - \frac{\beta-1}{\alpha+\beta-1}p_0\mathbb{1}\right), \quad (9.12)$$

under the requirement that $w_i = 0$, $\sum_{j=1}^{n} w_j = 1$ *and* $\phi \neq 0$.

Remark 21. *As detailed in Remark 9, when* $w_j = \frac{1}{n-1}$ *for all* $j \neq i$, *the formulation of a risk-adjusted ZD strategy for a symmetric social dilemma can be simplified to have 2n elements.*

***Theorem* 13** (**Enforcing a linear relation under uncertain disocunting**). *Assume the probabilistic discount factor has a fixed beta distribution with real parameters* $\alpha > 0$ *and* $\beta > 1$. *If a player applies a fixed risk-adjusted ZD strategy as in Definition 26 then, independent of the fixed strategies of the* $n-1$ *group members, expected payoffs obey the equation*

$$\pi_{-i} = s\pi_i + (1-s)l, \quad (9.13)$$

Proof. Substituting the expression for $\mathbf{p}^t$ into Eq. (9.10) we obtain

$$\frac{\beta-1}{\alpha+\beta-1}\sum_{t=0}^{\infty} d(t)v(t)\cdot\left(\phi\left[sg^i - \sum_{j\neq i}^{n} w_j g^j + (1-s)l\mathbb{1}\right] - \frac{\beta-1}{\alpha+\beta-1}p_0\mathbb{1}\right) = -\frac{\beta-1}{\alpha+\beta-1}p_0 \quad (9.14)$$

By the distributive and commutative properties of the dot product, this implies

$$\left(\phi\left[sg^i - \sum_{j\neq i}^{n} w_j g^j + (1-s)l\mathbb{1}\right] - \frac{\beta-1}{\alpha+\beta-1}p_0\mathbb{1}\right)\cdot\mathbf{v} = -\frac{\beta-1}{\alpha+\beta-1}p_0$$

$$\phi\left[s\pi_i - \sum_{j\neq i}^{n} w_j \pi_j + (1-s)l\right] - \frac{\beta-1}{\alpha+\beta-1}p_0 = -\frac{\beta-1}{\alpha+\beta-1}p_0,$$

$$\phi\left[s\pi_i - \sum_{j\neq i}^{n} w_j \pi_j + (1-s)l\right] = 0.$$

where we have used the fact that $\mathbf{v} \cdot \mathbb{1} = 1$. Finally, because $\phi \neq 0$, it follows that

$$s\pi_i - \sum_{j \neq i}^{n} w_j \pi_j + (1-s)l = 0 \xRightarrow{\pi_{-i} := \sum_{j \neq i}^{n} w_j \pi_j} s\pi_i + (1-s)l = \pi_{-i}. \tag{9.15}$$

This completes the proof. □

9.3 Existence of risk-adjusted ZD strategies

As in the case of deterministic discounting discussed in Chapter 6, the entries of risk-adjusted zero-determinant strategies are conditional probabilities and in order to obtain a well-defined strategy, they need to belong to the closed unit interval. Consequently, there are limitations on how a strategic player can choose the slope s and the baseline payoff l of the linear payoff relation.

Definition 27 (**Enforceable payoff relations under beta discounting**). *A linear relation $(s, l) \in \mathbb{R}^2$ with weights $w \in \mathbb{R}^n$, is enforceable under beta discounting if there exist real uncertainty parameters $\alpha > 0$ and $\beta > 1$ of the beta distribution, and strategy parameters $\phi > 0$ and $p_0 \in [0, 1]$ such that for all $t \geq 0$ the entries of $\mathbf{p}^t$ are in the closed unit interval.*

As we have seen in Chapter 6, the parameter $\phi > 0$ determines how fast the linear relation is enforced and plays a crucial role in determining threshold discount factors in deterministically discounted games [122, 146]. Because for beta discounting the discount rate is monotonically decreasing over time, the set of enforceable payoff relations of a risk-adjusted zero-determinant strategy is determined at time $t = 0$. This is formalized in the following lemma.

Lemma 14 (**Monotonically decreasing upper bounds**). *If the entries of $\mathbf{p}^0$ are in the closed unit interval, then also the entries of $\mathbf{p}^t$ are in the closed unit interval for all $t \geq 0$.*

Proof.

$$\begin{aligned}
&\mathbb{0} \leq \mathbf{p}^t \leq \mathbb{1}, \\
&\mathbb{0} \leq \frac{t+\alpha+\beta}{t+\alpha} \left(\mathbf{p}^{\text{rep}} + \phi \left[s g^i - \sum_{j \neq i}^{n} w_j g^j + (1-s) l \mathbb{1} \right] - \frac{\beta - 1}{\alpha + \beta - 1} p_0 \mathbb{1} \right) \leq \mathbb{1}, \\
&\mathbb{0} \leq \mathbf{p}^{\text{rep}} + \phi \left[s g^i - \sum_{j \neq i}^{n} w_j g^j + (1-s) l \mathbb{1} \right] - \frac{\beta - 1}{\alpha + \beta - 1} p_0 \mathbb{1} \leq \frac{t+\alpha}{t+\alpha+\beta} \mathbb{1}.
\end{aligned}$$

To satisfy this inequality for all $t \geq 0$ it needs to hold for the minimum upper bound in t. We continue to show that this minimum occurs at $t = 0$. To this end, observe that

$$\frac{\alpha}{\alpha+\beta} \leq \frac{t+\alpha}{t+\alpha+\beta} \xRightarrow{\alpha>0,\beta>0,t\geq 0} 0 \leq \beta t. \tag{9.16}$$

Clearly, this is satisfied for any $t \geq 0$ and $\beta > 0$. □

The result in Lemma 14 has an intuitive interpretation: for a strategic player, the possibilities for exerting control over uncertain future interactions are constrained by her *initial* possibilities for exerting control.

Lemma 14 implies that the existence problem for risk-adjusted ZD strategies with beta discounting can be solved by the implications of the inequality

$$\mathbb{0} \leq \mathbf{p}^{\text{rep}} + \phi\left[sg^i - \sum_{j\neq i}^{n} w_j g^j + (1-s)l\mathbb{1}\right] - \frac{\beta-1}{\alpha+\beta-1}p_0\mathbb{1} \leq \frac{\alpha}{\alpha+\beta}\mathbb{1}. \tag{9.17}$$

Let us first show that generous strategies cannot exist in the case of beta discounting.

Proposition 25 (**No possibilities for generosity**). *In the case of beta discounting, generous payoff relations are not enforceable in symmetric multiplayer social dilemma games.*

Proof. Suppose all players are cooperating e.g. $\sigma = (C, C, \ldots, C)$, then all players receive the one shot payoff a_{n-1}. By plugging these payoffs into the risk-adjusted ZD strategy in Definition 26, and using the fact that $\sum_{j\neq i} w_j = 1$, one obtains

$$\mathbf{p}^t(C, C, \ldots, C) = \frac{t+\alpha+\beta}{t+\alpha}\left(1 + \phi(1-s)(l - a_{n-1}) - \frac{\beta-1}{\alpha+\beta-1}p_0\right). \tag{9.18}$$

Using Lemma 14, the requirement that at $t = 0$ the entries of the risk-adjusted ZD strategy are in the unit interval implies

$$0 \leq \frac{\alpha+\beta}{\alpha}\left(1 + \phi(1-s)(l - a_{n-1}) - \frac{\beta-1}{\alpha+\beta-1}p_0\right) \leq 1, \tag{9.19}$$

$$\frac{\beta-1}{\alpha+\beta-1}p_0 - 1 \leq \phi(1-s)(l - a_{n-1}) \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0 - 1. \tag{9.20}$$

Now for the generous strategy it is required that $l = a_{n-1}$. From the above equation we obtain the requirement,

$$\frac{\beta-1}{\alpha+\beta-1}p_0 - 1 \leq 0 \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0 - 1 \tag{9.21}$$

Clearly, for any $p_0 \in [0,1]$, the lower bound is satisfied. However, the upper bound reads as

$$0 \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1} p_0 - 1. \tag{9.22}$$

Because $\beta > 1$ and $\alpha > 0$, we have $\frac{\beta-1}{\alpha+\beta-1} > 0$ and because $p_0 \in [0,1]$, a necessary condition for this inequality to hold for some $p_0 \in [0,1]$ is that it holds for $p_0 = 1$, i.e.,

$$0 \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1} - 1. \tag{9.23}$$

We proceed to show that Eq. (9.23) cannot be satisfied for $\alpha > 0$, $\beta > 1$. For this, we write Eq. (9.23) equivalently as

$$\begin{aligned} 0 \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1} - \frac{\alpha+\beta-1}{\alpha+\beta-1}&, \\ 0 \leq \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta-1}&, \\ \frac{\alpha}{\alpha+\beta-1} \leq \frac{\alpha}{\alpha+\beta}&. \end{aligned}$$

Because $\beta > 1$ and $\alpha > 0$ the left hand side is positive. Moreover, because α and β are reals they are finite and we arrive at a contradiction. This completes the proof.

Remark 22 (**Generosity in deterministic limits**). *In the deterministic limit $\alpha \to \infty$ and $\beta < \infty$, we obtain $1 \leq 1$ which is always satisfied and thus in the deterministic limit of infinitely repeated games without discounting, generous strategies can exist, which is consistent with the results in [64, 118]. If we additionally let $\beta = \frac{\alpha(1-\delta)}{\delta}$, then*

$$\lim_{\alpha\to\infty} \frac{\alpha}{\alpha + \frac{\alpha(1-\delta)}{\delta} - 1} \leq \lim_{\alpha\to\infty} \frac{\alpha}{\alpha + \frac{\alpha(1-\delta)}{\delta}} \Rightarrow \delta \leq \delta,$$

thus generous strategies also exist in the deterministic limit of exponential discounting, which is consistent with the results in [152].

□

We now continue to characterize the enforceable payoff relations of risk-adjusted zero determinant strategies. We begin by formulating necessary conditions.

Proposition 26. *In n-player symmetric social dilemma games with payoffs as in Chapter 6, the enforceable payoff relations of a risk-adjusted ZD strategy with beta* *discount factors require the following necessary conditions*

$$\phi > 0, \tag{9.24}$$

$$-\frac{1}{n-1} \leq -\underset{j\neq i}{min}\ w_j < s < 1, \tag{9.25}$$

$$b_0 \leq l < a_{n-1}. \tag{9.26}$$

Proof. In the following, we refer to the ZD strategist as player i. Let $t = 0$ and suppose all players are cooperating e.g. $\sigma = (C, C, \ldots, C)$. In this case, every player receives the one shot payoff a_{n-1}. By plugging these payoffs into the risk-adjusted ZD strategy in Definition 26, and using the fact that $\sum_{j\neq i} w_j = 1$, one obtains

$$\mathbf{p}^0(C, C, \ldots, C) = \frac{\alpha+\beta}{\alpha}\left(1 + \phi(1-s)(l - a_{n-1}) - \frac{\beta-1}{\alpha+\beta-1}p_0\right) \tag{9.27}$$

Now suppose that to the contrary all players defect; then all players receive the one shot payoff b_0 and the entry of the risk-adjusted ZD strategy is

$$\mathbf{p}^0(D, D, \ldots, D) = \frac{\alpha+\beta}{\alpha}\left(\phi(1-s)(l - b_0) - \frac{\beta-1}{\alpha+\beta-1}p_0\right) \tag{9.28}$$

The requirement that Eq. (9.27) and Eq. (9.28) belong to the closed unit interval results in the inequalities

$$\frac{\beta-1}{\alpha+\beta-1}p_0 - 1 \leq \phi(1-s)(l - a_{n-1}) \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0 - 1 < 0 \tag{9.29}$$

$$0 \leq \frac{\beta-1}{\alpha+\beta-1}p_0 \leq \phi(1-s)(l - b_0) \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0, \tag{9.30}$$

where the strict upper bound in Eq. (9.29) follows from the fact that Eq. (9.22) *cannot* be satisfied for $\alpha > 0$, $\beta > 1$ and $p_0 \in [0, 1]$. By multiplying Eq. (9.29) by -1 we obtain

$$0 < 1 - \frac{\alpha}{\alpha+\beta} - \frac{\beta-1}{\alpha+\beta-1}p_0 \leq \phi(1-s)(a_{n-1} - l) \leq 1 - \frac{\beta-1}{\alpha+\beta-1}p_0 \tag{9.31}$$

By adding Eq. (9.30) and Eq. (9.31) we obtain

$$0 < 1 - \frac{\alpha}{\alpha+\beta} \leq \phi(1-s)(a_{n-1} - b_0) \leq 1 + \frac{\alpha}{\alpha+\beta} \tag{9.32}$$

Combining this with the assumption in the main text that $a_{n-1} > b_0$, it follows that in order for the payoff relation to be enforceable it is necessary that

$$\phi(1-s) > 0. \tag{9.33}$$

Now suppose that there is a single defecting player, i.e., $\sigma = (C, C, \dots, D)$ or any of its permutations. In this case, the cooperators receive a_{n-2} and the single defector obtains b_{n-1}. In the case that the single defector is $j \neq i$, the entry of the risk-adjusted ZD strategy is

$$\mathbf{p}_\sigma^0 = \frac{\alpha+\beta}{\alpha}\left(1 + \phi[sa_{n-2} - (1-w_j)a_{n-2} - w_j b_{n-1} + (1-s)l] - \frac{\beta-1}{\alpha+\beta-1}p_0\right), \tag{9.34}$$

and if the unique defector is i, the entry of $\mathbf{p}^{\text{rep}}$ is equal to zero and thus, the entry of the risk-adjusted ZD strategy is

$$\mathbf{p}_\sigma^0 = \frac{\alpha+\beta}{\alpha}\left(\phi[sb_{n-1} - a_{n-2} + (1-s)l] - \frac{\beta-1}{\alpha+\beta-1}p_0\right). \tag{9.35}$$

The requirement that Eq. (9.34) and Eq. (9.35) belong to the closed unit interval results in the following inequalities

$$\frac{\beta-1}{\alpha+\beta-1}p_0 \leq \phi[sb_{n-1} - a_{n-2} + (1-s)l] \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0 \tag{9.36}$$

$$\begin{aligned}\frac{\beta-1}{\alpha+\beta-1}p_0 - 1 \leq \phi[sa_{n-2} - (1-w_j)a_{n-2} - w_j b_{n-1} + (1-s)l] \leq \\ \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0 - 1\end{aligned} \tag{9.37}$$

By combining these two conditions in a similar manner as was done for the homogeneous action profile case we obtain

$$0 < 1 - \frac{\alpha}{\alpha+\beta} \leq \phi[(s+w_j)(b_{n-1} - a_{n-2})] \leq 1 + \frac{\alpha}{\alpha+\beta} \tag{9.38}$$

From the assumption $b_{z+1} > a_z$ in the main text, it follows that

$$\forall j \neq i: \quad \phi(s+w_j) > 0. \tag{9.39}$$

Together with Eq. (9.33) this implies that

$$\phi(1+w_j) > 0 \xrightarrow{\exists\ j\neq i\ s.t. w_j>0} \phi > 0.$$

This also implies that

$$\phi(1-s) \overset{\phi>0}{\Longrightarrow} (1-s) > 0 \Rightarrow s < 1.$$

In combination with Eq. (9.39) it follows that

$$\forall j \neq i : s + w_j > 0 \Leftrightarrow \forall j \neq i : w_j > -s \Leftrightarrow \min_{j\neq i} w_j > -s. \tag{9.40}$$

Moreover, because it is required that $\sum_{j=1}^{n} w_j = 1$, it follows that $\underset{j \neq i}{min}\ w_j \leq \frac{1}{n-1}$. Hence the necessary condition turns into:

$$-\frac{1}{n-1} \leq -\min_{j \neq i} w_j < s < 1. \tag{9.41}$$

Let us now investigate the bounds on the baseline payoffs. From Eq. (9.30) we have

$$0 \leq \frac{\beta - 1}{\alpha + \beta - 1} p_0 \leq \phi(1-s)(l - b_0) \tag{9.42}$$

Thus, in order for Eq. (9.42) to hold it is required that

$$0 \leq \frac{\beta - 1}{\alpha + \beta - 1} p_0 \leq \phi(1-s)(l - b_0) \xrightarrow{\phi(1-s)>0} l \geq b_0. \tag{9.43}$$

From Eq. (9.29) we have

$$\phi(1-s)(l - a_{n-1}) \leq \frac{\alpha}{\alpha + \beta} + \frac{\beta - 1}{\alpha + \beta - 1} p_0 - 1 < 0 \tag{9.44}$$

It follows that it must hold that

$$l < a_{n-1}.$$

□

Let us now formulate the result that fully characterizes the enforceable payoff relations of risk-adjusted zero determinant strategies. For this let $w = (w_j) \in \mathbb{R}^{n-1}$ denote the vector of weights and let $\hat{w}_z = \min_{w_h \in w} (\sum_{h=1}^{z} w_h)$ denote the sum of the j smallest weights of $j \neq i$ and finally let $\hat{w}_0 = 0$. We note that the proof of the following theorem is very similar to the deterministic case in Chapter 6.

***Theorem* 14 (Characterizing enforceable sets under uncertain discounting).** *For the repeated n-player game with beta discount factors such that $\alpha > 0$ and $\beta > 1$ and payoffs as in the main text that satisfy the social dilemma assumptions in the main text, the payoff relation $(s, l) \in \mathbb{R}^2$ with weights $w \in \mathbb{R}^{n-1}$ is enforceable by the risk-adjusted ZD strategy in* Eq. (9.12) *if and only if $-\frac{1}{n-1} < s < 1$ and*

$$\underset{0 \leq z \leq n-1}{max} \left\{ b_z - \frac{\hat{w}_z(b_z - a_{z-1})}{(1-s)} \right\} \leq l < \underset{0 \leq z \leq n-1}{min} \left\{ a_z + \frac{\hat{w}_{n-z-1}(b_{z+1} - a_z)}{(1-s)} \right\}. \tag{9.45}$$

Proof. Let $t = 0$. In the following we refer to the key player, who is employing the ZD strategy, as player i. Let $\sigma = (x_1, \ldots, x_n)$ such that $x_k \in \mathcal{A}_k$ and let σ^C be the number of $i's$ co-players that cooperate and let $\sigma^D = n - 1 - \sigma^C$, be the number of

$i's$ co-players that defect. Also, let $|\sigma|$ be the total number of cooperators including player i. Using this notation, and using the fact that $\alpha, \beta > 0$, for some action profile σ we may write the ZD strategy in Eq. (9.12) as

$$\frac{\alpha}{\alpha+\beta}\mathbf{p}^0_\sigma = \mathbf{p}^{\text{rep}} + \phi[(1-s)(l-g^i_\sigma) + \sum_{j\neq i}^{n} w_j(g^i_\sigma - g^j_\sigma)] - \frac{\beta-1}{\alpha+\beta-1}p_0. \quad (9.46)$$

Also, note that

$$\sum_{j\neq i}^{n} w_j g^j_\sigma = \sum_{k\in\sigma^D} w_k g^k_\sigma + \sum_{h\in\sigma^C} w_h g^h_\sigma, \quad (9.47)$$

and because $\sum_{j\neq i}^{n} w_j = 1$ it holds that

$$\sum_{h\in\sigma^C} w_h = 1 - \sum_{k\in\sigma^D} w_k.$$

Additionally, note that because of the symmetric one shot payoffs, for all $h \in \sigma^C$ it holds that $g^h_\sigma = a_{|\sigma|-1}$, and for all $k \in \sigma^D$, $g^k_\sigma = b_{|\sigma|}$. It follows that Eq. (9.47) can be written as

$$\sum_{j\neq i}^{n} w_j g^j_\sigma = a_{|\sigma|-1} + \sum_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1}).$$

Accordingly, the entries of the ZD strategy $\frac{\alpha}{\alpha+\beta}\mathbf{p}^0_\sigma$ are given by Eq. (9.49). For all $\sigma \in \mathcal{S}$ we require that

$$0 \leq \mathbf{p}^0_\sigma \leq 1 \Rightarrow 0 \leq \frac{\alpha}{\alpha+\beta}\mathbf{p}^0_\sigma \leq \frac{\alpha}{\alpha+\beta}. \quad (9.48)$$

This leads to the inequalities in Eq. (9.50) and Eq. (9.51). Because $\phi > 0$ can be chosen arbitrarily small, the inequalities in Eq. (9.50) can be satisfied for some $\alpha > 0$ and $\beta > 1$ and $p_0 \in [0,1]$ if and only if for all σ such that $x_i = C$ the inequalities in Eq. (9.52) are satisfied.

$$\frac{\alpha}{\alpha+\beta}\mathbf{p}_\sigma = \begin{cases} 1+\phi\left[(1-s)(l-a_{|\sigma|-1}) - \sum\limits_{j\in\sigma^D} w_j(b_{|\sigma|}-a_{|\sigma|-1})\right] - \frac{\beta-1}{\alpha+\beta-1}p_0, & \text{if } x_i = C, \\ \phi\left[(1-s)(l-b_{|\sigma|}) + \sum\limits_{j\in\sigma^C} w_j(b_{|\sigma|}-a_{|\sigma|-1})\right] - \frac{\beta-1}{\alpha+\beta-1}p_0, & \text{if } x_i = D. \end{cases} \quad (9.49)$$

By substituting Eq. (9.49) into this requirement we obtain that for all σ such that

$x_i = C$ the following inequalities are required to hold:

$$0 < 1 - \frac{\alpha}{\alpha+\beta} - \frac{\beta-1}{\alpha+\beta-1}p_0 \leq \phi \left[(1-s)(a_{|\sigma|-1} - l) + \sum_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1})\right] \leq 1 - \frac{\beta-1}{\alpha+\beta-1}p_0. \quad (9.50)$$

On the other hand, when $x_i = D$ the following inequalities are required to hold:

$$0 \leq \frac{\beta-1}{\alpha+\beta-1}p_0 \leq \phi \left[(1-s)(l - b_{|\sigma|}) + \sum_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1})\right] \leq \frac{\alpha}{\alpha+\beta} + \frac{\beta-1}{\alpha+\beta-1}p_0. \quad (9.51)$$

Let us first derive the conditions that result from Eq. (9.50). From the lower bound we obtain

$$0 < (1-s)(a_{|\sigma|-1} - l) + \sum_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1}). \quad (9.52)$$

The inequality in Eq. (9.52) together with the necessary condition that $(1-s) > 0$ implies that

$$a_{|\sigma|-1} + \frac{\sum_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} > l, \quad (9.53)$$

and thus provides an upper-bound on the enforceable baseline payoff l. We now turn our attention to the inequalities in Eq. (9.51) that can be satisfied if and only if for all σ such that $x_i = D$ the following holds

$$\begin{aligned} 0 &\leq (1-s)(l - b_{|\sigma|}) + \sum_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1}) \\ &\xrightarrow{(1-s)>0} b_{|\sigma|} - \frac{\sum_{j\in\sigma^C} w_j(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \leq l. \end{aligned} \quad (9.54)$$

Combining Eq. (9.54) and Eq. (9.53) we obtain

$$\begin{aligned} \max_{|\sigma| s.t. x_i = D} \left\{ b_{|\sigma|} - \frac{\sum_{l\in\sigma^C} w_l(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \right\} &\leq l, \\ l < \min_{|\sigma| s.t. x_i = C} \left\{ a_{|\sigma|-1} + \frac{\sum_{k\in\sigma^D} w_k(b_{|\sigma|} - a_{|\sigma|-1})}{(1-s)} \right\}&. \end{aligned} \quad (9.55)$$

Because $b_{|\sigma|} - a_{|\sigma|-1} > 0$ and $(1-s) > 0$ the minima and maxima of the bounds in Eq. (6.42) are achieved by choosing the w_j as small as possible. That is, the extrema of the bounds on l are achieved for those states $\sigma|_{x_i=C}$ in which $\sum_{l\in\sigma^C} w_l$ is minimum and those $\sigma|_{x_i=D}$ in which $\sum_{k\in\sigma^D} w_k$ is minimum. By the above reasoning, Eq. (9.55) can be equivalently written as in the theorem in the main statement. This completes the proof.

□

Theorem 14 provides insights in the existence of risk-adjusted zero determinant strategies in repeated games with uncertain discounting. In particular, extortionate strategies, the baseline payoff is equal to the full defection payoff, that is $l = b_0$. In the lower bound of Eq. (9.45), this occurs when $z = 0$ and due to the conservative lower bound on the baseline payoff, one can conclude that extortionate strategies can exist in multiplayer social dilemmas with beta discounting. The multiplayer social dilemma thus remains vulnerable to extortionate behaviors even under uncertainty. Likewise, equalizing strategies with a slope $s = 0$ can be enforced as long as their baseline payoff l satisfies the lower and upper bounds. A crucial implication of uncertainty follows from the *strict* upper bound in Eq. (9.45), that implies that generous strategies, for which $l = a_{n-1}$, *cannot* exist in multiplayer social dilemmas. This has a revealing intuitive interpretation: if future interactions and their payoffs are uncertain, then one cannot guarantee that others will do well. The characterization of the enforceable payoff relations can also be applied to the repeated prisoners dilemma game. To see this, let $n = 2$, $b_1 = \mathcal{T}$, $b_0 = \mathcal{P}$, $a_1 = \mathcal{R}$, and finally $a_0 = \mathcal{S}$. Then, the enforceable slopes satisfy $-1 < s < 1$ and the enforceable payoff relations must satifsy

$$\max\left\{\mathcal{P}, \frac{\mathcal{S} - \mathcal{T}s}{1-s}\right\} \leq l < \min\left\{\mathcal{R}, \frac{\mathcal{T} - \mathcal{S}s}{1-s}\right\}.$$

It follows that the mutual cooperation baseline payoff $l = \mathcal{R}$ cannot be enforced and hence generous strategies do not exist in the repeated prisoners dilemma with beta discounting.

9.4 Uncertainty and the level of influence

The characterization of enforceable payoff relations does not specify conditions on the shape parameters other than $\alpha > 0$ and $\beta > 1$. But how do these shape parameters affect the payoff relations that a strategic player can enforce? When future interactions are at least as likely as a termination of the game, the beta distribution is symmetric or

Table 9.1: Existence of zero-determinant strategies for multiplayer social dilemma games without discounting, deterministic exponential discounting and probabilistic beta discounting.

ZD strategy	No discounting	Exponential	Beta
Fair	✓	✗	✗
Generous	✓	✓	✗
Extortionate	✓	✓	✓
Equalizing	✓	✓	✓

negatively skewed such that $\alpha \geq \beta > 1$. In this case, the *mean* of the beta distribution

$$\mu = \frac{\alpha}{\alpha + \beta},$$

is at least a half. If discounting would be deterministic, the players, in this case, would expect at least two rounds of play. In the following, we provide a simple condition on one shot payoffs and the ZD parameters s and l that suggest that in many social dilemmas, in order to enforce a payoff relation it is, in fact, required that $\alpha \geq \beta$. For any distribution with $\alpha < \beta$, the mean discount factor would simply not allow a player to exert enough influence because payoffs are discounted too fast. This additional requirement on the shape parameters of the beta distribution also provides insight into *how* uncertain a strategic player can be about the discount rate or continuation probability before losing the possibility to enforce some desired payoff relation. For $\alpha \geq \beta > 1$ the variance of the beta distribution is monotonically decreasing in α. Consequently, the *maximum* variance that a risk-adjusted zero-determinant strategy can handle in these situations occurs when $\alpha = \beta > 1$, and evaluates as

$$\sigma^2_{\max} = \frac{1}{4(2\beta + 1)} < \frac{1}{12}.$$

Now let us suppose the strategic player has estimated the shape parameters of the beta distribution. Then, exactly how extortionate can a payoff relation be? It is exactly here where the parameter $\phi > 0$ plays a crucial role for the level of influence of the strategic player. In particular, for a given μ, in order for the risk-adjusted zero-determinant strategy to be well-defined, additional requirements on $\phi > 0$ are necessary that in turn determine if the strategic player can enforce the linear payoff relation *fast enough.* This is formalized in the following theorem which is related to the deterministic case in Theorem 9.

***Theorem* 15** (Mean discount rates and the level of influence)**.** *Assume* $p_0 = 0$ *and* $(s, b_0) \in \mathbb{R}^2$ *satisfy the conditions in Theorem 14. Then* $\overline{\rho}^C > 0$ *and* $\overline{\rho}^D + \underline{\rho}^C > 0.$

Moreover, the threshold mean μ *above which the extortionate payoff relation can be enforced is given by*

$$\mu_{\mathcal{T}} = max\left\{\frac{\overline{\rho}^C - \underline{\rho}^C}{\overline{\rho}^C}, \frac{\overline{\rho}^D}{\overline{\rho}^D + \underline{\rho}^C}\right\}.$$

Proof. For $\alpha > 0, \beta > 1$ for the existence of extortionate payoff relations with $l = b_0$ we know $p_0 = 0$ is required (this fact immediately follows from the lower bound in Eq. (9.30)). By substituting this into Eq. (9.50) it follows that in order for the payoff relation to be enforceable it is required that for all σ such that $x_i = C$ the following holds:

$$0 < 1 - \frac{\alpha}{\alpha+\beta} \leq \phi\left[(1-s)(a_{|\sigma|-1} - l) + \sum_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1})\right] \leq 1.$$

$$\rho^C(\sigma) := (1-s)(a_{|\sigma|-1} - l) + \sum_{j\in\sigma^D} w_j(b_{|\sigma|} - a_{|\sigma|-1}) > 0. \tag{9.56}$$

Hence, Eq. (6.37) with $p_0 = 0$ implies that for all σ such that $x_i = C$ it holds that

$$\frac{1-\mu}{\rho^C(\sigma)} \leq \phi \leq \frac{1}{\rho^C(\sigma)} \Rightarrow \frac{1-\mu}{\underline{\rho}^C(z,\hat{w}_z)} \leq \phi \leq \frac{1}{\overline{\rho}^C(z,\tilde{w}_z)}. \tag{9.57}$$

Naturally, $\overline{\rho}^C \geq \underline{\rho}^C$. In the special case in which equality holds, it follows from Eq. (7.3) that $\mu \geq 0$, which is satisfied for any $\alpha, \beta > 0$. We continue to investigate the case in which $\overline{\rho}^C > \underline{\rho}^C$. In this case, a solution to Eq. (9.57) for some $\phi > 0$ exists if and only if

$$\frac{1-\mu}{\underline{\rho}^C(z,\hat{w}_z)} \leq \frac{1}{\overline{\rho}^C(z,\tilde{w}_z)} \Rightarrow \mu \geq \frac{\overline{\rho}^C - \underline{\rho}^C}{\overline{\rho}^C}, \tag{9.58}$$

which leads to the first expression in the theorem. Now, from Eq. (9.51) with $p_0 = 0$, it follows that in order for the payoff relation to be enforceable it is necessary that

$$\forall\sigma \text{ s.t. } x_i = D: \quad 0 \leq \phi\rho^D(\sigma) \leq \mu \Rightarrow 0 \leq \phi\overline{\rho}^D(z,\tilde{w}_z) \leq \mu. \tag{9.59}$$

Because $\phi > 0$ is necessary for the payoff relation to be enforceable, it follows that $\rho^D(\sigma) \geq 0$ for all σ such that $x_i = D$. Let us first investigate the special case in which $\overline{\rho}^D(z,\tilde{w}_z) = 0$. Then Eq. (9.59) is satisfied for any $\phi > 0$ and $\mu \in (0,1)$. Now, assume $\overline{\rho}^D(z,\tilde{w}_z) > 0$. Then, Eq. (9.59) and Eq. (9.57) imply

$$\frac{1-\mu}{\underline{\rho}^C(z,\hat{w}_z)} \leq \phi \leq \frac{\mu}{\overline{\rho}^D(z,\tilde{w}_z)}. \tag{9.60}$$

In order for such a ϕ to exist it needs to hold that

$$\frac{1-\mu}{\underline{\rho}^C(z,\hat{w}_z)} \leq \frac{\mu}{\overline{\rho}^D(z,\tilde{w}_z)} \xRightarrow{\overline{\rho}^D,\ \underline{\rho}^C>0} \mu \geq \frac{\overline{\rho}^D}{\overline{\rho}^D+\underline{\rho}^C}. \tag{9.61}$$

This completes the proof. □

Corollary 9 (symmetric and negatively skewed distributions)**.** *If $\underline{\rho}^C = \overline{\rho}^D$, enforceable payoff relations require $\frac{\alpha}{\alpha+\beta} \geq \frac{1}{2}$ and hence $\alpha \geq \beta$.*

Relation to deterministic discounting and generous strategies

Numerical examples of Theorem 15 for the n-player snowdrift game and n-player linear public goods game are shown in Figure 9.3. These figures are related to Propositions 9 and 6 in which existence conditions for extortionate strategies in these games are provided. The n-player linear public goods game and n-player snowdrift game are also illustrative examples of how uncertainty in the probability for a future interaction can influence opportunities to enforce *generous* payoff relations. In the case of deterministic discounting, in the n-player linear public goods game, generous strategies can enforce the same slopes as extortionate strategies. For n-player snowdrift games it can even be shown that a generous strategist can enforce *any* slope $0 < s < 1$ provided that $\delta < 1$ is sufficiently high. In this deterministic setting, it is the fixed discount factor δ that determines one's possibilities for the level of control. There is however a subtle but *crucial* difference between the effects of μ and δ: only in the deterministic limit can one enforce generous payoff relations in multiplayer social dilemmas games.

9.5 Final Remarks

The discovery of zero-determinant strategies by Press and Dyson [64] showed that in the absence of discounting, individuals can deterministically exert control over the outcome of 2×2 games without imposing any restriction on the strategy of the other player. This surprising finding motivated others to investigate how such strategies hold up under a variety of circumstances [114, 116, 118, 152]. Zero-determinant strategies were first studied within the traditional deterministic discounting framework in [121]. One of the conclusions was that with discounting, the strategic player's initial probability to cooperate remains important for her opportunities to influence the outcome of the game. Perhaps a more important conclusion was that, fair strategies, which enforce an equal payoff for everyone, do not exist in games with finite but undetermined time horizons. The existence of extortionate, generous and equalizing strategies however remained unchanged. And thus, even with a finite expected number of rounds, a ZD strategy could promote or maintain cooperative behavior.

Also in an evolutionary setting, generous strategies have since been studied both theoretically and empirically for their ability to *maintain* cooperation [114, 135, 146]. However, independent of how one interprets the discount rate in traditional models of repeated games, in many real-world scenarios it is likely that there is some degree of *uncertainty*. Indeed, interest rates are subject to change over time and decision makers do not always know the exact probability of a following mutual interaction.

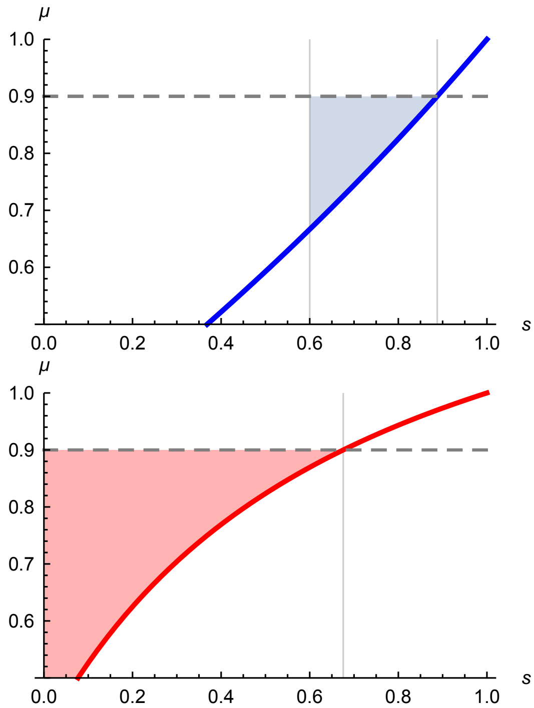


Figure 9.3: Numerical examples of enforceable slopes by extortionate strategies when $\mu = \frac{9}{10}$. Top: In an n-player snowdrift game with $b = \frac{5}{4}$, $c = 1$ and $n = 3$ extortionate strategies can only enforce slopes after the vertical line at $s = 1 - \frac{c}{b(n-1)} = \frac{6}{10}$. Every slope s for which the blue curve is below μ is enforceable, this is indicated by the blue region. Bottom: in a linear public goods game with $r = \frac{5}{4}$, $c = 1$ and $n = 3$. extortionate strategies can enforce any slope s for which the red curve is under $\mu = \frac{9}{10}$, this is indicated by the red region.

This uncertainty in turn can influence decisions made during the repeated interactions [147, 153]. Herein, we have extended the theory of repeated n-player games in two directions. First, from a more general perspective, we provided a unifying framework of discounting in repeated games that is able to capture infinitely repeated games without and with traditional deterministic exponential discounting *and* can capture the uncertainty in the discount rate or continuation probability. This additional layer of psychological complexity can be useful in predicting real-world behaviors.

From an empirical point of view, it can be interesting to investigate how the effective discounting function of this generalized framework holds up in laboratory or field experiments of human interaction, knowing the promising fact that it supports the monotone time-inconsistency property of hyperbolic discounting. From a more theoretical point of view, it can be interesting to investigate classic folk theorems in this uncertain setting, or to extend the proposed framework to individual beliefs about the discount rate. For instance, what would happen if there is only one individual that is uncertain about the continuation probability?

In addition, we have extended the theory of zero determinant strategies to repeated games with uncertain discount rates or continuation probabilities. We have shown how the mean discount factor can influence one's level of control in repeated interactions and how the amount of uncertainty affects one's possibility to exert control. An important consequence is that generous strategies seize to exist in this uncertain setting. In some sense this theoretical finding is in line with the conclusions of [147]. Namely, when a witty strategic player aims at enforcing a generous payoff relation, if the uncertain co-players tend to cooperate even when they "should not", their increased tendency to cooperate prevents them to profit maximally from the strategic player's generous actions. Consequently, the generous strategist *cannot* enforce that her co-players do better than herself. In sharp contrast, when this witty strategic player employs an extortionate strategy, she *can* enforce that others are worse off.

Chapter 10

Conclusion and Future Research

> It is our task, both in science and in society at large, to prove the conventional wisdom wrong and to make our unpredictable dreams come true
>
> *Freeman Dyson*

10.1 Conclusion

Part I: Rationality and social influence in network games

Based on theoretical and behavioral economics in Chapter 3, we proposed novel dynamics for finite and convex network games that result from an intuitive mix of rational best responses and social learning. We have shown that for a general class of games these dynamics converge to a generalized Nash equilibrium and that the corresponding decision-making process is "compatible" with rational best response dynamics. That is, a mix of best responders and h-relative best responders will eventually reach an equilibrium action profile. These results make it possible to rigorously study how relative performance considerations of "irrational" or *conforming* decision makers affect the behavior and equilibrium profiles of complex socio-technical and socio-economic processes. Considering these effects is important because many technological challenges require increasingly complex models of large social systems that, in reality, are often affected by social learning effects that are not captured by best response dynamics.

In Chapter 4, we have shown that rational imitation dynamics in a general class of asynchronous public goods games on networks converge to an imitation equilibrium in finite time. By means of a counter-example, we have shown that this general case of convergence is not guaranteed when imitation is *unconditional*. For regular spatial structures and linear production functions, we have proven convergence either directly from the payoff functions or by using an algorithmic proof technique that takes advantage of the regularity of the network. We have shown that in the case of rational imitation, convergence is also guaranteed when the group structures are determined by a bipartite graph. Such a representation of a spatially structured social dilemma can, for instance, be used when the group structures are obtained from data that does not contain information about the entire social network. Next to convergence, we have provided evidence that in contrast to best response dynamics, rational imitation can effectively facilitate the evolution of cooperation via network reciprocity. Our results indicate that through the combination of rationality and imitation, beneficial dynamic features can arise that are able to sustain the availability of a publicly available good, providing new insights in the design of solutions to the tragedy of the commons.

In Chapter 5, we have shown how network games can be extended to include a subset of players that can employ different actions against different opponents. When the local games in the network admit a weighted potential function, convergence of the strategically differentiated version with myopic best response dynamics is guaranteed. For both imitation and best response dynamics the topology of the network, the existence and location of differentiators in the networks can crucially alter the action profile at an equilibrium of groupwise public goods games. When differentiators are plentiful, the equilibrium action profile becomes less sensitive to changes in the values of the payoff parameters and cooperation can exist for very low values benefit-to-cost ratios.

Part II: Strategic play and control in repeated games

In Chapter 6, we have extended the existing results for ZD strategies in repeated two-player two-action games to n-player two-action games. We focused on n-player social dilemma games because of their importance to the current literature. However, the fundamental relation between the memory-one strategy and the limit distribution is independent of the structure of the game and thus the results in this chapter can be extended by considering n-player games that are not social dilemmas. Our theory supports the finding that due to the finite number of expected rounds or discounting of the payoffs, the initial probability to cooperate of the key player remains important, and we have shown that for the existence of generous strategies the ZD strategist must start to cooperate with probability one. Likewise, for extortionate strategies, this initial probability must be zero. These results indicate that even in large groups of players, a single player can unilaterally enforce the mutual cooperation payoff independent of the strategies of the other players. Especially the latter is an important feature of our results that distinguishes them from classical folk theorems in which it is assumed all players are *rational*. If one however assumes that other players are rational, the positive payoff relations that generous and extortionate ZD strategies enforce ensure that the collective best response of the selfish co-players is to maximize the ZD strategists payoff by cooperating every round.

In Chapter 7, a theory is developed that characterizes the efficiency of exerting control in terms of the minimum required number of expected interactions in social dilemmas. Based on the necessary conditions on the initial probability to cooperate, we derived expressions for the minimum discount factors above which a ZD strategist can enforce some desired generous or extortionate payoff relation. Because equalizer strategies do not impose such conditions on the initial probability to cooperate, one can identify a multitude of p_0 regions in the unit interval for which there exist different threshold discount factors. Consequently, we have derived an expression that ensures the desired equalizer strategy to be enforceable for any initial probability to cooperate in the open unit interval. The derived necessary and sufficient conditions for existence and the thresholds discount factors presented in this chapter may also be helpful in designing novel control techniques for repeated decision making processes in which the objective is to achieve a desired relative performance within a given number of rounds.

In Chapter 8 we have studied the evolutionary stability of ZD strategies in a finite population in which players interact in randomly formed n-player repeated contests. Necessary and sufficient conditions are provided for a resident ZD strategy to be stable with respect to a single mutant ZD strategy. These conditions can characterize when ZD strategies can enforce cooperation to evolve in a finite population. In particular, they suggest that under the classic Maynard-Smith conditions ($N = \infty, n = 2$) extortionate strategies cannot be evolutionarily stable. In this case, only generous strategies and equalizers with generous slopes are favored by evolution. In sharp contrast, when the population size is equal to the group size ($n = N$), only extortionate strategies can be evolutionarily stable. In a finite population in which the group size of the contests is smaller than the population size ($n < N$) both generosity and extortion can be stable, however, this highly depends on the benefit-to-cost ratio, the population size N, and group size of the contests n.

In Chapter 9 we have proposed a novel discounting method for repeated games that takes into account the added psychological complexity of uncertainty about the discount factor or continuation probability of the repeated game. With this generalized discounting framework it is possible to recover deterministic discounting methods that exist in the current literature, as well as hyperbolic discounting that has time-inconsistent discount rates. We have shown how ZD strategies, that are normally fixed memory-one strategies, can be adapted to time-varying memory-one strategies that take into account the changing discount rates that result from uncertainty in the continuation probability. In deterministic limits these novel risk-adapted ZD strategies recover the formulations of ZD strategies under deterministic discounting methods and ZD strategies for repeated games with an infinite number of expected rounds. Characterization of the enforceable slopes shows that in this uncertain setting, generous strategies cannot be enforced. This result highlights that certain continuation probabilities are necessary for mutual cooperation to be enforceable by a strategic player.

10.2 Recommendations for future research

Part I: Rationality and social influence in network games

For the h-RBR dynamics proposed in Part I many challenging open problems and future research directions can be identified. In this work, we have focused on deterministic decision-makers that do not deviate from their decision rule. In reality, trembling hands [154] or random explorations are inevitable. For myopic best response dynamics, these effects have been studied under a variety of noise models such as constant noise as in adaptive play [45], or a noise that is proportional to one's *expected payoff* as in the log-linear response model [45]. These ideas can also be applied to relative best responses, and it is interesting to characterize how stochastically stable equilibria may change under the influence of social learning. For rational imitation dynamics, proportional noise models from imitation processes may be incorporated [37, 155] as well. However, even for deterministic dynamics, the effects of social influence and network structure on the equilibria of network games are not yet fully characterized. In particular, it could be interesting to identify network structures that enhance the opportunities for rational cooperation to evolve in social dilemmas on networks. The same holds for the mechanism strategic differentiation.

From a more technical perspective, it would be interesting to study the convergence properties of *synchronous* h-RBR and rational imitation dynamics. In this case, the existence of a potential function is not immediately helpful in the convergence analysis and because the constrained sets imposed by the h-RBR dynamics are not jointly convex, new analysis techniques must be developed to characterize the convergence properties of this class of synchronous dynamics. Nevertheless, it can also be interesting to study the effect that synchronous revisions have on the effectiveness of network reciprocity under rational imitation and h-RBR dynamics. Finally, it would be interesting to apply the idea of relative performance considerations in rational imitations and relative best responses to opinion dynamics. In this case, the relative performance of the players can, for instance, be associated with the differences in opinions (3). In such a model, players will only take into account the opinions of neighbors that are *relatively close* to their own opinion. In particular, it could be of interest to investigate under which conditions polarization or clustering of opinions will occur in such "relative" opinion dynamics.

Part II: Strategic play and control in repeated games

The theory developed in Part II can be extended in several ways. We will begin with the most immediate extensions. We have seen how uncertainty in the continuation probability or discount rate prevents the opportunities of an individual to enforce a generous payoff relation, but it remains an open problem how the parameters of the probability distribution affect the equilibrium payoffs of repeated games. In particular, the uncertain discounting framework from Chapter 9 can used to investigate how classical folk theorems hold up under uncertain discounting.

In chapter 6, we have focused on characterizing the enforceable payoff relations in social dilemmas, but as mentioned before, the fundamental relation between memory-one strategies and mean distributions of the repeated game do not require the social dilemma assumptions. In particular, the results in [63] indicate that ZD strategies possibly exist in the class of symmetric *potential games*. It would be interesting to extend the theory in this thesis by characterizing the enforceable payoff relations in this widely studied class of games.

Perhaps a more difficult research direction is to extend the theory of ZD strategies in repeated games with *individual discount factors*. Indeed, the folk theorem has been studied under these rather complex settings [53].

It would be also interesting to include other sources of psychological complexity and uncertainty in n-player games. For instance, psychologists and game theorists have recently studied the effect of uncertainty in the group size n [156, 157]. It will be interesting to study how this will affect the strategic behavior of individuals in repeated games. Lastly, an interesting and challenging direction for future research is to study ZD strategies in *continuous-time* repeated games [158]. For this, a new analysis tool must be developed first. In particular, it is not yet clear how mean distributions of continuous time stochastic processes can be related to a strategy of a strategic player in continuous time repeated games. With such a relation, one could apply continuous-time discounting methods [150] to repeated games. This, in turn, would allow us to formalize how the time "spent" in a certain action profile can affect the strategic decisions of individuals.

Bibliography

[1] P. A. Van Lange, J. Joireman, C. D. Parks, and E. Van Dijk, "The psychology of social dilemmas: A review," *Organizational Behavior and Human Decision Processes*, vol. 120, no. 2, pp. 125–141, 2013.

[2] G. Hardin, "The tragedy of the commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968.

[3] R. Hardin, "Collective action as an agreeable n-prisoners' dilemma," *Behavioral Science*, vol. 16, no. 5, pp. 472–481, 1971.

[4] W. F. Lloyd, *Two Lectures on the Checks to Population: Delivered Before the University of Oxford, in Michaelmas Term 1832*. JH Parker, 1833.

[5] J.-F. Bonnefon, A. Shariff, and I. Rahwan, "The social dilemma of autonomous vehicles," *Science*, vol. 352, no. 6293, pp. 1573–1576, 2016.

[6] E. Ostrom, *Governing the commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, 1990.

[7] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[8] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[9] J. Nash, "Non-cooperative games," *Annals of Mathematics*, pp. 286–295, 1951.

[10] J. F. Nash *et al.*, "Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[11] T. Yamagishi, "Exit from the group as an individualistic solution to the free rider problem in the united states and japan," *Journal of Experimental Social Psychology*, vol. 24, no. 6, pp. 530–542, 1988.

[12] H. Brandt, C. Hauert, and K. Sigmund, "Punishment and reputation in spatial public goods games," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. 1519, pp. 1099–1104, 2003.

[13] E. Fehr and S. Gächter, "Altruistic punishment in humans," *Nature*, vol. 415, no. 6868, pp. 137–140, 2002.

[14] D. Balliet, L. B. Mulder, and P. A. Van Lange, "Reward, punishment, and cooperation: a meta-analysis." *Psychological Bulletin*, vol. 137, no. 4, pp. 594–615, 2011.

[15] H. Ohtsuki and Y. Iwasa, "How should we define goodness?– reputation dynamics in indirect reciprocity," *Journal of Theoretical Biology*, vol. 231, no. 1, pp. 107–120, 2004.

[16] H. Brandt and K. Sigmund, "The logic of reprobation: assessment and action rules for indirect reciprocation," *Journal of Theoretical Biology*, vol. 231, no. 4, pp. 475–486, 2004.

[17] R. Alexander, *The Biology of Moral Systems.* Taylor & Francis Inc, 2017.

[18] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity by image scoring," *Nature*, vol. 393, no. 6685, pp. 573–577, 1998.

[19] C. Wedekind and M. Milinski, "Cooperation through image scoring in humans," *Science*, vol. 288, no. 5467, pp. 850–852, 2000.

[20] D. G. Rand and M. A. Nowak, "Human cooperation," *Trends in Cognitive Sciences*, vol. 17, no. 8, pp. 413–425, 2013.

[21] W. D. Hamilton, "The genetical evolution of social behaviour. ii," *Journal of Theoretical Biology*, vol. 7, no. 1, pp. 17–52, 1964.

[22] K. R. Foster, T. Wenseleers, and F. L. Ratnieks, "Kin selection is the key to altruism," *Trends in Ecology & Evolution*, vol. 21, no. 2, pp. 57–60, 2006.

[23] S. A. Frank, *Foundations of social evolution.* Princeton University Press, 1998.

[24] R. Dawkins, *28. The Selfish Gene.* Oxford University Press, 1976.

[25] E. O. Wilson, *Sociobiology.* Harvard University Press, 2000.

[26] A. Traulsen and M. A. Nowak, "Evolution of cooperation by multilevel selection," *Proceedings of the National Academy of Sciences*, vol. 103, no. 29, pp. 10 952–10 955, 2006.

[27] D. S. Wilson, "A theory of group selection," *Proceedings of the national academy of sciences*, vol. 72, no. 1, pp. 143–146, 1975.

[28] P. B. Rainey and K. Rainey, "Evolution of cooperation and conflict in experimental bacterial populations," *Nature*, vol. 425, pp. 72–74, 2003.

[29] J. M. Smith, *Evolution and the Theory of Games.* Cambridge University Press, 1982.

[30] M. A. Nowak and K. Sigmund, "Evolutionary dynamics of biological games," *Science*, vol. 303, no. 5659, pp. 793–799, 2004.

[31] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bulletin of the American Mathematical Society*, vol. 40, no. 4, pp. 479–519, 2003.

[32] J. W. Weibull, *Evolutionary game theory.* MIT press, 1997.

[33] R. Cressman, C. Ansell, and K. Binmore, *Evolutionary dynamics and extensive form games.* MIT Press, 2003, vol. 5.

[34] E. Lieberman, C. Hauert, and M. A. Nowak, "Evolutionary dynamics on graphs," *Nature*, vol. 433, pp. 312–316, 2005.

[35] R. M. May, "Network structure and the biology of populations," *Trends in Ecology & Evolution*, vol. 21, no. 7, pp. 394–399, 2006.

[36] M. P. Hassell, H. N. Comins, and R. M. May, "Species coexistence and self-organizing spatial dynamics," *Nature*, vol. 370, pp. 290–292, 1994.

[37] G. Szabó and G. Fath, "Evolutionary games on graphs," *Physics Reports*, vol. 446, no. 4-6, pp. 97–216, 2007.

[38] B. Allen and M. A. Nowak, "Games on graphs," *EMS Surveys in Mathematical Sciences*, vol. 1, no. 1, pp. 113–151, 2014.

[39] M. Nowak and R. Highfield, *Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed.* Simon and Schuster, 2011.

[40] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," *Nature*, vol. 441, pp. 502–505, 2006.

[41] H. A. Simon, *Models of Bounded Rationality: Empirically Grounded Economic Reason.* MIT Press, 1997, vol. 3.

[42] D. Monderer and L. S. Shapley, “Potential games,” *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[43] P. Dubey, O. Haimanko, and A. Zapechelnyuk, “Strategic complements and substitutes, and potential games,” *Games and Economic Behavior*, vol. 54, no. 1, pp. 77–94, 2006.

[44] M. Voorneveld, “Best-response potential games,” *Economics Letters*, vol. 66, no. 3, pp. 289–295, 2000.

[45] H. P. Young, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, 2001.

[46] J. R. Marden and J. S. Shamma, “Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation,” *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.

[47] S. Grammatico, “Proximal dynamics in multi-agent network games,” *IEEE Transactions on Control of Network Systems*, vol. 5, pp. 1707–1716, 2018.

[48] L. Pavel, *Game theory for control of Optical Networks*. Springer Science & Business Media, 2012.

[49] P. Yi and L. Pavel, “An operator splitting approach for distributed generalized nash equilibria computation,” *Automatica*, vol. 102, pp. 111–121, 2019.

[50] D. Helbing, M. Schönhof, H.-U. Stark, and J. A. Hołyst, “How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the prisoner’s dilemma,” *Advances in Complex Systems*, vol. 8, no. 01, pp. 87–116, 2005.

[51] M. M. Flood, “Some experimental games,” *Management Science*, vol. 5, no. 1, pp. 5–26, 1958.

[52] R. L. Trivers, “The evolution of reciprocal altruism,” *The Quarterly Review of Biology*, vol. 46, no. 1, pp. 35–57, 1971.

[53] D. Fudenberg and J. Tirole, *Game theory, 1991*. MIT Press, 1991.

[54] D. M. Kreps, P. Milgrom, J. Roberts, and R. Wilson, “Rational cooperation in the finitely repeated prisoners’ dilemma,” *Journal of Economic Theory*, vol. 27, no. 2, pp. 245–252, 1982.

[55] J. W. Friedman, “A non-cooperative equilibrium for supergames,” *The Review of Economic Studies*, vol. 38, no. 1, pp. 1–12, 1971.

[56] C. Hilbe, K. Chatterjee, and M. A. Nowak, "Partners and rivals in direct reciprocity," *Nature Human Behaviour*, pp. 469–477, 2018.

[57] R. Selten and P. Hammerstein, "Gaps in harley's argument on evolutionarily stable learning rules and in the logic of ?tit for tat?" *Behavioral and Brain Sciences*, vol. 7, no. 1, pp. 115–116, 1984.

[58] R. Boyd and J. P. Lorberbaum, "No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game," *Nature*, vol. 327, pp. 58–59, 1987.

[59] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[60] M. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game," *Nature*, vol. 364, pp. 56–58, 1993.

[61] P. Molander, "The optimal level of generosity in a selfish, uncertain environment," *Journal of Conflict Resolution*, vol. 29, no. 4, pp. 611–618, 1985.

[62] C. Hilbe, L. A. Martinez-Vaquero, K. Chatterjee, and M. A. Nowak, "Memory-n strategies of direct reciprocity," *Proceedings of the National Academy of Sciences*, vol. 114, no. 18, pp. 4715–4720, 2017.

[63] P. Duersch, J. Oechssler, and B. C. Schipper, "When is tit-for-tat unbeatable?" *International Journal of Game Theory*, vol. 43, no. 1, pp. 25–36, 2014.

[64] W. H. Press and F. J. Dyson, "Iterated prisoner?s dilemma contains strategies that dominate any evolutionary opponent," *Proceedings of the National Academy of Sciences*, vol. 109, no. 26, pp. 10 409–10 413, 2012.

[65] P. Yi and L. Pavel, "A distributed primal-dual algorithm for computation of generalized nash equilibria via operator splitting methods," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 3841–3846.

[66] L. E. Blume, "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, vol. 5, no. 3, pp. 387–424, 1993.

[67] G. Ellison, "Learning, local interaction, and coordination," *Econometrica: Journal of the Econometric Society*, pp. 1047–1071, 1993.

[68] M. A. Nowak and R. M. May, "The spatial dilemmas of evolution," *International Journal of Bifurcation and Chaos*, vol. 3, no. 01, pp. 35–78, 1993.

[69] H. P. Young, "The evolution of conventions," *Econometrica*, pp. 57–84, 1993.

[70] M. Friedman and M. Friedman, *Essays in Positive Economics.* University of Chicago Press, 1953.

[71] L. Zino, G. Como, and F. Fagnani, "On imitation dynamics in potential population games," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC).* IEEE, 2017, pp. 757–762.

[72] J. Barreiro-Gomez and H. Tembine, "Constrained evolutionary games by using a mixture of imitation dynamics," *Automatica*, vol. 97, pp. 254–262, 2018.

[73] A. Govaert, P. Ramazi, and M. Cao, "Convergence of imitation dynamics for public goods games on networks," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC).* IEEE, 2017, pp. 4982–4987.

[74] D. K. Levine and W. Pesendorfer, "The evolution of cooperation through imitation," *Games and Economic Behavior*, vol. 58, no. 2, pp. 293–315, 2007.

[75] F. Vega-Redondo, "The evolution of Walrasian behavior," *Econometrica*, vol. 65, no. 2, pp. 375–384, 1997.

[76] A. A. Alchian, "Uncertainty, evolution, and economic theory," *Journal of Political Economy*, vol. 58, no. 3, pp. 211–221, 1950.

[77] D. D. Davis, "Advance production and Cournot outcomes: an experimental investigation," *Journal of Economic Behavior & Organization*, vol. 40, no. 1, pp. 59–79, 1999.

[78] P. van den Berg, L. Molleman, and F. J. Weissing, "Focus on the success of others leads to selfish behavior," *Proceedings of the National Academy of Sciences*, vol. 112, no. 9, pp. 2912–2917, 2015.

[79] G. Ellison and D. Fudenberg, "Word-of-mouth communication and social learning," *The Quarterly Journal of Economics*, vol. 110, no. 1, pp. 93–125, 1995.

[80] J. Webster and L. K. Trevino, "Rational and social theories as complementary explanations of communication media choices: Two policy-capturing studies," *Academy of Management journal*, vol. 38, no. 6, pp. 1544–1572, 1995.

[81] L. G. Tornatzky and K. J. Klein, "Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings," *IEEE Transactions on Engineering Management*, no. 1, pp. 28–45, 1982.

[82] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management science*, vol. 46, no. 2, pp. 186–204, 2000.

[83] L. Molleman, P. Van den Berg, and F. J. Weissing, "Consistent individual differences in human social learning strategies," *Nature Communications*, vol. 5, p. 3570, 2014.

[84] A. Bandura and R. H. Walters, *Social learning theory.* Prentice-hall Englewood Cliffs, NJ, 1977, vol. 1.

[85] A. Masini and E. Menichetti, "Investment decisions in the renewable energy sector: An analysis of non-financial drivers," *Technological Forecasting and Social Change*, vol. 80, no. 3, pp. 510–524, 2013.

[86] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[87] N. E. Friedkin and E. C. Johnsen, "Social influence and opinions," *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193–206, 1990.

[88] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences*, vol. 38, no. 10, pp. 886–893, 1952.

[89] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *4or*, vol. 5, no. 3, pp. 173–210, 2007.

[90] C. Shapiro and H. R. Varian, *Information Rules: a Strategic Guide to the Network Economy.* Harvard Business Press, 1998.

[91] F. M. Bass, "A new product growth for model consumer durables," *Management Science*, vol. 15, no. 5, pp. 215–227, 1969.

[92] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[93] A. Ordanini, G. Rubera, and R. DeFillippi, "The many moods of inter-organizational imitation: A critical review," *International Journal of Management Reviews*, vol. 10, no. 4, pp. 375–398, 2008.

[94] R. Reed and R. J. DeFillippi, "Causal ambiguity, barriers to imitation, and sustainable competitive advantage," *Academy of Management Review*, vol. 15, no. 1, pp. 88–102, 1990.

[95] R. Cappelli, D. Czarnitzki, and K. Kraft, "Sources of spillovers for imitation and innovation," *Research Policy*, vol. 43, no. 1, pp. 115–120, 2014.

[96] F. C. Santos, M. D. Santos, and J. M. Pacheco, "Social diversity promotes the emergence of cooperation in public goods games," *Nature*, vol. 454, no. 7201, pp. 213–216, 2008.

[97] P. Ramazi, "Convergence analysis and control of evolutionary matrix-game dynamics," dissertation, University of Groningen, 2017.

[98] J. Riehl, P. Ramazi, and M. Cao, "A survey on the analysis and control of evolutionary matrix games," *Annual Reviews in Control*, vol. 45, pp. 87–106, 2018.

[99] J. Gomez-Gardenes, M. Romance, R. Criado, D. Vilone, and A. Sánchez, "Evolutionary games defined at the network mesoscale: The public goods game," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 21, no. 1, p. 016113, 2011.

[100] J. R. Riehl and M. Cao, "Towards optimal control of evolutionary games on networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 458–462, 2017.

[101] K. H. Schlag, "Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits," *Journal of Economic Theory*, vol. 78, no. 1, pp. 130–156, 1998.

[102] C. Hauert, S. De Monte, J. Hofbauer, and K. Sigmund, "Volunteering as red queen mechanism for cooperation in public goods games," *Science*, vol. 296, no. 5570, pp. 1129–1132, 2002.

[103] C. S. Gokhale and A. Traulsen, "Evolutionary games in the multiverse," *Proceedings of the National Academy of Sciences*, vol. 107, no. 12, pp. 5500–5504, 2010.

[104] M. Bergen and M. A. Peteraf, "Competitor identification and competitor analysis: a broad-based managerial approach," *Managerial and Decision Economics*, vol. 23, no. 4-5, pp. 157–169, 2002.

[105] E. J. Zajac and M. H. Bazerman, "Blind spots in industry and competitor analysis: Implications of interfirm (mis) perceptions for strategic decisions," *Academy of Management Review*, vol. 16, no. 1, pp. 37–56, 1991.

[106] G. Facchini, F. Van Megen, P. Borm, and S. Tijs, "Congestion models and weighted bayesian potential games," *Theory and Decision*, vol. 42, no. 2, pp. 193–206, 1997.

[107] A. Fabrikant, C. Papadimitriou, and K. Talwar, “The complexity of pure nash equilibria,” in *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing.* ACM, 2004, pp. 604–612.

[108] M. Bagnoli and M. McKee, “Voluntary contribution games: Efficient private provision of public goods,” *Economic Inquiry*, vol. 29, no. 2, pp. 351–366, 1991.

[109] M. van Veelen and M. A. Nowak, “Multi-player games on the cycle,” *Journal of Theoretical Biology*, vol. 292, pp. 116–128, 2012.

[110] A. Szolnoki and M. Perc, “Impact of critical mass on the evolution of cooperation in spatial public goods games,” *Physical Review E*, vol. 81, p. 057101, May 2010.

[111] H. Hamburger, M. Guyer, and J. Fox, “Group size and cooperation,” *Journal of Conflict Resolution*, vol. 19, no. 3, pp. 503–531, 1975.

[112] M. A. Nowak and R. M. May, “Evolutionary games and spatial chaos,” *Nature*, vol. 359, pp. 826–829, 1992.

[113] K.-K. Kleineberg and D. Helbing, “Topological enslavement in evolutionary games on correlated multiplex networks,” *New Journal of Physics*, vol. 20, no. 5, p. 053030, 2018.

[114] A. J. Stewart, T. L. Parsons, and J. B. Plotkin, “Evolutionary consequences of behavioral diversity,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 45, pp. E7003–E7009, 2016.

[115] A. McAvoy and C. Hauert, “Autocratic strategies for alternating games,” *Theoretical population biology*, vol. 113, pp. 13–22, 2017.

[116] ——, “Autocratic strategies for iterated games with arbitrary action spaces,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 13, pp. 3573–3578, 2016.

[117] L. Pan, D. Hao, Z. Rong, and T. Zhou, “Zero-determinant strategies in iterated public goods game,” *Scientific Reports*, vol. 5, p. 13096, 2015.

[118] C. Hilbe, B. Wu, A. Traulsen, and M. A. Nowak, “Cooperation and control in multiplayer social dilemmas,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 46, pp. 16 425–16 430, 2014.

[119] B. Kerr, P. Godfrey-Smith, and M. W. Feldman, “What is altruism?” *Trends in Ecology & Evolution*, vol. 19, no. 3, pp. 135–140, 2004.

[120] A. Diekmann, "Volunteer's dilemma," *Journal of conflict resolution*, vol. 29, no. 4, pp. 605–610, 1985.

[121] C. Hilbe, A. Traulsen, and K. Sigmund, "Partners or rivals? strategies for the iterated prisoner's dilemma," *Games and Economic Behavior*, vol. 92, pp. 41–52, 2015.

[122] G. Ichinose and N. Masuda, "Zero-determinant strategies in finitely repeated games," *Journal of Theoretical Biology*, vol. 438, pp. 61–77, 2018.

[123] E. Akin, "The iterated prisoner?s dilemma: good strategies and their dynamics," *Ergodic Theory, Advances in Dynamical Systems*, pp. 77–107, 2016.

[124] M. Van Veelen, "Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong," *Journal of Theoretical Biology*, vol. 259, no. 3, pp. 589–600, 2009.

[125] U. Fischbacher, S. Gächter, and E. Fehr, "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, vol. 71, no. 3, pp. 397–404, 2001.

[126] E. Fehr and U. Fischbacher, "The nature of human altruism," *Nature*, vol. 425, pp. 785–791, 2003.

[127] D.-F. Zheng, H. Yin, C.-H. Chan, and P. Hui, "Cooperative behavior in a model of evolutionary snowdrift games with n-person interactions," *EPL (Europhysics Letters)*, vol. 80, no. 1, 2007.

[128] H. Liang, M. Cao, and X. Wang, "Analysis and shifting of stochastically stable equilibria for evolutionary snowdrift games," *Systems & Control Letters*, vol. 85, pp. 16–22, 2015.

[129] M. O. Souza, J. M. Pacheco, and F. C. Santos, "Evolution of cooperation under n-person snowdrift games," *Journal of Theoretical Biology*, vol. 260, no. 4, pp. 581–588, 2009.

[130] C. Hilbe, M. A. Nowak, and K. Sigmund, "Evolution of extortion in iterated prisoner's dilemma games," *Proceedings of the National Academy of Sciences*, p. 201214834, 2013.

[131] L. A. Imhof and M. A. Nowak, "Stochastic evolutionary dynamics of direct reciprocity," *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no. 1680, pp. 463–468, 2009.

[132] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg, "Emergence of cooperation and evolutionary stability in finite populations," *Nature*, vol. 428, pp. 646–650, 2004.

[133] A. J. Stewart and J. B. Plotkin, "From extortion to generosity, evolution in the iterated prisoner's dilemma," *Proceedings of the National Academy of Sciences*, pp. 15 348–15 353, 2013.

[134] A. Traulsen, M. A. Nowak, and J. M. Pacheco, "Stochastic dynamics of invasion and fixation," *Physical Review E*, vol. 74, no. 1, p. 011909, 2006.

[135] C. Hilbe, B. Wu, A. Traulsen, and M. A. Nowak, "Evolutionary performance of zero-determinant strategies in multiplayer games," *Journal of Theoretical Biology*, vol. 374, pp. 115–124, 2015.

[136] M. E. Schaffer, "Evolutionarily stable strategies for a finite population and a variable contest size," *Journal of Theoretical Biology*, vol. 132, no. 4, pp. 469–478, 1988.

[137] M. Shubik and R. Levitan, *Market Structure and Behavior.* Harvard University Press, 1980.

[138] P. D. Sozou, "On hyperbolic discounting and uncertain hazard rates," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 265, no. 1409, pp. 2015–2020, 1998.

[139] L. Green and J. Myerson, "A discounting framework for choice with delayed and probabilistic rewards." *Psychological Bulletin*, vol. 130, no. 5, pp. 769–792, 2004.

[140] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-offs.* Cambridge University Press, 1993.

[141] H. Rachlin, A. Raineri, and D. Cross, "Subjective probability and delay," *Journal of the Experimental Analysis of Behavior*, vol. 55, no. 2, pp. 233–244, 1991.

[142] P. Ostaszewski, L. Green, and J. Myerson, "Effects of inflation on the subjective value of delayed and probabilistic rewards," *Psychonomic Bulletin & Review*, vol. 5, no. 2, pp. 324–333, 1998.

[143] L. Green, A. F. Fry, and J. Myerson, "Discounting of delayed rewards: A life-span comparison," *Psychological Science*, vol. 5, no. 1, pp. 33–36, 1994.

[144] J. Andreoni, W. T. Harbaugh, and L. Vesterlund, "Altruism in experiments," *The New Palgrave Dictionary of Economics: Volume 1–8*, pp. 134–138, 2008.

[145] A. Dreber, D. Fudenberg, and D. G. Rand, "Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics," *Journal of Economic Behavior & Organization*, vol. 98, pp. 41–55, 2014.

[146] C. Hilbe, T. Röhl, and M. Milinski, "Extortion subdues human players but is finally punished in the prisoner?s dilemma," *Nature Communications*, vol. 5, p. 3976, 2014.

[147] A. W. Delton, M. M. Krasnow, L. Cosmides, and J. Tooby, "Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters," *Proceedings of the National Academy of Sciences*, vol. 108, no. 32, pp. 13 335–13 340, 2011.

[148] R. M. Dawes and R. H. Thaler, "Anomalies: cooperation," *Journal of Economic Perspectives*, vol. 2, no. 3, pp. 187–197, 1988.

[149] E. Fehr, U. Fischbacher, and S. Gächter, "Strong reciprocity, human cooperation, and the enforcement of social norms," *Human Nature*, vol. 13, no. 1, pp. 1–25, 2002.

[150] M. L. Weitzman, "Gamma discounting," *American Economic Review*, vol. 91, no. 1, pp. 260–271, 2001.

[151] D. J. MacKay and D. J. Mac Kay, *Information Theory, Inference and Learning Algorithms.* Cambridge University Press, 2003.

[152] A. Govaert and M. Cao, "Zero-determinant strategies in finitely repeated n-player games," *IFAC-PapersOnLine*, vol. 52, no. 3, pp. 150–155, 2019.

[153] J. K. Murnighan and A. E. Roth, "Expecting continued play in prisoner's dilemma games: A test of several models," *Journal of conflict resolution*, vol. 27, no. 2, pp. 279–300, 1983.

[154] R. S. Bielefeld, "Reexamination of the perfectness concept for equilibrium points in extensive games," in *Models of Strategic Rationality.* Springer, 1988, pp. 1–31.

[155] M. A. Nowak, S. Bonhoeffer, and R. M. May, "More spatial games," *International Journal of Bifurcation and Chaos*, vol. 4, no. 01, pp. 33–56, 1994.

[156] L. Boosey, P. Brookins, and D. Ryvkin, "Contests with group size uncertainty: Experimental evidence," *Games and Economic Behavior*, vol. 105, pp. 212–229, 2017.

[157] W. Mill and M. M. Theelen, “Social value orientation and group size uncertainty in public good dilemmas,” *Journal of Behavioral and Experimental Economics*, vol. 81, pp. 19–38, 2019.

[158] J. Bergin and W. B. MacLeod, “Continuous time repeated games,” *International Economic Review*, pp. 21–37, 1993.

Summary

Human decisions have a central role in emerging engineering applications such as smart energy grids and intelligent transportation systems. The decisions that impact the performance of these complex systems are often made in situations in which the immediate individual benefits conflict with the long-term performance of the overall system. For instance, the system performance can rely on individuals to share their energy resources, accept delays in energy consumption or product deliveries, etc. Without strategic or structural influence on individual decisions, in these social dilemmas selfish economic trade-offs can easily lead to an undesirable collective behavior. It is therefore crucial to identify mechanisms that can promote cooperative decisions that lead to better collective performance and sustainable outcomes. In this thesis, decision-making processes in social dilemmas are studied using the framework of mathematical games or game theory.

Part I of the thesis is concerned with network games, network reciprocity, and potential game theory. Based on economic and behavioral studies, novel decision-making dynamics are defined and studied that combine rationality principles with social learning through imitation (Chapter 4 and 5). It is shown how selfish decisions can be moderated by social influence to promote sustainable outcomes. Moreover, the mechanism called strategic differentiation is proposed through which players can react differently to their various neighbors in the network (Chapter 5). In this setting, at equilibrium cooperative decisions are promoted if players with a relatively high degree in the network (e.g. individuals with a large social network) differentiate their actions. However, strategic differentiation can become detrimental when it is applied by players that have a relatively low degree in the network (e.g. individuals with a small social network).

Part II is concerned with strategic solutions to social dilemmas in which players repeatedly interact with each other in a multiplayer game. A theory is developed that characterizes the level of control that a player can unilaterally exert in the eventual outcome of a multiplayer game with a finite number of expected rounds (Chapter

6). The theory covers a broad class of social dilemmas, that have been extensively studied in a variety of research disciplines including sociology, psychology, biology, and economics, and can capture a variety of complex situations in which the player's benefits (non-linearly) depend on the decisions of others. Through unilateral strategic influence cooperative behavior of selfish, rational co-players can be promoted and sustained. However, in contrast to classic theories, it is not necessary to assume rational behavior of the co-players. By making virtually no assumptions on the decision-making behavior, the theory can still ensure a relative payoff performance between the strategic player and the co-players. By characterizing this level of unilateral control we provide a robust framework through which the performance of systems that rely on repeated human decisions can be studied and improved. Expressions are given for the *efficiency* of strategic influence in terms of the minimum number of expected interactions required to enforce a desired behavior or relative performance (Chapter 7). This is useful, for instance, in designing additional benefits when strategic influence in collective outcomes must be achieved within a given time-frame. The evolutionary performance of these manipulative strategies is studied in Chapter 8. It is shown that under classic Maynard-Smith conditions (i.e. infinite population and pairwise interactions) only generous strategies, that typically enforce a linear payoff relation in which others do better, can be evolutionarily stable against an arbitrary mutant strategy. In sharp contrast, when "playing the field" (i.e. the entire population interacts with each other in a multiplayer game), only extortionate strategies, that typically enforce a linear payoff relation in which the strategic player outperforms others, are favored by evolution. In a finite population with a variable interaction size, we show how the evolutionary stability of a strategy depends on the population size, the number of players in each interaction, and the payoff parameters of the social dilemma. Finally, a general framework is proposed through which the interaction between strategic decision-making and uncertainty about the valuation of the future is studied (Chapter 9). With this novel framework, classic and modern theories of strategic play can be recovered in deterministic limits. More importantly, it enables to unveil, for the first time, how one might strategically influence the collective behavior of a large group of decision-makers that are uncertain about events in the future. This framework exhibits the characteristics of empirically validated time-inconsistent discounting observed in social, temporal, and probabilistic discounting frameworks, and indicates how strategic decisions and the possibilities for strategic influence must be adjusted to the level of uncertainty in the future.

Samenvatting

De keuzes van mensen spelen een centrale rol in opkomende engineering applicaties zoals slimme energie netwerken en intelligente transportsystemen. De keuzes die de prestaties van dergelijke complexe systemen beïnvloeden worden vaak genomen onder omstandigheden waarin onmiddellijke individuele belangen conflicteren met de langere termijn prestaties van het algehele systeem, dat bijvoorbeeld kan afhangen van het delen van energiebronnen of het uitstellen van energieverbruik of productleveringen. Zonder enige vorm van strategische of structurele invloed op individuele keuzes kunnen in dit soort situaties zelfzuchtige economische overwegingen leiden tot een ongewenst collectief gedrag. Het is hierdoor cruciaal om mechanismen te identificeren die coöperatieve keuzes kunnen promoten die leiden tot betere collectieve prestaties en duurzame uitkomsten. In deze thesis worden de beslissingsprocessen in dergelijke sociale dilemma's bestudeerd aan de hand van wiskundige spellen ofwel speltheorie.

Deel I van de thesis is begaan met netwerk spellen, netwerk wederkerigheid en potentieel speltheorie. Gebaseerd op economische en gedragskundige studies wordt een nieuw type dynamica gedefinieerd en bestudeerd dat rationaliteitsprincipes combineert met sociaal leren door middel van imitatie (Hoofdstuk 4 en 5). Het wordt aangetoond hoe zelfzuchtige keuzes gemodereerd kunnen worden door sociale invloed om duurzame uitkomsten te promoten. Daarnaast wordt een mechanisme, genaamd strategische differentiatie, voorgesteld waarmee spelers anders kunnen reageren op de keuzes van verbonden spelers in het netwerk (Hoofdstuk 5). In deze setting, worden in de evenwichtstoestand coöperatieve keuzes gepromoot als spelers met een relatief hoge graad in het netwerk (bijvoorbeeld individuen met een groot sociaal netwerk) hun keuzes differentiëren. Strategische differentiatie kan echter een negatieve invloed hebben op coöperatie als het wordt toegepast door spelers met een relatief lage graad in het netwerk (ofwel individuelen met een klein sociaal netwerk).

Deel II gaat over strategische oplossingen voor sociale dilemmas waarin spelers herhaaldelijk participeren in een meerpersoons spel. Een theorie wordt ontwikkeld die de mate van controle karakteriseert die een speler eenzijdig kan uitoefenen op de uiteindelijke uitkomst van het beslissingsproces met een eindig maar onbepaald aantal

interacties (Hoofdstuk 6). Deze theorie omvat een groot aantal sociale dilemma's die uitvoerig bestudeerd zijn in onderzoek disciplines zoals sociologie, psychologie, biologie, en economie, en kan een verscheidenheid aan complexe situaties beschrijven waarin het resultaat van een speler (niet-lineair) afhangt van de keuzes van anderen. Door de eenzijdige strategische invloed kunnen coöperatieve keuzes van zelfzuchtige en rationele medespelers worden gepromoot. In tegenstelling tot klassieke theorieën, is het echter niet noodzakelijk om aan te nemen dat de medespelers zich rationeel gedragen. Hoewel er praktisch geen aannames worden gemaakt over het beslissingsgedrag van anderen, kan de theorie alsnog een bepaalde relatieve prestatie van de strategische speler ten opzichte van de medespelers garanderen. Door de mate van dergelijke eenzijdige controle te karakteriseren wordt een robuust raamwerk ontwikkeld waarmee de prestaties van een systeem dat afhankelijk is van herhaaldelijke menselijke beslissingen bestudeerd en verbeterd kunnen worden. Analytische expressies worden gegeven voor de efficiëntie van de strategische invloed in termen van het minimum aantal interacties dat een strategische spelers nodig heeft om een bepaald gewenst gedrag of relatieve prestatie af te dwingen (Hoofdstuk 7). Dit kan bijvoorbeeld gebruikt worden voor het bepalen van extra voordelen wanneer strategische invloed in collectieve uitkomsten behaald moet worden binnen een gegeven tijdsvlak. De evolutionaire prestaties van deze manipulative strategieën wordt bestudeerd in Hoofdstuk 8. Er wordt aangetoond dat onder klassieke Maynard-Smith condities (i.e. een oneindige populatie en paarsgewijze interactie) alleen genereuze strategieën, die typisch ten voordele van de medespelers zijn, evolutionair stabiel zijn ten opzichte van een willekeurige mutant strategie. Echter, wanneer alle spelers in de populatie deelnemen in het spel (playing the field condities) kunnen alleen afpersing strategieën, die typisch ten voordele van de strategische speler zijn, bevoordeeld worden door evolutie. In een populatie met een eindig aantal spelers en variabele interactie groottes, tonen we aan hoe de evolutionaire stabiliteit van een strategie afhankelijk is van de populatiegrootte, de interactiegrootte en de payoff parameters van het sociale dilemma. Als laatste wordt een algemeen raamwerk ontwikkeld waarmee de interacties tussen strategische keuzes en onzekerheid over de waarde van de toekomst bestudeerd wordt (Hoofdstuk 9). Met dit nieuwe raamwerk kunnen klassieke en moderne theorieën over strategische beslissingen worden afgeleid in deterministische limieten. Van groter belang is dat dit raamwerk het mogelijk maakt om, voor de eerste keer, aan te tonen hoe strategische invloed kan worden uitgeoefend op het collectieve gedrag van een grote groep beslissingsnemers die niet geheel zeker zijn over gebeurtenissen in de toekomst. Het raamwerk bevat de karakteristieken van empirisch gevalideerde tijdsinconsistente verdiscontering dat geobserveerd is bij sociale, temporele, en probabilistische verdiscontering, en toont aan hoe strategische keuzes en de mogelijkheden voor strategische invloed aangepast moeten worden aan de onzekerheid in de toekomst.