INSTITUT POLYTECHNIQUE DE PARIS

HEC PARIS

NNT : 2021EHEC0009

Thèse de doctorat

Information, Coordination and Cooperation: Essays on Learning in Dynamic Games.

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à École des Hautes Etudes Commerciales de Paris

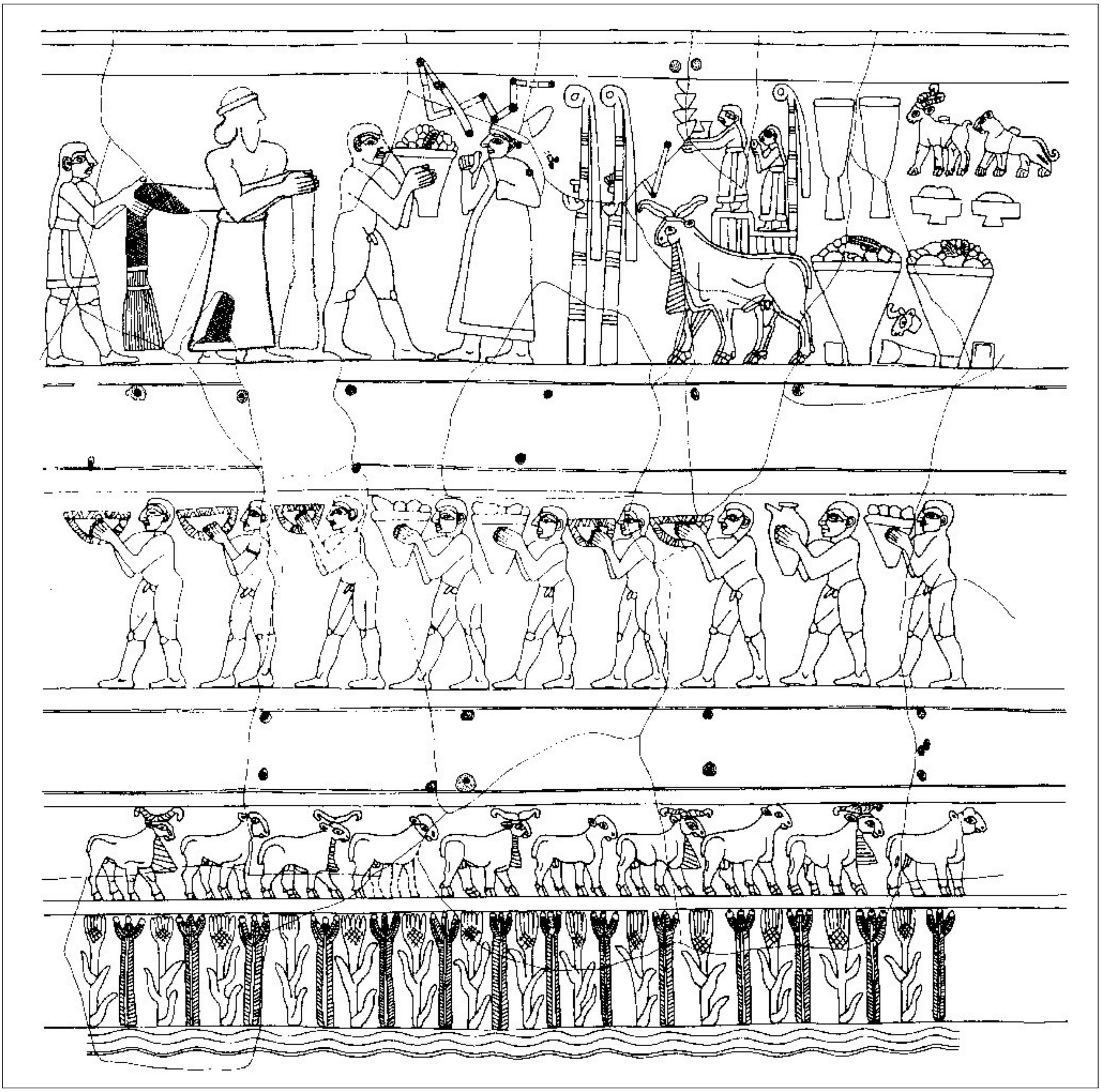
École doctorale n° 626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : Sciences de Gestion

Thèse présentée et soutenue à Jouy-en-Josas, le 07/12/2021, par

ÉMILIEN MACAULT

Composition du Jury :

Frédéric Koessler Directeur de recherche CNRS, École d'économie de Paris	Président
Penelope Hernandez Professeure, Universitat de València (LINEEX)	Rapporteure
Lucie Ménager Professeure, Université Paris 2 Panthéon-Assas (LEMMA)	Rapporteure
Marco Scarsini Professeur, LUISS Guido Carli (Economia e Finanza)	Examinateur
Nicolas Vieille Professeur, HEC Paris (E&DS)	Examinateur
Tristan Tomala Professeur, HEC Paris (E&DS)	Directeur de thèse



The Uruk vase, c. 3200-3000 BC, National Museum of Irak.

A René et Paulette,

A Babeth,

REMERCIEMENTS

Alors que je m'attelle à l'écriture de ces remerciements, je pense à la succession vertigineuse de circonstances et de coups de chance ou de malchance qui font que je me retrouve ce matin, devant mon bureau, un fichier LaTeX ouvert et un café à la main. S'il devait y avoir un point initial, forcément arbitraire, à cette trajectoire, il se trouverait sûrement près de quinze ans en arrière, en classe de Seconde. Je faisais à l'époque un constat sans appel: les seules choses dignes d'intérêt dans la vie sont celles suffisamment fortes pour rester immuables pendant des millénaires et il n'y a donc qu'une seule voie à envisager: l'archéologie. Je décidais donc de toute mettre en oeuvre pour me consacrer à des études littéraires et, surtout, pour éviter absolument les voies sans saveur des mathématiques et de l'ingénierie. Raté.

Je tiens avant tout à remercier mon directeur de thèse, Tristan Tomala, qui a été le moteur essentiel de ce travail, pour la qualité de son encadrement et la confiance totale qu'il m'a accordée. Je me souviens de notre première rencontre, un après-midi d'hiver dans les locaux sombres du boulevard Jourdan où j'étais venu te présenter mon projet de recherche. Déjà là, j'avais été impressionné par ta capacité non seulement à comprendre mes idée décousues, mais aussi à les reformuler en des termes intelligibles. S'en sont suivies 5 années de thèse à HEC Paris au cours desquelles je t'ai bombardé d'idées bancales et de questions floues, auxquelles tu as toujours su répondre avec précision et bienveillance, et je t'en suis infiniment reconnaissant. Tu as su m'aiguiller dans mes travaux, parfois dans des domaines très éloignés du tien avec un savoir et une patience impressionnants. Je me rends compte de tout ce que j'ai appris sous ta direction, qu'il s'agisse de connaissances techniques, pédagogiques ou autres, et je t'en sais gré. Merci de m'avoir laissé une liberté totale durant ces années tout en fournissant un soutien constant.

A l'été 2016, en fin de première année de thèse, j'ai rencontré Marco Scarsini à l'occasion d'une *summer school* en Bretagne. Deux mois plus tard, je le retrouvais à Rome pour travailler ensemble à la LUISS. Véritable *shadow co-advisor*, Marco a suivi la quasi-totalité de mes travaux de thèse avec intérêt et enthousiasme. Je tiens à le remercier pour m'avoir invité par deux fois à passer un semestre à Rome. J'étais loin de me douter de l'influence que ces séjours auraient tant sur mon travail que sur le reste de ma vie. Marco m'a appris la rigueur, aussi bien dans le raisonnement que dans l'écriture scientifique, et les joies d'un préambule LaTeX bien construit. Je tiens à le remercier pour tout ce qu'il m'a transmis, tant scientifiquement qu'humainement. Je connais peu de personnes qui prennent le soin de fouiller YouTube pour partager, à 2 heures du matin un soir de semaine, leur version favorite du Canto General de Theodorakis! Les homards de Port Jefferson et les poissonniers de la Canée sont des souvenirs impérissables.

Je tiens aussi à remercier les membres de mon jury, Lucie Ménager et Penelope Hernandez qui ont accepté de rapporter cette thèse, Frédéric Koessler qui présidera la soutenance, et Nicolas Vieille. J'ai rejoint Lucie au LEMMA en septembre en tant qu'ATER, et je la remercie du soin qu'elle a pris d'alléger mon emploi du temps pendant ce premier semestre pour me permettre de venir au bout de cette thèse. C'est avec beaucoup d'enthousiasme que j'aborde ce "début de l'après" au sein de son équipe. J'ai eu l'occasion d'assister à plusieurs conférences données par Penelope et je suis heureux, et terrifié, qu'une chercheuse de son niveau accepte d'évaluer mes travaux. Il en est de même pour Frédéric, dont les questions qu'il m'a posées à l'occasion de séminaires ont largement contribué à clarifier et faire avancer mes modèles. Que ce soit à HEC ou ailleurs, Nicolas a toujours été d'une infinie gentillesse et d'une grande disponibilité lorsque j'avais des questions à lui poser. Ses travaux, remarques et conseils ont été d'une grande richesse et je l'en remercie. Les savoir tous membres de mon jury est un grand honneur.

Ces années à HEC n'auraient pas été aussi fructueuses sans le concours d'un grand nombre de personnes. En premier lieu les membres du département d'économie et de sciences de la décision qui ont su me consacrer du temps. Je pense particulièrement à Eric pour ses nombreux conseils sur le *Job Market* et la pédagogie et pour ses retours, impitoyables certes mais essentiels, sur mes présentations. Je pense également à Raphaël et Stefania avec qui j'ai eu la chance d'apprendre à enseigner, Gae-

tano, Dinah, Victoria et tous les autres. Merci aussi à Gilles pour sa gentillesse et la qualité de son enseignement à l'occasion de son cours à Orsay qui m'a beaucoup marqué. Je tiens également à remercier l'équipe du programme doctoral, Kristine, Britta, Françoise et Mélanie qui déploient des efforts titanesques pour nous. Merci à Kristine de nous avoir soutenus dans nos combats de doctorants, de nous avoir consacré son temps et son énergie dans une mesure que je peine à imaginer, et d'avoir été une *Sinterklaas* mémorable lorsque je refusais d'endosser ce rôle. Merci à Britta pour toutes nos conversations. Je tiens à te rappeler d'ailleurs que j'ai un jeu de pétanque et une revanche à prendre. Merci à Françoise pour avoir toujours eu la bonne information à mes questions désespérées et tardives et d'avoir su jongler avec mes choix de cours aléatoires. Et merci à Mélanie d'avoir su faire preuve d'une infinie patience lorsque je rendais des documents qui n'étaient pas les bons et/ou incomplets et/ou en piteux état et/ou en retard et/ou impossibles à entrer dans Notilus. Merci aussi aux participants du séminaire parisien de théorie des jeux à l'Institut Henri Poincaré pour leur accueil bienveillant, particulièrement à Bruno, Xavier, Gaëtan et Nikos.

Mais au-delà du travail, une thèse représente aussi 5 années de vie, qui auraient pu être bien mornes sans le soutien de mes camarades doctorants. Dans le désordre, merci à Martin et Tam (et à Raphaëlle, évidemment) pour la belle amitié franco-canadienne, à Veronica pour sa sagesse infinie, à Shuv, à Claudia, à Yang, à Elena pour les plus belles heures de la FNDC, à Xucheng pour les pauses café, à Wei avec qui j'ai exploré le monde étrange de la théorie économique, à Atulya qui vient prendre la relève, à Thomas qui est venu avant, à Fan et tous les autres. Des soirées de Noël à Chaville à une nouvelle vie aux quatre coins du monde, *long live the FNDC*! C'est un honneur absolu d'avoir atteint le 77è palier du méta-Jean-Louis en si bonne compagnie. Merci aussi à ceux rencontrés à Rome, sur les tatamis comme sur les parquets de la LUISS: Tristan, Ludovic, Federica, Daniele, Giovanni (x2), Roberto *et al.* et qui ont contribué à faire de ce séjour une période si épanouissante.

Alors qu'une thèse fait entrevoir un avenir d'enseignant, je mesure l'influence que beaucoup ont eu sur moi. Merci à mes enseignantes du lycée Saint Louis de Bar-le-Duc, mesdames Rylko, Barros et Barthélémy, qui m'ont appris, dans l'ordre, que faire des mathématiques n'empêchait pas d'aimer tout le

reste, à avoir toujours pour soi le plus haut niveau d'exigence et, surtout, que le dilettantisme érigé en art est le meilleur moyen d'y parvenir. Merci à Manon Didry et Valérie Leblanc, professeures de mathématiques pendant mes années de B/L au lycée Fustel de Coulanges, qui m'ont appris à faire des mathématiques un objet familier, visuel, et à aimer leur art. Merci enfin à Marc Bachraty, *sensei* implacable, esthète du geste minimal, pour son enseignement et nos conversations sur les liens entre théorie des jeux et arts martiaux qui ont fortement influencé cette thèse.

Evidemment, la trajectoire menant à cette thèse ne commence pas en septembre 2016 et nombreux sont ceux qui ont contribué à l'infléchir, avant comme pendant. Je pense ainsi à mes confrères naufragés de la Meuse et ceux qui ont rejoint notre radeau de la Méduse par la suite. Merci à Paul, Eléonore, Valentin, Christelle, Jean-Aimé, Claire, Antoine, Chat, Djo', Aurélien et Solène pour ces 15 années de rigolade et celles à suivre. Je pense également à mes camarades strasbourgeois, avec qui j'ai découvert coup sur coup les plaisirs de l'économie et les fleischnackas. Merci donc à Gaëtan, Julia, Manon, Elise, Aloïs, Caroline, Colin, Manuel, Camille, Thomas, Maxence et JG. Et enfin, pour boucler la boucle, merci à mes amis du Malakoff Institute of Technology, cis 3 avenue Pierre Larousse: Victor, Nedde, Nathan, Jeanne, Raphaël, Jad, Hugo et Alexander, avec qui j'ai bravé vents, marées, la théorie de la mesure, SAS, l'AGRAF et tant d'autres choses. Et puis il y a tous ceux que je ne mentionne pas, par oubli, par souci de place, par pudeur ou simplement par peur de devoir organiser un pot de thèse qui ne tiendrait pas dans les murs d'HEC. Ils sauront se reconnaître.

Et puis, aussi, il y a la famille. Mes grands-parents, René et Paulette, à la fois phares dans la tempête et cause première: c'est eux qui m'ont appris à lire sans même s'en rendre compte, qui m'ont transmis leur curiosité insatiable et la faculté de s'émerveiller sans cesse, qui m'ont porté pendant toutes ces années de vie et d'études et qui continueront de le faire bien après. Mes parents, aussi, pour leur soutien constant et leur confiance. Jamais ils n'ont douté de mes choix même lorsqu'ils essayaient à grand peine de comprendre ce que je faisais et ce qui m'intéressait, et ils ont tout mis en œuvre pour que je puisse suivre la voie que je me suis tracée. Merci à Tantine, Delphine, Mostafa, Marion, Nawel, Marwane et Noam, à mon frère

Aurélien et mes sœurs Charlotte et Gabrielle ainsi qu'à mon grand-père Jean d'avoir été là, d'avoir partagé et compris mes galères de thésard. Merci aussi à Gilles, Mireille, Matthias et Babeth d'avoir toujours pensé à nous même lorsqu'ils traversaient, pour de vrai, la pire des tempêtes.

Enfin, et surtout, merci à Elise, qui aura tout bravé à mes côtés: les hauts, les bas, les très-bas. Merci pour tout ce que tu as fait, accepté et enduré pour arriver au bout de cette thèse. Tu en as vu chacune des étapes, tu as su t'y intéresser et comprendre ce dont je parlais, tu as su donner un sens moral à mon travail et tu as su me donner le sourire lorsqu'il n'y avait plus vraiment de raisons de l'avoir. Sans toi, je ne sais pas si j'aurais été capable d'y mettre un point final. Il t'appartient. Merci de permettre que, chaque matin, un Emilien meilleur que celui de la veille émerge. Merci de faire de chaque jour une aventure où l'ennui n'existe pas. Nous rirons vieux.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

WE Wardrop equilibrium

BWE Bayes-Wardrop equilibrium

NRG nonatomic routing game

RGUNS nonatomic routing game with unknown network state

DRGUNS dynamic nonatomic routing game with unknown network state

S.P. series-parallel

P.O.A. price of anarchy

P.O.S. price of stability

SLI series of linearly independent

I.I.D. independent and identically distributed

N.E. Nash equilibrium

S.P.E. subgame perfect equilibrium

S.A.A. stochastic approximation algorithm

SA stochastic approximation

R.V. random variable

O.D.E. ordinary differential equation

M.P.E. Markov Perfect equilibrium

INTRODUCTION

1

GENERAL INTRODUCTION

1.1 SCOPE OF THE THESIS

1.1.1 Introduction to game theory

Game theory emerged as an independent field in the early 1940s, with the publication of Von Neumann and Morgenstern [152], as a meeting point between mathematics and social sciences. Since then, it has expanded its scope to a variety of other fields, from political to computer science and biology. Nearly 80 years after the publication of the first milestone in game theoretic history, new problems have become central in expanding game theory as a field *per se* and the range of situations it can handle. Such a successful spread is the consequence of a simple, transposable approach to things: any situation where some entities interact while pursuing their individual interests can be abstracted as a *game* on which theorists can apply a panel of methods and results to derive predictive statements. Such predictions are based on the concept of *equilibrium*, a state of balance between individuals optimizing their outcomes where no agent can improve her situation by changing her decision. Equilibria are proved to exist under standard conditions and, when they do, they offer a reference point to policy makers as to what are the most likely outcomes to expect from rational and strategic agents. This dual nature of being both a field of study on its own and a methodological suite with interdisciplinary views is essential in understanding how game theory evolved and incorporated problems and methods from neighboring disciplines.

This thesis subsumes the work done during my PhD studies on three different topics that caught my interest as being of instantaneous relevance to explore current problems we observe as well as having some methodological appeal. The three resulting studies are mostly independent, be it in the nature of their questions or in the methods deployed to address them. They feature as three independent parts that are presented in the chronological order in which they were developed.

1.1.2 MISINFORMATION AND LEARNING IN GAMES

The first part of this thesis originated from the large-scale disinformation phenomena that came to light in the early 2010s. The recurring call to fake information created a breach in the trust that the general public was putting in news agencies, public institutions or even academia which generated a general state of doubt and mistrust with harsh economic and institutional consequences. This naturally raised the question of how does one proceed in setting up a disinformation strategy and, further, how is it possible to prevent such phenomenon to occur.

There exists a vast literature on information in games. Informational asymmetries and externalities occupy a central place in economic reasoning and game theory have provided insights on the topic since the pioneering papers by R. Aumann and M. Maschler on games with incomplete information, reedited in Aumann, Maschler, and Stearns [13]. Information acquisition and exploitation in games is generally well understood within the boundaries of the *Bayesian* framework. Such models analyze interactions between agents that rationally process information in the sense that they hold probabilistic beliefs on the uncertain elements at play and update them using Bayes rule. This strand of literature has proved that informational refinements – e.g. allowing one agent to be informed and not the other – strongly change the set of equilibria in games. Yet, Bayesian modeling offers limited insights when it comes to analyzing large populations: it assumes strong computational requirements on agents and offers solutions whose complexity explodes when the number of agents grows, unless some specific game structure is assumed as in Smith and Sørensen [144]. For these reasons, Bayesian models are fit to explain how a limited number of agents behave when it comes to forming beliefs and exploiting some informational advantage, but offer a limited insight to mass phenomena like disinformation and the spread of fake news.

An alternative approach, referred to as *bounded rationality* or *non-Bayesian*, refers to the nebula of models which rejected – partly or entirely – the Bayesian approach. This line of work stands on the idea that agents have bounded computational capacities that limit their treatment of information. This general statement entails a whole diversity of approaches but, when it comes to modeling large-scale opinion formation and belief exchange, one of them stands as the canonical option: the DeGroot model. This model considers a network of agent that iteratively average their beliefs with those of their neighbors. This process is shown to converge under mild conditions to a stable consensus that is easily tractable from the structure of the communication network and the initial beliefs. Many variations of the model have been considered but most of them rely on this simple local averaging dynamics. While its tractability plays as a comparative advantage with respect to Bayesian models, the DeGroot model is not entirely satisfactory either as it does not leave room to situations where some marginal opinion spreads to an important share of a population. To our view, this is mostly due to the belief averaging process, for reasons we detail in Chapter 3.

What we propose in Part i is an alternative approach to DeGroot, where agents communicate states drawn according to their beliefs instead of directly exchanging their beliefs. This model extends DeGroot and, at the same time, questions the nature of its predictions. As such, it contributes to the literature on the emergence and characterization of consensus.

1.1.3 ROUTING NETWORKS AND CONGESTION EXTERNALITIES

The second part of this thesis is also connected to the literature on learning in games but considers the other side of the coin: while Part i focuses on the exchange of exogenous opinions, Part ii considers a model where information is endogenously obtained as the outcome of equilibria in a dynamic game. This work originated from a study of the literature at the interface between game theory and computer science. Both fields have displayed a growing interest for learning-related problems while their approaches are radically different. Computer science is a natural home for the study of algorithmic learning procedures like Multiplicative Weights or the replicator dynamics. This non-behavioral approach to learning has been proved to operate efficiently in several classes of games, among which congestion games stand as one of the most studied. These games model situations where congestion externalities arise, that is, resource allocation problems where the more agents use one resource, the higher its cost. Several properties explain the particular interest for this class of games: equilibria in pure strategies exist under almost no condition, unicity is ensured if costs are strictly monotonic, and a strong connection exists between these games and convex optimization. Yet, despite these helpful properties and although learning algorithms perform generally well in expectation, they may be subject to chaotic behavior. This phenomenon was showed to appear in simple, non-degenerate instances of routing games in Chotibut et al. [41].

Routing games are a subclass of congestion games that model flows of selfish agents that move on a routing network and minimize their total travel time from an origin to a destination. What Chotibut et al. [41] show is that in a network with only two paths connecting an origin point to a destination, if travel times of each path are linear functions of the share of agents using those paths, then there exist a range of parameters for which the Multiplicative Weights algorithm displays chaotic behavior in the sense of Li and Yorke [108]. While this behavior may occur in the algorithmic realm, this is not the case for Bayesian beliefs that converge almost-surely as they are martingales. In Part i we proposed a non-Bayesian approach to a problem where standard Bayesian models failed to provide efficient answers. Symmetrically, in Part ii, we propose a Bayesian approach to a problem where standard algorithmic methods show concerning limits. We study a dynamic routing game model with incomplete information and explore the conditions under which a public Bayesian belief converges to the truth.

Our main question may be formulated as follows: consider a non-strategic navigation system that aggregates and broadcasts the routing choices and subsequent travel times of a new population each day. If the actual functions that map shares of users to travel times are unknown, will they be efficiently learned by the navigation system? Is a finite number of observations enough to accurately determine these functions? And how does learning connect to the structure of the routing network? We offer an answer to each of these questions using standard game theoretical modeling.

1.1.4 GAME THEORETICAL MODELING FOR ENVIRONMENT ECONOMICS

The third part of the thesis departs from considerations on learning and networks to explore alternative approaches to the modeling of time preferences and their consequences in terms of environmental decision making. The starting point to this study is the failure of states in meeting the terms of the Kyoto Protocol and their progressive disengagement as they drew closer to the expiration dates. The Protocol was negotiated in December 1997 as an extension of the United Nations Framework Convention on Climate Change, and became effective as of February 2005. It expired in December 2012 at the end of a first commitment period, then was extended for an additional 8-year period. In December 2020, the second commitment period reached its end with a contrasted efficiency. The Kyoto Protocol is a complex treaty that received harsh criticism, part of which led to the drafting of the Paris agreement, and many are the reasons invoked for this insufficiency.

Game theoretical modeling has provided an array of explanations to this problem. A first insight is to be found in the incomplete nature of the agreement itself, which generated hold-up problems. A second reason for the limited efficiency of the treaty is its lack of enforcement power: only countries listed as Annex B had binding targets and a punishment scheme in case they do not meet their objectives. Further, punishments were limited and did not deter deviations. Finally, it has been argued that when studying such long-term agreements, standard models of time preferences, i. e. exponential discounting with constant rate, do not fit time preferences as observed in experiments. Rather, countries may display some time inconsistency: today's policymaker may disagree with the choices of a future policy maker, as they do not evaluate outcomes in the same way. Many models already exist on this topic and non-constant discounting is fairly widespread in environmental economics. But these models rarely incorporate the concept of a deadline.

While most treaties fix a date at which some objective has to be met, and the scientific communication on global warming usually works on the forecast state of the world at a specific horizon, little is known as to how fixing these dates influences decision making over time.

In this part, we propose to approach this problem by building a model where two states compete over a bounded time interval. They jointly control one common state of the world *via* their actions. They play at times determined by a Poisson process. That way, we endogenize a decreasing discount factor: the closer they are to the deadline, the less likely they will play again. The objective of the study is to determine optimal trajectories of actions and assess the influence of the rate of the Poisson process.

1.1.5 GENERAL CONSIDERATIONS

While each part of the thesis addresses a specific independent problem, several underlying topics emerge throughout as common threads. All the models we consider rest on uncertain environments: there is an unknown state of the world in Part i and Part ii, and the sequence of times at which agents play in Part iii is unknown. As a consequence, in all three parts, our analysis relies on stochastic methods. Part i and Part ii are thematically related: both consider learning models, although under different heuristics, and are interested in the formation and behavior of a public belief. Network externalities appear in Part i as the diffusion of beliefs in the population is shaped by the existence of a communication network, and in Part ii as the combinatorial structure of the action space is characterized by a routing network. Finally, both Part ii and Part iii are influenced by considerations on cooperative behavior: both models seek to establish results on cooperative behavior in environments where Folk theorems fail to apply.

1.2 SUMMARY OF THE MAIN CONTRIBUTIONS

We summarize below the contributions made in the different parts of the thesis. Those summaries also appear in their respective parts.

1.2.1 Part I

In Part i, we propose a stochastic extension to non-Bayesian models of opinion exchange. We build an opinion formation model where agents communicate by drawing states according to their beliefs instead of directly communicating subjective probabilities. To do so, we model beliefs as a system of interacting urns. Balls of different colors represent the possible values of the state of the world. At each iteration of the communication process, the number of balls in urns grows as they add new balls based on their neighbors' draws. We assimilate the evolution of beliefs to the evolution of urns' compositions.

In this setup, we prove that under very general conditions, the dynamics of beliefs converge to a rest point. Our proof relies on stochastic approximation techniques. We then show that at the steady-state, all agents in a connected component of the communication network share the exact same belief on the state of the world. This result follows from algebraic properties of graphs. Finally, we show that as long as initial beliefs cover the whole state space, the consensus is drawn from a distribution with full support. This strongly contradicts the predictions of DeGroot's and similar models. We then try to characterize this limit distribution using simulations. We establish a set of conjectures regarding the limit consensus. First, we believe that the limit belief follows a beta distribution. This conjecture is supported by the close connection between our model and the Polya urn model. Second, we believe that the expected value of the limit belief on the wrong state of the world is equal to the initial probability that an urn is misinformed. In other terms, while not being a martingale, the vector of proportions in urns behaves as if it were one. This result is proved for regular graphs where we prove that the sum of proportions across urns is a martingale, and conjectured in the general case.

This part contributes to the literature on opinion exchange and emergence of consensus by extending one of the canonical models to a stochastic framework, and showing that this drastically changes its predictions. An existing literature has already criticized the robustness of the DeGroot model for reasons detailed in Chapter 3. We add new elements to the controversy, while providing a model where extreme events that are observed in reality but were absent of DeGroot can occur with positive probability. The original purpose of the paper was to provide tractable metrics on disinformation to help build models of strategic disinformation. While the exact nature of the limit consensus remains a conjecture, we have good hope to be able to prove these results and use them in a broader model.

1.2.2 Part II

In Part ii, we consider a repeated symmetric nonatomic routing game where the cost functions of each edge of the routing network depend on the load of the edge and on an unknown state parameter that is invariant over time. The set of states is finite and endowed with a common prior. At each period of time, a short-lived generation of users with a given total demand plays the game and realizes a Wardrop equilibrium with respect to the expected costs on edges: each path that receives positive load has the least expected cost. For every used edge, its load and the corresponding realized cost become public information for the following generations. There is perfect recall, so each generation knows the entire past history of the game and updates its beliefs in a Bayesian way. The sequence of different generations' demands is assumed to be random, independent and identically distributed (i.i.d.).

We consider two concepts of social learning: under strong learning, players eventually learn the true state of the world; under weak learning they learn to play the game as if the true state of the world was known. We show that weak learning is a strictly weaker concept than strong learning and that the conditions to achieve either of them depend on the topology of the network and on the support of the random demand. Our main theorem proves that weak learning occurs if the routing network is series-parallel and both the cost functions and the support of the demand are unbounded. Further, we show that strong learning is achieved under the same prerequisites and the additional condition that the demand has full support over $\mathbb{R}_+$. The intuition behind this result is the following: when the demand is stochastic, equilibrium flows vary. This generates observations of the cost functions for different values of loads. Based on results from Cominetti, Dose, and Scarsini [44] on the variation of equilibrium flows with respect to the demand, we prove that in a series-parallel network, as the demand goes high, all edges are used in equilibrium and equilibrium loads are unbounded. This implies that the cost functions will be observed at levels which allow distinguishing between the cost-relevant states with probability one. Finally, we prove that the condition on the network topology is necessary: for typical networks that do not satisfy it, we show that there exists an assignment of cost functions and capacities such that weak learning fails for any distribution of the demand.

This part contributes to the literature on social learning by offering a case of a *large game* where sequences of continuous player sets achieve learning. We also contribute to the literature on routing games by reconnecting it to a more traditional approach on learning in games. As such, we offer a learning model that is immune to chaos where beliefs converge in finite time.

1.2.3 Part iii

In Part iii, our objective is two-fold. From a theoretical perspective, we intend to build a model of revision games with flow payoffs and a cumulative state to structure the trade-off faced by states involved in environmental transition. Two players act over a finite time interval and are offered at stochastic dates to revise their decision. This yields a flow of payoffs and determines a common state of the world that captures the share of time each player played the non-cooperative action. This state is a proxy to model the impact of countries emissions over the course of the game. When time reaches the end of the interval, a terminal payoff is determined according to the state of the world.

This model remains an early-state project and most of our efforts have been spent in identifying the adequate model. Nevertheless, we provide a detailed agenda of the results we are working on. In practical terms, we first intend to prove the existence of an optimal symmetric strategy profile in the form of a threshold strategy: prior to a time threshold that negatively depends on the expected frequency of revision times, both players play the Nash equilibrium of the one-shot game, then shift to a cooperative action profile. In that respect, the higher the frequency at which players revise their strategies, the shortest the duration of the cooperative regime. This relationship between decision frequency and cooperative behavior is already observed in revision games without flow payments.

This project contributes to the recent literature on revision games by offering an alternative version of the model where players balance a flow of payoffs and a terminal payoff. It also contributes on the literature on environmental decision making by connecting theoretically the frequency at which policies are revised with the enforcing of cooperative behavior. Finally, while being equivalent to stochastic games, the class of revision games is immune to existing Folk theorems. Proving the existence of subgame perfect equilibria sustaining some cooperation is then an addition to what is known on cooperation in non-cooperative dynamic games.

1.3 GUIDELINES FOR THE READER

This section details the organization of the remainder of the thesis. In particular, we refer the reader to chapters where our main contributions are stated.

The next chapter, Chapter 2 is a continuation of this general introduction. It offers a complete overview on mathematical tools used throughout the thesis. We first review general results of game theory as a mean to fix our terminology and notation. Then we recall some important definitions and facts related to graph theory. Finally, we also provide some material on the study of dynamical systems and stochastic approximation. The reading of these contents is absolutely dispensable for the reader, and any of these sections may be skipped independently of the others.

Part i is devoted to our first paper on the emergence of consensus and reinforcement learning. It comprises two chapters. Chapter 3 is mostly introductory, as it offers a detailed review of the literature on the emergence of consensus and motivates our main problem consequently. Chapter 4 contains our contribution on the topic. We first detail the model we study, and then provide our main results and their proofs. A section is devoted to results obtained from simulations, and additional outputs feature in Appendix a.

Part ii is based on a paper written in collaboration with Tristan Tomala and Marco Scarsini on social learning in routing games. Chapter 5 is devoted to an introduction of the problem and provides information on our modeling choices and the related literature. Chapter 6 contains our main model, results and examples.

Part iii contains our ongoing project on revision games. It consists in a single chapter, Chapter 7, that motivates the main problem and details our modeling process. Our main model is fully stated and we included a detailed perspective on our ongoing research.

Section 7.4 contains an extensive summary of the thesis in French.

Appendix a provides further information on the simulations used in Part i. We present the main code elements used to run the simulations as well as additional outputs that support the content of Chapter 4.

Thus, chapters in this thesis may be split as follows: Chapter 2 is a technical preliminary, Chapter 3 and Chapter 5 are mostly introductory material that cast light on what is contained in other chapters. The reader seeking our own contributions is referred to Chapter 4 and Chapter 6 that only contain our models and results. Finally, Chapter 7 is a hybrid, containing contextual elements as well as our model.

2

THEORETICAL PRELIMINARIES

In order to ease the reading and make this thesis as much of a self-contained document as possible, we devote this chapter to a technical exposition of the major mathematical concepts, tools and results that will be used throughout. Most of the definitions and results in this chapter are well-known, hence an informed reader may prefer to skip some sections. Featured results will not be proved but references containing the proofs will be mentioned. Section 2.1 covers the basic elements of game theory: normal-form games, repeated games and games with incomplete information. Section 2.2 presents the main elements of graph theory that will be of use in the thesis. Section 2.3 provides some background on dynamic systems and stochastic approximation methods which we use. Table 2.1 details which part of the thesis require each section of the present chapter.

Sections of Chapter 2	Part i	Part ii	Part iii
Section 2.1	✓	✓	✓
Section 2.2	✓	✓	✗
Section 2.3	✓	✗	✗

Table 2.1: Correspondence between Sections and Parts of the Thesis.

2.1 NON-COOPERATIVE GAME THEORY

Game theory aims at modeling and analyzing strategically interactive environments in which *agents* make decisions with mutual outcomes. In this thesis we will not cover *cooperative* – or *coalitional* – games, i. e. situations where agents seek to form coalitions. Our focus will be entirely on *non-cooperative* games, with agents acting for their own good. All the results featured in this section are textbook material. This section is greatly inspired by Laraki, Renault, and Sorin [101] and Laraki, Renault, and Tomala [102] as well as the lecture notes by Bruno Ziliotto.

2.1.1 GAMES, PLAYERS AND EQUILIBRIA

A strategic game is defined by the following elements:

- A set of *players* N – assumed finite for the moment – of cardinal n which identifies the agents (e. g.individuals, countries, populations) operating;

- A set of *pure strategies* $\mathcal{S}_i$ for each $i \in N$ – also referred to as *actions* – that contains all the elements player i chooses from;

- A *payoff function* g mapping $\mathcal{S} = \prod_{i \in N} \mathcal{S}_i$ to a vector in $\mathbb{R}^n$.

A game in *normal form* is described as a tuple $G = (N, (\mathcal{S}_i)_{i \in N}, g)$. All sets are assumed to be non-empty. If the player set and every action set are finite, the game G is called *finite*. An *action profile* is a vector $s = (s_1, \ldots, s_n)$ of strategy choices for each player. We write $s = (s_i, s_{-i})$ where s_{-i} denotes the choice of strategies s_j for any player $j \neq i$ and $\mathcal{S}_{-i} = \prod_{j \neq i} \mathcal{S}_j$. For any player $i \in N$, the i-th entry $g_i(s)$ of the payoff vector corresponds to i's individual gain (or loss) when the strategy profile s is played. Throughout, we do not distinguish the *utility* of an agent from her payoff. Normal form games are simultaneous move games: every player chooses an action independently from the other players.

The payoff map g creates a natural hierarchy over strategies: for any player $i \in N$, for a fixed profile s_{-i}, the higher the payoff $g_i(s_i, s_{-i})$ a strategy s_i grants, the "better" it is. This leads to the concept of domination.

Definition 1. A strategy $s_i \in \mathcal{S}_i$ is *dominated* if there exists another strategy $s_i' \in \mathcal{S}_i$ such that

$$\forall s_{-i} \in \mathcal{S}_{-i}, g_i(s_i', s_{-i}) \geq g_i(s_i, s_{-i}) \tag{2.1}$$

If the inequalities are strict, we say s_i is *strictly dominated* by s_i'.

Intuitively, a rational player is expected not to play any strictly dominated strategy.

Definition 2. A strategy $s_i \in \mathcal{S}_i$ is *dominant* if for every other strategy $s'_i \in \mathcal{S}_i$ such that $s'_i \neq s_i$,

$$\forall s_{-i} \in \mathcal{S}_{-i}, g_i(s_i, s_{-i}) \geq g_i(s_i, s_{-i}) \tag{2.2}$$

If the inequalities are strict, we say s_i is *strictly dominant*.

Similarly, a strictly dominant strategy is the only rational choice when it exists.

Players are assumed to be rational in the sense that they maximize their utility (respectively, minimize their disutility) using every available information. Utility maximization leads naturally to the definition of the *best-response correspondence*.

Definition 3. For any $i \in N$, given a profile s_{-i} of opponents' actions, a strategy $s \in \mathcal{S}_i$ is a *best-response* to s_{-i} if $g_i(s, s_{-i}) \geq g_i(s', s_{-i})$ for any $s' \in \mathcal{S}_i$.

We define the best-response correspondence of player i, BR_i, as the set-valued function mapping profiles in $\mathcal{S}_{-i}$ to best-responses of player i. Building on this definition, we can introduce our main equilibrium concept.

Definition 4. An action profile $s = (s_1, \ldots, s_n)$ is a *Nash equilibrium* (N.e.) of the game G if, for every $i \in N$ and $s'_i \in \mathcal{S}_i$,

$$g_i(s_i, s_{-i}) \geq g_i(s'_i, s_{-i}) \tag{2.3}$$

Equivalently, an action profile $s = (s_1, \ldots, s_n)$ is a Nash equilibrium if for every player $i \in N$, $s_i \in BR_i(s_{-i})$.

At times, we will have to use the following weaker definition of the best-response:

Definition 5. Let $\varepsilon > 0$. For any $i \in N$, given a profile s_{-i} of opponents' actions, a strategy $s \in \mathcal{S}_i$ is a *ε-best-response* to s_{-i} if $g_i(s, s_{-i}) \geq g_i(s', s_{-i}) - \varepsilon$ for any $s' \in \mathcal{S}_i$.

For $\varepsilon = 0$, we fall back to Definition 3.

Definition 6. A *mixed strategy σ_i* of player $i \in N$ is a probability distribution over $\mathcal{S}_i$.

Mixed strategies of player $i \in N$ belong to the simplex $\Delta(\mathcal{S}_i)$ which we write Σ_i. The set of mixed strategy profiles is $\Sigma = \prod_{i \in N} \Sigma_i$. For any mixed strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$, for any player $i \in N$, we define the extended payoff $g_i(\sigma)$ as the expected payoff:

$$g_i(\sigma) = \sum_{s \in \mathcal{S}} \left[\prod_{j=1}^{n} \sigma_j(s_j) \right] g_i(s) \tag{2.4}$$

Definition 7. The *mixed extension* Γ of the normal form game $G = (N, \mathcal{S}, g)$ is the normal form game (N, Σ, g).

Having defined the mixed extension of a game, we can state the main existence theorem from Nash et al. [119].

Theorem 8. *Any finite game G admits a N.e. in mixed strategies.*

2.1.2 REPEATED GAMES, FOLK THEOREM AND OBSERVABILITY

In the section above, we only consider a game played once but many strategically interactive situations involve some degree of repetition. Repeated games aim at modeling those that fit a discrete time setup. A repeated game proceeds as follows: let G be a normal form game, at discrete times $t \geq 1$,

- Every player $i \in N$ chooses independently and simultaneously a (possibly mixed) action s_i^t;

- Every player $i \in N$ receives her payoff $g_i(s_1^t, \ldots, s_n^t)$;
- The action profile $s^t = (s_1^t, \ldots, s_n^t)$ is publicly observed.

The game G is referred to as the *stage game*. As action profiles are observed publicly and perfectly, at any stage t, players remember past plays. This leads to the natural definition of histories of the game.

Definition 9. Fix $t \geq 1$. A *history* of length t of the repeated game is an element $(s^1, \ldots, s^t) \in \mathcal{S}^t$.

For any $t \geq 1$, let H_t be the set of histories of length t, with $H_0 = \{\emptyset\}$. Let $H = \cup_{t \geq 0} H_t$ be the set of histories and $H_\infty = \mathcal{S}^\infty$ the set of infinite plays. We can define pure strategies in the repeated game as follows.

Definition 10. A pure strategy of player i in the repeated game selects an element in $\mathcal{S}_i$ for any period t and history $h \in H_{t-1}$.

Repeating a game greatly expands the set of pure strategies. Mixed strategies of the repeated game are defined as probability distributions over the set of pure strategies. This concept may be difficult to handle when dealing with explicit construction of strategies, hence the following alternative concept is often preferred.

Definition 11. A *behavior strategy* of player $i \in N$ is a map from H to Σ_i.

In other terms, a behavior strategy associates a mixed strategy of the stage game to any history of the game. Observe that any measurable profile of strategies σ induces a unique probability $\mathbb{P}_\sigma$ on the set of histories of finite length, as players' choices are independent. This probability uniquely extends to H_∞ by Kolmogorow extension theorem. The next result, due to Kuhn [99] and Aumann [11] connects mixed and behavior strategies.

Theorem 12. *We have the following equivalences:*

1. *Every behavior strategy of player i is equivalent to a mixed strategy of player i in the following sense: for every behavior strategy β_i of player i there exists a mixed strategy σ_i such that for any measurable profile σ_{-i}, $\mathbb{P}_{\beta_i,\sigma_{-i}} = \mathbb{P}_{\sigma_i,\sigma_{-i}}$;*

2. *Every mixed strategy of player i is equivalent to a behavior strategy of player i.*

Having discussed strategies in the repeated game, we now turn to payoffs.

Definition 13. Let $T \geq 1$. The T-repeated game is the game $G_T = (N, \Sigma, \gamma^T)$ where, for every $i \in N, \sigma \in \Sigma$,

$$\gamma_i^T = \mathbb{E}_\sigma \left[\frac{1}{T} \sum_{t=1}^{T} g_i(s^t) \right] \tag{2.5}$$

The T-repeated game is the finitely repeated game where payoffs are evaluated at their average value.

Definition 14. Let $\delta \in [0, 1)$. The δ-discounted game is the game $G_T = (N, \Sigma, \gamma^\delta)$ where, for every $i \in N, \sigma \in \Sigma$,

$$\gamma_i^\delta = \mathbb{E}_\sigma \left[(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(s^t) \right] \tag{2.6}$$

The
the value of 1 at stage $t+1$ is $1-\delta$ at stage t. Nash equilibria of both the T-repeated game and the δ-discounted game may display suboptimal choices, in particular in the case of punishment strategies. This requires a refinement of the equilibrium concept suited for repeated games. For any history $h \in H$ and strategy σ_i of player i, let $\sigma_i[h]$ be the strategy induced by σ_i after history h.

Definition 15. Let $T \geq 1$. A *subgame perfect equilibrium* (S.P.e.) of G_T is a N.e. σ such that for any time $t \leq T$ and any history $h \in H_t$, $\sigma[h]$ is a N.e. of G_{T-t+1}.

This equilibrium definition is stronger than Definition 4 as it requires that the equilibrium strategy profile is a N.e. at every history, in particular after some players potentially deviated. A similar definition exists for the δ-discounted game.

Definition 16. Let $\delta \in [0,1)$. A S.P.e. of G_δ is a N.e. σ such that for every $t \geq 1$ and every history $h \in H_t$, $\sigma[h]$ is a N.e. of G_δ.

Finally, we introduce the uniform approach to infinitely repeated games.

Definition 17. A strategy profile σ is a uniform equilibrium of G_∞ if:

1. $\forall \varepsilon > 0$, σ is a ε-N.e. of any sufficiently long finitely repeated game, i. e. :

$$\exists T_0, \forall T \geq T_0, \forall i \in N, \forall \tau \in \Sigma_i, \gamma_i^T(\tau_i, \sigma_{-i}) \leq \gamma_i^T(\sigma) + \varepsilon, \text{ and} \tag{2.7}$$

2. The sequence $(\gamma^T(\sigma))_T$ converges to a vector $\gamma(\sigma)$ of $\mathbb{R}^n$, called the uniform equilibrium payoff of G_∞.

We note E_∞, E_T and E_δ the sets of equilibrium of, respectively, G_∞, G_T and G_δ. The three sets are compact and G_T and G_δ are non-empty. Further, one has $E_1 \subset E_T \subset E_\infty$ and $E_1 \subset E_\delta \subset E_\infty$. We now turn to the characterization of equilibrium payoffs. The first criteria for a payoff vector to be sustained in an equilibrium of the repeated game is that there exists a (mixed) Nash equilibrium generating said payoff vector.

Definition 18. A payoff profile $u \in \mathbb{R}^n$ is *feasible* if there exists $\pi \in \Delta(\mathcal{S})$ such that, for every $i \in N$,

$$u_i = \sum_{s \in \mathcal{S}} \pi(s) g_i(s) \tag{2.8}$$

The set of feasible payoffs is $Co(\mathcal{S}) = g(\delta(\mathcal{S}))$.

The set $Co(\mathcal{S})$ is a convex compact polytope containing E_∞. The second criteria for the sustainability of a payoff vector is that players deem it rational to receive such payoffs. To formalize that idea, we introduce the following definition.

Definition 19. For any $i \in N$, we define the *punishment level* of player i as

$$v_i = \min_{\sigma_{-i} \in \prod_{j \neq i} \Delta(\mathcal{S}_j)} \max_{\sigma_i \in \Delta(\mathcal{S}_i)} g_i(\sigma_i, \sigma_{-i}) \tag{2.9}$$

For any player i, v_i is the lowest payoff that non-coordinated opponents may impose on i.

Definition 20. A payoff profile $u \in \mathbb{R}^n$ is *individually rational* if, for every player i, $u_i \geq v_i$. We denote by IR the set of individually rational payoffs.

This sequence of definitions leads to the all-important *Folk theorem*. Of unknown lineage, the first written version of this result is attributed to Aumann and Shapley [14] and Rubinstein [136].

Theorem 21.

$$E_\infty = Co(\mathcal{S}) \cap IR \tag{2.10}$$

This result states that any feasible and individually rational payoff vector can be obtained as the outcome of a uniform equilibrium. We will conclude by citing the equivalent results for finitely repeated and discounted games. Sorin [146] prove the convergence of the set E_δ to $Co(\mathcal{S}) \cap IR$.

Theorem 22 (Sorin [146])**.** *If there are only two players or there exists $u \in Co(\mathcal{S}) \cap IR$ with $u_i > v_i$ for every player i, then*

$$E_\delta \to Co(\mathcal{S}) \cap IR \text{ as } \delta \to 1. \tag{2.11}$$

Fudenberg and Maskin [64] provides an equivalent result for S.P.e. of the discounted game. For finitely repeated games, the following result is due to Benoit and Krishna [23].

Theorem 23 (Benoit and Krishna [23]). *If for every player i there exists $u \in E_1$ with $u_i > v_i$, then*

$$E_T \to Co(\mathcal{S}) \cap IR \text{ as } T \to \infty. \tag{2.12}$$

An equivalent result for S.P.e. of finitely repeated games features in Benoit and Krishna [22].

2.1.3 GAMES WITH INCOMPLETE INFORMATION

We finish this overview of game theory by introducing the general framework of games with incomplete information. As uncertainty is at the heart of most economic situations, they form a canonical yet heterogeneous class of games within the literature. Just like repeating a game makes the set of equilibria explode, introducing uncertainty in games made the literature sprawl in uncountable and sometimes strongly different directions, let alone considering repeated games with incomplete information. For the curious reader, Aumann, Maschler, and Stearns [13] contains most of the major – though dated – references on the subject. In particular, it includes the seminal works for the U.S. Arms Control and Disarmament Agency where major results on games with lack of information on one side were originally stated.

The foundations of the framework analyzing games with incomplete information – or *Bayesian games* – were established in Harsanyi [78] and Mertens and Zamir [113]. They propose to model incomplete information *via types* that belong to a "universal type space" over which players form Bayesian beliefs. In its general statement, a *Bayesian game* is defined as a tuple $G = (N, (\mathcal{S}_i)_{i \in N}, (\Theta_i)_{i \in N}), (\mu_i)_{i \in N}, g$ where:

- N is a set of *players*;

- $\mathcal{S}_i$ is the set of *pure strategies* of player $i \in N$;
- Θ_i is the set of *types* of player $i \in N$;
- $\mu_i : \Theta_i \to \Delta(\Theta_{-i})$ maps each type of player i to a distribution over types of other players;
- A *payoff function* g mapping $\mathcal{S} \times \Theta$ to a vector in $\mathbb{R}^n$.

Initially, a type vector $(t_1, \ldots, t_n)$ is drawn. Each player $i \in N$ knows her realized type t_i, as well as the game G. Her belief is given by the distribution $\mu_i(t_i)$, which may be seen as the conditional probability distribution of the *common prior distribution* of types $\mu \in \Delta(\Theta)$ conditional on t_i being realized. This framework may also include uncertainty on the set of actions of other players by fixing arbitrarily low payoffs to an action for certain types. We now turn to strategies.

Definition 24. A *behavior strategy* of player $i \in N$ in the Bayesian game G is a mapping $\sigma_i : \Theta_i \to \Delta(\mathcal{S}_i)$. We denote by Σ_i the set of mixed strategies of player i and by $\Sigma = \prod_{i \in N} \Sigma_i$ the set of strategy profiles.

We now define the adequate equilibrium concept.

Definition 25. A strategy profile $\sigma \in \Sigma$ is a *Bayesian equilibrium* – or *Bayes-Nash equilibrium* – of the game G if for each player $i \in N$ and type $t_i \in \Theta_i$,

$$\mathbb{E}_\sigma(g_i | t_i) \geq \mathbb{E}_{\tau_i, \sigma_{-i}}(g_i | t_i) \quad \forall \tau_i \in \Sigma_i \tag{2.13}$$

In other terms, Bayesian equilibria of the game G are Nash equilibria of the normal-form game with expected payoffs. Consequently, Theorem 8 applies and ensure existence of an equilibrium when types and actions sets are finite. More generally, an equilibrium exists when the type spaces are countable, action sets are compact metric and payoffs are continuous.

We stop our description of game theoretical results here, as this framework is important for the understanding of Part i, Part ii and Part iii. Each of these parts carry the basic Bayesian model to very different directions and will provide the necessary elements for their own understanding independently.

2.2 GRAPH THEORY

In this section, we introduce the essential definitions and results from graph theory which we will use in later parts. Graphs – or networks – have become a staple of economic modeling, as a tool to structure local externalities and interdependencies. Such models capitalize on a vast corpus of literature in discrete mathematics and combinatorics. In this section we will present the elements which are useful for further chapters. More information is to be found in graph theory monographs like Wilson [156] or Diestel [54]. For economic applications of graph theory, see Jackson [85].

2.2.1 ELEMENTARY DEFINITIONS

The language of graph theory is not necessarily standardized. To avoid confusions, we recall the basic definitions and the notations used thereafter.

Definition 26. A *graph* $\mathcal{N}$ consists in:

- A non-empty countable set $\mathcal{V}$ of *nodes* (or *vertices*);
- A countable set $\mathcal{E}$ of pairs of elements of $\mathcal{V}$, the *edges*.

If the set of vertices $\mathcal{V}$ is finite, $\mathcal{N}$ is said to be *finite*. For any $a, b \in \mathcal{V}$, if $\{a, b\} \in \mathcal{E}$, then a is connected to b by an edge of the graph. We may also write $ab \in \mathcal{E}$ when there is no risk of confusion. This leads to the following distinction.

Definition 27. For any graph $\mathcal{N}$,

- If the edge set $\mathcal{E}$ is unordered – i. e. $\{a, b\} = \{b, a\}$ for any $a, b \in \mathcal{V}$, we say that $\mathcal{N}$ is *undirected*;

- If, instead, the edge set is ordered, in which case $\{a, b\} \neq \{b, a\}$, we say that $\mathcal{N}$ is *directed* – equivalently, $\mathcal{N}$ is called a *digraph*.

This distinction involves strong differences both in the analysis and in terms of modeling. In a directed graph, node a being connected to node b does not imply the converse. If, for instance, edges are defined as an observation structure, this implies that observation is asymmetric.

In Definition 26, we do not forbid the presence of self-loops – i. e. an edge connecting a node a to itself – nor do we preclude the existence of several edges connecting the same pair of nodes. To avoid any confusion, we say that a graph $\mathcal{N}$ is a *simple graph* when it contains no self-loops and only a single edge may connect an ordered pair of nodes. Else, we say that $\mathcal{N}$ is a *multigraph*. In Part i, we consider undirected simple graphs while in Part ii we consider directed multigraphs without self-loops.

Graphs are an essential tool to model indirect relations. To that end we define the following notions.

Definition 28. A *walk* in $\mathcal{N}$ from node a to node b is a sequence of edges $v_1v_2, v_2v_3, \dots, v_kv_{k+1} \in \mathcal{E}$ such that $a = v_1$ and $b = v_{k+1}$. If every node in the sequence $v_1, \dots, v_k$ is distinct, the walk from a to b is called a *path*. A walk from one node to itself is called a *cycle*.

A graph is *connected* if for every two nodes $a, b \in \mathcal{V}$, there exists a path from a to b.

It is frequent to study networks in terms of local interactions.

Definition 29. Two nodes $a, b \in \mathcal{V}$ are said to be *adjacent* if:

- The graph $\mathcal{N}$ is undirected and $ab \in \mathcal{E}$;
- The graph $\mathcal{N}$ is directed and either ab or $ba \in \mathcal{E}$.

Definition 30. The *neighborhood* $\mathcal{N}(a)$ of node a is the set $\{b \in \mathcal{V} | ab \in \mathcal{E}\}$. The neighborhood of a set S of nodes is the union of their neighborhoods.

The size of the neighborhood of a node is an important tool in assessing the importance a node plays in a graph. This involves the following definitions.

Definition 31. In an undirected graph $\mathcal{N}$, the *degree* d_a of a node a is the cardinality of the neighborhood of a, that is

$$d_a = \text{card}\,\{b : ba \in \mathcal{E}\} \tag{2.14}$$

When dealing with directed graphs, it is important to distinguish between edges pointing to and from a node, hence we make the following distinction.

Definition 32. Let $\mathcal{N}$ be a directed network. For any $a \in \mathcal{V}$,

- The *in-degree* of a is the cardinal $\text{card}\,\{b : ba \in \mathcal{E}\}$;
- The *out-degree* of a is the cardinal $\text{card}\,\{b : ab \in \mathcal{E}\}$.

2.2.2 SUBGRAPHS, COMPONENTS AND MINORS

It is often useful to restrict the analysis of a graph to a limited number of nodes and edges.

Definition 33. A graph $\mathcal{N}' = (\mathcal{V}', \mathcal{E}')$ if a *subgraph* of a graph $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ if $\mathcal{V}' \subseteq \mathcal{V}$ and $\mathcal{E}' \subseteq \mathcal{E}$.

Subgraphs can be obtained by deleting edges, nodes, etc. A frequent and useful way to decompose a graph into subgraphs is through connectedness.

Definition 34. A (connected-)*component* of a network $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ is a non-empty subgraph $\mathcal{N}' = (\mathcal{V}', \mathcal{E}')$ such that:

- $\mathcal{N}'$ is connected, and
- If $a \in \mathcal{V}'$ and $ab \in \mathcal{E}$, then $b \in \mathcal{V}'$ and $ab \in \mathcal{E}'$.

The components of a graph are the maximal connected subgraphs. It allows to decompose a graph into sets of path-connected nodes. Two nodes belonging to different components generally have no influence over each other. There is no convention regarding whether isolated nodes are components or not.

Another way to build a relation between graphs is through the concept of *minors*. For a moment, we will only consider undirected graphs. For such a graph $\mathcal{N}$, consider the following operations:

- (D): Deleting an edge between two nodes;
- (C): Contracting an edge ab by merging its two end-nodes into a new node which is adjacent to all nodes in the former neighborhoods of a and b;
- (R): Remove an isolated node.

Definition 35. Any undirected subgraph $\mathcal{N}'$ that can be obtained from $\mathcal{N}$ by a finite application of $(D), (C)$ and (R) is called a *minor* of $\mathcal{N}$ and we say that $\mathcal{N}'$ is *embedded* in $\mathcal{N}$. A minor $\mathcal{N}'$ of $\mathcal{N}$ such that no path in $\mathcal{N}'$ has an inner vertex on another path is called a *topological minor* of $\mathcal{N}$.

For directed networks, these definitions extend in the following sense: an undirected graph $\mathcal{N}'$ is a (topological) minor of a directed graph $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ if $\mathcal{N}'$ is a (topological) minor of the undirected graph $\bar{\mathcal{N}} = (\mathcal{V}, \bar{\mathcal{E}})$ where $ab \in \bar{\mathcal{E}}$ if either $ab \in \mathcal{E}$ or $ba \in \mathcal{E}$. Both minor and topological minor relations define a partial ordering on the set of finite graphs. Some classes of graphs can be defined by the exclusion of certain minors. For instance, the class of series-parallel introduced in Definition 83 in Chapter 5 can be characterized as the class of graphs that do not admit the graph depicted in Fig. 2.1 as a topological minor.

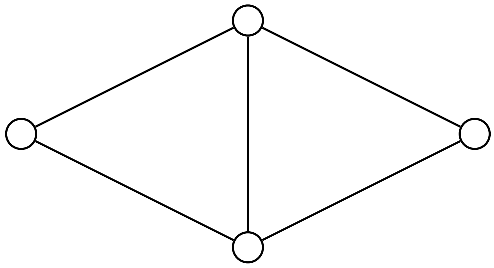


Figure 2.1: A forbidden minor.

2.2.3 FAMILIES OF NETWORKS

We now introduce several families of graphs which will be used in the thesis.

Definition 36. A finite simple network $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ is a *complete graph* if for any pair of nodes $a, b \in \mathcal{V}$, $ab \in \mathcal{E}$.

In other terms, a complete graph is a simple graph in which all the possible edges exist. Another important family is circle graphs.

Definition 37. A finite simple undirected (resp. directed) network $\mathcal{N}$ is a *circle* if it is connected and every node has a degree equal to 2 (resp. in- and ou- degree equal to 2).

Both circles and complete graphs belong to the broader class of k-regular graphs.

Definition 38. Let $\mathcal{N}$ be a finite undirected (resp. directed) simple graph of n nodes, and let $k \leq n$. $\mathcal{N}$ is a *k-regular graph* if every node has a degree equal to k (resp. in- and out-degree equal to k).

Thus, circles are 2-regular networks and complete networks are regular networks of maximal index of regularity k. When not required, we omit the k.

Definition 39. A (n, m)*-bipartite* graph $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ of $n + m$ nodes is a network whose vertex set $\mathcal{V}$ can be divided into two sets A of size n and B of size m such that if there are two nodes $a, b \in \mathcal{V}$ such that $ab \in \mathcal{E}$, then either $a \in A$ and $b \in B$ or $a \in B$ and $b \in A$. A bipartite graph is *complete* if every vertex in A is connected to every vertex in B and *vice versa*.

Finally, we define the family of star graphs.

Definition 40. A finite simple graph $\mathcal{N}$ of n nodes is a *star* if it is a complete $(1, n-1)$-bipartite graph.

Alternatively, star graphs consist of one single node called the *center* connected to every other node, called *leaves*, and such that no two leaves are connected.

2.2.4 NETWORK ALGEBRA

An alternative representation of graphs involve algebraic tools. Indeed, considering a finite network $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ of size n, we can represent $\mathcal{N}$ by a $n \times n$ squared matrix A called the *adjacency matrix* of $\mathcal{N}$.

Definition 41. For any finite network $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ of size n, the *adjacency matrix* A of $\mathcal{N}$ is the $n \times n$ squared matrix with entries

$$A_{ij} = \begin{cases} 1 \text{ if } ij \in \mathcal{E}; \\ 0 \text{ else.} \end{cases} \tag{2.15}$$

If $\mathcal{N}$ is undirected, then A is symmetric. To model situations where two connections do not have the same weight, it is possible to allow A to have values in $\mathbb{R}$ instead of $\{0, 1\}$ only. In that case, entries A_{ij} contain the weights that are put on the edges and the graph $\mathcal{N}$ is said to be *weighted*. We may also present the degrees of an undirected graph in matrix form.

Definition 42. The degree matrix of an undirected graph $\mathcal{N}$ of size n is the $n \times n$ diagonal matrix D with diagonal entries

$$D_{ii} = d_i \tag{2.16}$$

And $D_{ij} = 0$ if $i \neq j$.

The degree matrix of a graph is not involved as often as its adjacency matrix, but both serve in defining the *Laplacian matrix* concept.

Definition 43. Let $\mathcal{N}$ be a finite unweighted undirected graph. The *Laplacian matrix* of $\mathcal{N}$ is the $n \times n$ squared matrix L defined as

$$L = D - A. \tag{2.17}$$

The Laplacian matrix of a graph $\mathcal{N}$ carries essential information about the structure of $\mathcal{N}$ and has the following important properties.

Property 44. *For any undirected network $\mathcal{N}$ with non-negative weights,*

1. *L has only real eigenvalues;*
2. *L is positive semidefinite;*
3. *The smallest eigenvalue of L is $\lambda_1 = 0$ with corresponding eigenvector $(1, \ldots, 1)'$. The multiplicity of λ_1 is equal to the number K of connected components of $\mathcal{N}$.*

Those properties will be recalled when used in Section 4.2.2 where a proof is featured. For further properties and results on Laplacian matrices, we refer the reader to the survey Mohar [115]. There stops our brief overview of graph theoretical terminology. When required, further chapter will provide additional information on specific points.

2.3 STOCHASTIC APPROXIMATION ALGORITHMS

In this section, we present the main methods and terminology used in stochastic approximation theory. In essence, stochastic approximation aims at approaching limit sets of a discrete stochastic differential system by trajectories of a well-defined continuous-time differential system. The objective is not to determine analytical solutions but rather to characterize the set where, in the limit, trajectories will converge and remain stable. This requires a certain amount of definitions to precisely fix the nature of such limit points. There exists a vast literature on the subject, where Kushner and Yin [100] and Borkar [29] stand as high-quality textbooks. The survey by Pemantle [124] gives a good overview of the connection between dynamical systems theory and the study of stochastic processes. For a detailed review of dynamical systems theory, see Hirsch, Smale, and Devaney [83]. Most of this section is inspired by these references.

2.3.1 DYNAMICAL SYSTEMS

Dynamical systems theory is the field of mathematics that studies the long-term behavior of systems described by a time-dependent state. As an independent field, it possesses its own terminology, which requires proper defining.

We begin by defining the *state space* Θ as a subset of some Euclidian space, typically $\Theta \subseteq \mathbb{R}^n$ for some $n < \infty$. A *dynamical system* is a description of the evolution of points in Θ along time. When time is measured at integer values, the system is said to be *discrete*, and when it is measured continuously the system is said to be *continuous*. When the system evolves continuously with respect to time, it is said to be *smooth*. Formally,

Definition 45. A *smooth dynamical system* on $\mathbb{R}^n$ is a continuously differentiable function $\Phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ where $\Phi(t, x) = \Phi_t(x)$ satisfies:

- Φ_0 is the identity function, i. e. $\Phi_0(x) = x \forall x \in \mathbb{R}^n$;
- $\Phi_t \circ \Phi_s = \Phi_{t+s}$ for any $t, s \in \mathbb{R}$.

Given some point $x \in \mathbb{R}^n$, the map $t \to \Phi_t(x)$ is called a *trajectory* and describes the evolution over time of a system starting in state x. When the system is smooth, it can be formulated as the *ordinary differential equation* (o.d.e.) $\dot{x} \equiv \frac{dx}{dt} = f(x)$ with

$$f(x) = \frac{d}{dt}\bigg|_{t=0} \Phi_t(x). \tag{2.18}$$

The map Φ_t is the time flow associated to the differential system $\dot{x} = f(x)$. When f is C^1, by Cauchy-Lipschitz theorem, we immediately have existence and uniqueness of solutions and the continuity of Φ with respect to t and x. Although these solutions exist, in the general case, they seldom have a computable closed-form expression, especially when the o.d.e. system is non-linear. Instead of solving the differential equation, dynamical systems theory studies the asymptotic behavior of the trajectories defined by the o.d.e..

Definition 46. The *orbit* of a point $p \in \mathbb{R}^n$ is the set

$$\gamma(p) = \left\{ x \in \mathbb{R}^n \middle| p = \Phi_t(x) \text{ for some } t \in \mathbb{R} \right\}. \tag{2.19}$$

In other terms, the orbit of a point p is the trajectory of Φ passing at p. Note that two orbits containing a same point p are necessarily equal. Our objective is to characterize orbits that bear suitable stability properties in the limit. To that end, we first define invariant sets of points.

Definition 47. A set $\mathcal{P} \subseteq \mathbb{R}^n$ is called a *positively invariant* set if for all $x \in \mathcal{P}$ and $t \geq 0$, $\Phi_t(x) \in \mathcal{P}$.

Invariant sets are, trivially, sets of points whose trajectories remain within themselves. We will need an additional property of invariant sets called *chain transitivity*.

Definition 48. An invariant set $\mathcal{P}$ is *internally chain transitive* if for any $x, y \in \mathcal{P}$, any $\epsilon > 0$ and $T > 0$ there exist $n \geq 1$ and points $x_0 = x, x_1, \ldots, x_{n-1}, x_n = y$ in $\mathcal{P}$ such that the trajectory initiated at point x_i meets with the ϵ-neighborhood of x_{i+1} for $0 \leq i \leq n$ after a time greater or equal to T.

The next definition concerns the forward – or $\omega-$ – limits of the system, but equivalent statements exist for the backward – or $\alpha-$ – limits.

Definition 49. A point $p \in \mathbb{R}^n$ is an *ω-limit point* for an orbit $\gamma(x)$ if there exists a sequence $t_n \to \infty$ such that $\lim_{n\to\infty} \Phi_{t_n}(x) = p$.

The set $\omega(x)$ of all ω-limit points of trajectories going through a point x is called the *ω-limit set* of x.

They bear the following properties.

Property 50. *The ω-limit set of any orbit γ is:*

- *A closed set;*
- *An invariant set;*
- *Transitive, in the sense that if $z \in \omega(y)$ and $y \in \omega(x)$ then $z \in \omega(x)$.*

Moreover, if the orbit γ is bounded, then its ω-limit set is non-empty and connected.

Two asymptotic behavior within ω-limit sets are to be distinguished.

Definition 51. An equilibrium of the o.d.e. $\dot{x} = f(x)$ is a point $p \in \mathbb{R}^n$ such that $f(p) = 0$.

Equilibria of an o.d.e. are ω-limit points. If a trajectory reaches an equilibrium, it remains there forever. Alternatively, when they converge, trajectories may enter a periodic orbit.

Definition 52. An orbit $\gamma(x)$ of the dynamical system Φ is *periodic* if there exists $p \in \gamma(x)$ and $T > 0$ such that $\Phi_t(p) = \Phi_{t+T}(p)$

Although thanks to Poincaré-Bendixon theorem, in dimensions 1 and 2 ω-limit sets are easily characterized as being either equilibrium points or periodic orbits, this is not the case in dimension 3 or higher where much richer behavior may emerge. Thus, it is essential to identify within the limit sets of an o.d.e. those where trajectories will converge and remain stable.

Definition 53. A compact invariant set M is an *attractor* if there exists an open neighborhood O of M such that every trajectory in O remains in O and converges to M. The largest such O is called the *domain of attraction* of M.

Definition 54. A compact invariant set M is *Lyapunov stable* if, for any $\epsilon > 0$ there exists $\delta > 0$ such that every trajectory in the δ- neighborhood of M remains in its ϵ-neighborhood.

A compact invariant set M that is both Lyapunov stable and an attractor is *asymptotically stable*. When M is equal to a single equilibrium point x, then the equilibrium x^* is asymptotically stable. There exist several criteria to check asymptotic stability of equilibria, but the most widespread is *Lyapunov's second method*. Assume that there exists a continuously differentiable map V defined on a neighborhood O of x^* such that $\langle \nabla V(x), f(x) \rangle < 0$ for $x^* \neq x \in O$ and $\langle \nabla V(x^*), f(x^*) \rangle = 0$, with $V(x) \to \infty$ as x goes to the boundary of O. Then x^* is asymptotically stable. Conversely, if an equilibrium x^* is asymptotically stable, then such a function V exists. The same method may be applied to compact invariant sets instead of equilibrium points. Finding a *Lyapunov function* V may be difficult in practice but, when the o.d.e. is defined as a linear system with constant coefficients, i. e.

$$\dot{x} = Ax \tag{2.20}$$

where A is a squared matrix, then equilibria are asymptotically stable if and only if A is semi-definite negative.

An equilibrium point that is asymptotically stable and such that every trajectory of the o.d.e. converges to it is said to be *globally stable*. In that case the set O above may be replaced by the entire state space $\mathbb{R}^n$. More generally, if there exists a continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ such that $\langle \nabla V(x), f(x) \rangle \leq 0$ for every $x \in \mathbb{R}^n$ then any trajectory of the o.d.e. converges to the largest invariant set in $\left\{ x \in \mathbb{R}^n \middle| \langle \nabla V(x), f(x) \rangle = 0 \right\}$. When the o.d.e. is a linear system with constant coefficients, if the matrix A is semi-definite negative, then equilibrium points are globally stable and trajectories converge at exponential rate.

2.3.2 Monro-Robbins algorithm

So far, we only discussed continuous dynamical systems. In Part i though, we need to analyze the behavior of discrete stochastic systems. Approaching the latter by trajectories of continuous systems is a method known as *stochastic approximation* (SA). Consider a stochastic sequence (x_t) in $\mathbb{R}^n$. A *stochastic approximation algorithm* (s.a.a.) consists in finding a deterministic map f, a sequence of positive scalars γ_t and a stochastic sequence u_t such that

$$x_{t+1} = x_t + \gamma_t \left[f(x_t) + u_t \right], \quad t \geq 0. \tag{2.21}$$

Eq. (2.21) is known as the *Monro-Robbins algorithm*, introduced in Robbins and Monro [128]. While it strongly resembles a Newton discretization scheme, it is specifically designed to handle cases where the latter fails. Their original idea was to build a method to find roots of a regression function when observations are affected by random errors that preclude convergence of Newton's scheme. The recursive scheme is designed to average out the errors u_t in the limit. Then, convergence happens with probability 1 at the cost of the following assumptions.

Assumption 1. The map $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz, i. e.

$$\exists 0 < L < \infty, \| f(x) - f(y) \| \leq \| x - y \| \tag{2.22}$$

This assumption is standard in differential calculus and ensures that the o.d.e. $\dot{x} = f(x)$ is well-posed.

Assumption 2. The sequence of positive scalars (γ_t) satisfy

$$\sum_t \gamma_t = \infty, \text{ and} \tag{2.23}$$

$$\sum_t \gamma_t^2 < \infty \tag{2.24}$$

The first condition follows from the discretization intuition: $(\gamma_t)_t$ is a sequence of step-sizes on the time interval, which we want to cover fully. The second condition is essential in order to ensure that the noise u_t vanishes in the limit, jointly with the next assumption.

Assumption 3. The sequence $(u_t)_t$ is a sequence of square-integrable martingale difference noises with respect to the filtration $\mathcal{F} = (\mathcal{F}_t, t \geq 0) = (\sigma(x_s, u_s, s \leq t)t \geq 0)$, i. e.

$$\mathbb{E}\left[u_{t+1}\middle|\mathcal{F}_t\right] = 0 \tag{2.25}$$

Finally, we want the sequence of observations $(x_t)_t$ to satisfy the following assumption.

Assumption 4. The iterates of Eq. (2.21) are bounded almost-surely, i. e.

$$\sup_t \|x_t\| < \infty \tag{2.26}$$

This is generally a tricky assumption although when we will use SA in Part i, we will work on bounded random variables only. Under Assumptions 1–4 we have convergence of Eq. (2.21) with probability 1 in the following sense.

Theorem 55. *The sequence $(x_t)_t$ converges almost surely to a compact connected internally chain transitive invariant set of the o.d.e.*

$$\dot{x} = f(x) \tag{2.27}$$

Studying limit values of a s.a.a. is then equivalent to characterizing invariant sets of the associated o.d.e.. Popularity of this method is explained by the high versatility of its main result. Numerous variations of Eq. (2.21) are shown to behave in a similar fashion, when stepsizes are stochastic, or f is time-dependent, or the o.d.e. is projected on a particular set, or even when Assumption 4 only holds on some set with positive probability. Part i will require such extensions but, in flavor, the method remains very close to what is exposed in the present section.

Part I

MISINFORMATION, LEARNING AND CONSENSUS IN NETWORKS

This part is based on the paper *Stochastic Consensus and the Shadow of Doubt*. It proposes a stochastic approach to non-Bayesian learning models and on the emergence of consensus in networks. Namely, we propose a stochastic model of opinion exchange in networks, where agents communicate by signaling states drawn according to their beliefs. A finite set of agents is organized in a fixed network structure. There is a binary state of the world and, *ex ante*, each agent receives a random signal informing them about one of the two values of the state of the world. Based on reinforcement learning models and stochastic approximation techniques, we show that if some fraction of the agents are misinformed with positive probability at the beginning of the process, then in the limit, all beliefs converge to the same value, which is a random variable with full support. This result strongly departs from standard in the literature and shows that even with a marginal fraction of misinformed agents, false information may spread within the whole society with positive probability. We further provide characterization elements on the distribution of the random limit belief through simulations.

Chapter 3 introduces the problem through an overview of the literature on information exchange in networks. Chapter 4 details the model and presents the main results.

While this paper is single-authored it greatly benefited from the comments of members of the department of Economics and Finance at LUISS, the department of Economics and Decision Science at HEC Paris and participants in the Parisian seminar of Game Theory at Institut Henri Poincaré.

3

ON OPINION EXCHANGE IN NETWORKS

Doubt is our product since it is the best means of competing with the "body of fact" that exists in the mind of the general public. It is also the means of establishing a controversy.

Smoking and Health Proposal, 1969
(Brown and Williamson Tobacco
Corporation Records)

This chapter introduces the main problem studied in Chapter 4 and the connected literature. It explores major models of opinion exchange, their results and limitations. It is largely based on the surveys Acemoglu and Ozdaglar [5], Mossel and Tamuz [118] and Golub and Sadler [75], and on the textbook Jackson [86]. Section 3.1 introduces and motivates the main problem. Section 3.2 presents Condorcet's early model of opinion aggregation. Section 3.3 presents the Bayesian approach on opinion exchange in networks. Section 3.4 introduces bounded-rationality modeling and major results on the DeGroot model of belief exchange. Section 3.5 develops a critical view on naive learning and motivates the use of reinforcement learning in opinion formation models.

3.1 INTRODUCTION

Historical instances of misinformation date back as far as the existence of written testimonies of battles of opinion. A long-lasting example is to be found in the late years of the Roman Republic. After the death of Julius Caesar, his adopted son Octavian and his general Marcus Antonius both claimed the legitimacy of the power and engaged in a civil war. Throughout years of conflict, they also waged war to gain the opinion of the Roman public. In an attempt at discrediting Marcus Antonius, Octavian spread poems and slogans depicting his enemy as a frivolous defector engaged in a romance with queen Cleopatra Philopator against Roman moral principles. Octavian eventually won the war, leading to the suicide of Marcus Antonius and Cleopatra and the progressive establishment of the Roman Empire. Yet, over two millenia later, the myths spread by the future emperor Augustus still outshadow historical facts in the minds of the public, reinforced through artistic depictions of these events.

This story highlights three key points in the functioning of misinformation. First, we observe the call to emotions and irrational thinking over actual facts – later coined *post-truth politics* in Tesich [150] – in a situation where the public had no means to check whether the conveyed information was true or not. Second, we see that in the long-run, the consensual belief may remain away from the truth. Despite efforts in reestablishing the facts, the romanticized picture of Marcus Antonius and Cleopatra remains dominant in the views of today's public. Finally, we note that this piece of misinformation has spread through almost the entire population although the civil war officially lasted for two years, after which efforts in disinforming the population stopped.

It appears important at this point to establish a distinction between *misinformation* that is, the spread of a false information, and *disinformation*, which additionally requires a purpose sought in manipulating one's belief. Two thousand years after the death of Augustus, methods and means used to disinform have evolved drastically. Yet, the mechanics at the heart of misinformation have remained very similar. Bernays [25] even argues that the presence of beliefs manipulation is a characteristic element of democratic societies. What changed, though, is the speed and scale at which information transits. The generalization of social media as the main means of communication and information has conducted to a noticeable increase both in the occurrences of and the attention given to instances of misinformation.

Up to this date, it remains difficult to assess the exact scale of misinformational content on social platforms, as pointed out in Lazer et al. [104]. The authors acknowledge the limited and sometimes contradictory measures of the impact of fake news in the 2016 American elections. On this particular point, Allcott, Gentzkow, and Yu [7] provides a partial answer suggesting a strong increase in the production of fallacious content until 2017, after which a slight decrease is observed. Yet, despite public investments in media education and the development of counter-measures that followed 2016, misinformation remains an ongoing issue with tangible consequences. The recent examples of the COVID-19 pandemics and the following vaccination campaign have highlighted how quickly inaccurate, deceptive or politically biased information spreads in a context of distrust towards experts and institutions.

Extensive work on the mechanisms behind mis- and disinformation in cognitive and political sciences highlighted an array of explanatory factors for this recent outburst – e.g. cognitive biases, mistrust in institutions and media, political motives – which fall beyond the scope of this thesis. Under pressure to limit the spread of deceptive content, media and open web companies have put in place a set of policies to regulate news contents on their platforms. Such policies mostly include source highlighting, fact checking and advertisement campaigns, all of which have proved to have a debatable efficiency as highlighted in Levin [106]. In this part, we claim that the failure of counter-disinformation policies may be explained by theoretical modeling shortcomings.

3.2 CONDORCET AND THE FOUNDATION OF OPINION EXCHANGE MODELS

The question of information aggregation stands as one of the oldest and most driving problems in economic theory in particular, and social sciences in general. Attempts at understanding how one shared public opinion emerges from the interaction of individuals' private information through the use of mathematical tools can be traced back to Condorcet [46] at least. In this early model, as summed up in Chamley [36], Condorcet considers an unknown binary state of the world and a set of agents each receiving a private signal informing them of the true state with some fixed probability p. He then studies the asymptotics of majority voting with respect to the size of the population, which led to the well known jury theorem.

Formally, consider a set of $n \in \mathbb{N}$ identical players – the juries – having to decide whether a person if guilty or not. There is a state of the world $\theta \in \{G, I\}$ corresponding to the true nature of the defendant: guilty (G) or innocent (I). Prior to making a decision, each member of the jury receives a private signal drawn as follows: with probability p, she is informed of the true state, and with probability $1 - p$, she is informed of the wrong state. In the deliberation, each member decides according to the value of her signal and the final decision is taken by majority rule. We then have the following:

Theorem 56 (Condorcet's Jury Theorem)**.** *Let $h_n(p)$ be the probability that the final decision is correct,*

1. *If $p > 1/2$, then $h_n(p) \to 1$ as $n \to \infty$;*
2. *If $p < 1/2$, then $h_n(p) \to 0$ as $n \to \infty$.*

This simple example does not involve any assumption on the beliefs of the members of the jury and only considers a simple aggregation rule. Yet, it introduces the basic framework upon which several strands of literature emerged studying the dynamics of information aggregation. It is important to note at this point that although the situation described above does not define a game *per se*, as there are no payoffs, one may translate this problem into a strategic framework by introducing some utility function $u : \{G, I\}^2 \to \mathbb{R}$. A profile $a = (a_1, \ldots, a_n) \in \{G, I\}^n$ of decisions made by juries yields a final choice $m(a)$ where m is the majority function. With the true state being θ, players receive a payoff given by $u(m(a, \theta))$ and it is assumed that:

$$u(G, G) > u(G, I) \tag{3.1}$$

$$\text{and } u(I, I) > u(G, I). \tag{3.2}$$

Condorcet's model is an early example of belief aggregation. It misses many elements that appear in the recent literature: a dynamic setup, learning heuristics and network effects. Yet, most of the literature dealing with learning in networks carries elements of this example as a modeling basis: some state of the world and a population of agents each having private information on this state and some communication happening either directly or by observing the choices of surrounding players. In the nebula of models building on this initial setup, two main families are distinguished based on how beliefs are modeled: frameworks with agents updating their beliefs according to Bayes rule on the one hand, and models where, on the contrary, agents are assumed to bear some limitations in their rational computation of probabilities. In those two strands of models, similar questions arise naturally: how does opinion exchange behave dynamically, do beliefs converge in the long-run and, if so, do agents end up agreeing on a consensus? Further, when a consensus emerges, do agents learn the true state of the world, or put differently, is the consensus belief equal to a Dirac mass on the true state of the world?

3.3 BAYESIAN LEARNING IN NETWORKS

In this section, we present the foundational elements of Bayesian learning models in networks, an overview of the existing literature, and critical elements which support the use of bounded rationality modeling.

3.3.1 THE BAYES RULE AND BELIEF FORMATION

The Bayes rule stands as one of the most intuitive yet subtle results in the realm of probability theory. In its simplest form, it states that if A and B are two events of a probability space, then the probability that A is true conditional on B being true is given by

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{3.3}$$

Or, by symmetry,

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)} \tag{3.4}$$

In a economic framework, this translates as the foundation of rational information computing. Consider a agent facing uncertainty on some unknown state of the world θ drawn from a set Θ. This agent holds some *prior* belief on the possible values, denoted $\mathbb{P}(\theta)$. Upon receiving some state-relevant *signal* s drawn from some set S, a Bayesian agent is assumed to update her belief on θ by computing the conditional probability $\mathbb{P}(\theta|s)$, that is:

$$\mathbb{P}(\theta|s) = \frac{\mathbb{P}(s|\theta) \cdot \mathbb{P}(\theta)}{\mathbb{P}(s)} \tag{3.5}$$

This requires that the agent knows the right-hand side probability measures $\mathbb{P}(s|\theta)$, $\mathbb{P}(\theta)$and $\mathbb{P}(s)$. Assuming both knowledge and exact computational capability is of course not innocuous and largely debated. For foundational material, we refer the reader to Savage [141]. For a critical perspective on the connection between the Bayesian paradigm and rationality, see e.g. Gilboa [72]. Beyond foundational critiques and tractability issues, especially in the context of a network of agents, Bayesian models remain an essential benchmark of an idealized rational behavior.

3.3.2 NETWORKS OF BAYESIAN AGENTS

The Bayesian literature on the emergence of consensus mostly started with Aumann [12] and its seminal result that two agents with equal prior beliefs and common knowledge posteriors must have equal beliefs. Generalizations have been proposed by Geanakoplos and Polemarchakis [70] and Parikh and Krasucki [123] who showed respectively that two players repeatedly communicating must agree in the long run and that a finite number of players communicating in pairs will eventually agree. To some extent, Parikh and Krasucki [123] may be seen as the first model characterizing the emergence of consensus in a network of Bayesian agents.

They consider a finite set N of n players. A state of the world θ is drawn from a finite set Θ according to a common-knowledge prior distribution $\mu \in \Delta(\Theta)$. Additionally, players receive some private information: in Aumann's framework, each player $i \in N$ has a partition P_i on Θ. A key element in the model is the communication *protocol*: at discrete time t, a pair of players $(s(t), r(t))$ is selected and $s(t)$ transmits a message to $r(t)$. A protocol is a pair of functions (s, r) mapping the natural numbers to the player set N. At a given time t, let $C(\theta, i, t)$ be the set of possible states for player i given that the true state is θ. When a pair $(s(t), r(t)) = (i, j)$ is selected, player i displays some information to player j in the form of a number $f(C(\theta, i, t))$ given by a mapping f from 2^Θ to some arbitrary domain D. A protocol (s, r) generates a directed graph $\mathcal{N}$ whose vertices are elements of N and, for any pair $(i, j) \in N^2$, there is an edge from i to j if i sends a message to j infinitely often.

Theorem 57 (Parikh and Krasucki [123])**.** *There exists a time t_0 such that for any $\theta \in \Theta$, $i \in N$ and $t, t' > t_0 \in \mathbb{N}$, $C(\theta, i, t) = C(\theta, i, t')$. Moreover, if f is convex and $\mathcal{N}$ is strongly connected, $f(C(\theta, i, t)) = f(C(\theta, j, t))$ for all $i, j \in N$.*

This theorem states that if a finite set of Bayesian agents communicate truthfully infinitely often with each other both as a sender and as a receiver, then in the limit they must all share the same belief. In a similar direction, Bala and Goyal [15] proves that when agents are embedded in a connected undirected network and observe the outcome of their actions with some noise, rather than communicating truthfully, players are able to learn the true payoff distributions for actions that their neighbors take infinitely often, hence all actions converge to a consensual action. It is worth noting that although Bala and Goyal [15] considers Bayesian agents, they limit their ability to compute beliefs by assuming that they do not make inferences on unobserved players and behave myopically.

Gale and Kariv [69] extends Bala and Goyal [15] by showing that in any connected network of privately informed Bayesian players observing the actions of their neighbors, actions chosen by all players converge to the same value in finite time with probability one. Note that, as is Bala and Goyal [15], while actions converge, beliefs may not and the limit consensual action may be suboptimal. These results are closely related to social learning models and the observational learning literature as Banerjee [16], Smith and Sørensen [145] or Rosenberg, Solan, and Vieille [130]. Mossel et al. [117] generalizes those results to a large class of social learning models by introducing the concept of *social learning equilibrium* to study the asymptotic properties of learning processes and characterize conditions that agreement and herding behavior. In the same line, Acemoglu et al. [3] and Acemoglu, Bimpikis, and Ozdaglar [2] connect the emergence of social learning with Bayesian agents and the topology of the communication network.

The Bayesian framework is arguably at the core of policies aimed at countering misinformation. Most of them are based on the assumption that agents behave rationally when it comes to information processing. By displaying the limited trustworthiness level of a spurious source, they assume agents will revise their beliefs over secure sources and naturally evacuate false news. As such, when platforms and news sources developed fact-checking programs, it was expected that when confronting the public with factual evidences against false information, viewers would revise their beliefs and prevent further spread. Similarly, Facebook implemented the "Trust Indicator", which consists in associating each news content-maker to a score of trustworthiness established by an independent agency. The underlying idea was that agents would filter spurious content by themselves if provided a proxy of the probability that the source is accurate. The persistence of detrimental spread of misinformation despite the implementation of such measures, as well as the identification of critical biases in belief formation by behavioral studies, show the limits of this modeling in a large-scale context.

3.4 LEARNING WITH BOUNDEDLY RATIONAL AGENTS

Issues in the tractability of Bayesian models as well as critical views on rationality have pushed the exploration of non-Bayesian – or bounded rationality – models. Non-Bayesian consensus models emerged through DeGroot [52], where the author introduces a model where agents living in a network repeatedly exchange their beliefs over some state of the world. At each stage, each agent replaces her belief by the average of her neighbors' beliefs. It is shown that if the communication network is connected, beliefs converge to a consensus which depends on initial beliefs and the network topology only.

3.4.1 CONSENSUS IN THE DEGROOT MODEL

Consider a finite set $N = \{1, \dots, n\}$ of agents embedded in a network $\mathcal{N} = (N, \mathcal{E})$ with edge set $\mathcal{E}$. The network $\mathcal{N}$ is possibly weighted and directed, and we denote by $A \in \mathcal{M}_n(\mathbb{R})$ its weighted adjacency matrix. In other terms, any entry A_{ij} denotes the weight agent i puts on the opinion of agent j. Entries are assumed to be non-negative and such that for any $i \in N$, $\sum_j A_{ij} = 1$. There is an unobservable state of the world θ drawn from a binary state space Θ. For clarity, we will assume that $\Theta = \{0, 1\}$. For any $i \in N$, let $\mu_i(t)$ be the belief of i on the event $\{\theta = 0\}$ at time t, and let $\mu(t) = (\mu_1(t), \dots, \mu_n(t))$ be the probability vector of beliefs at time t. Initially, each agent has a private belief $\mu_i(0) \in [0, 1]$.

So far, the model is close to any Bayesian updating model. Where it departs from this literature is in the definition of the updating process. Namely, time is discrete and, at any stage t, each agent i updates her belief according to the following rule:

$$\mu_i(t+1) = \sum_{j=1}^{n} A_{ij}\mu_j(t) \tag{3.6}$$

or, in matrix form,

$$\mu(t+1) = A\mu(t) \tag{3.7}$$

At each period, every agent replaces her opinion by a weighted average of her neighbors, hence this learning dynamics is often called "local averaging dynamics". Variations on the updating rule have been proposed, for instance in Friedkin and Johnsen [62] where authors allow some persistence on agents beliefs by including one's persistent belief in the averaging process. Namely, each agent i has a private belief ν_i and a weight $\alpha_i \in [0,1]$ and updates her belief as follows:

$$\mu_i(t+1) = \alpha_i \sum_{j=1}^{n} A_{ij}\mu_j(t) + (1-\alpha_i)\nu_i \tag{3.8}$$

Other models involve a time-varying and/or non-deterministic matrix $A(t)$ (see e.g. Chatterjee and Seneta [37] and DeMarzo, Vayanos, and Zwiebel [53]). Throughout the present chapter and Chapter 4, we will not consider such extensions and keep all the network and possible weights fixed. Similar to the Bayesian case, the objective is to characterize the conditions under which the opinion exchange dynamics converge and to determine whether this limit consists in a consensus value or not. Hopefully, this study is heavily simplified by Markovian properties of the process. Indeed, observe that the matrix A is (row-)stochastic and, as such, is the transition matrix of a finite discrete time Markov chain. Thus, convergence properties of the belief exchange process can be pinned down by properties of A. Before presenting major results on convergence, we require the following definitions:

Definition 58. A subset of agents C is *closed* if and only if there does not exist a pair $i, j \in N$ such that $i \in C$, $j \notin C$ and $A_{ij} > 0$.

In the case of undirected networks, a strongly-connected closed set of agents is referred to as a *connected component*. In Markov chain theory, closed sets of agents correspond to the communicating classes of the associated chain. We recall that a chain is said to be *irreducible* if it has a single communicating class. Similarly, we introduce the definition of *aperiodic* player sets:

Definition 59. For any player $i \in N$, we define the *period* d_i of i as the length of the longest cycle containing i: $d_i = \max \left\{ d \in \mathbb{N} | A_{ii}^d > 0 \right\}$. For any set of players C, the period d_C of C is defined as the greatest common denominator of the set $\{d_i | i \in C\}$.

For any two $i, j \in N$ belonging to the same closed set C, $d_i = d_j$.

Definition 60. A player $i \in N$ is said to be *aperiodic* if $d_i = 1$. Similarly, a closed set of players is said to be *aperiodic* if $d_C = 1$.

We then have the following convergence property:

Theorem 61. *The opinion exchange process converges if and only if every strongly-connected and closed set of nodes is aperiodic.*

A full proof of this result is available in Golub and Jackson [74]. After having established convergence, consensus arises naturally. In the limit of the belief averaging process, it is trivial that any strongly-connected and aperiodic group of players cannot disagree. Before detailing the conditions behind the emergence consensus, we formally define it:

Definition 62. We say a subset $C \subseteq N$ of agents reaches a *consensus* if, given the matrix A and a vector of prior beliefs μ_0, we have that for any two $i, j \in C$, $\lim_t \mu_i(t) = \lim_t \mu_j(t)$.

We then have the following:

Proposition 63. *A consensus is reached in the DeGroot model if and only if the matrix A is irreducible and aperiodic.*

3.4.2 INSTANCES OF NAIVE LEARNING

To illustrate Proposition 63, we present two examples of a converging and a non-converging communication networks, borrowed from Jackson [86].

Consider three agents connected as depicted in the communication network displayed in Fig. 3.1.

$$A = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

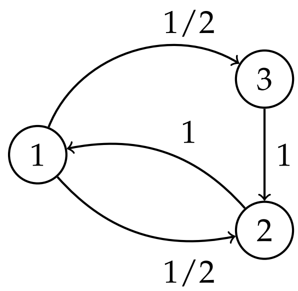


Figure 3.1: Irreducible and aperiodic communication network

In this first example, the matrix A is irreducible and aperiodic. Powers of A converge to a limit

$$A^\infty = \begin{pmatrix} 2/5 & 2/5 & 1/5 \\ 2/5 & 2/5 & 1/5 \\ 2/5 & 2/5 & 1/5 \end{pmatrix} \tag{3.9}$$

This implies that for any initial vector of beliefs $\mu(0) = (\mu_1(0), \mu_2(0), \mu_3(0))$, the communication process converges to a consensus, which is given by $\lim_t A^t \mu(0)$:

$$\mu_1(\infty) = \mu_2(\infty) = \mu_3(\infty) = \frac{2}{5}\mu_1(0) + \frac{2}{5}\mu_2(0) + \frac{1}{5}\mu_3(0) \quad (3.10)$$

But if, instead, the communication network is given by the one in Fig. 3.2, then there is no convergence anymore.

$$A = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

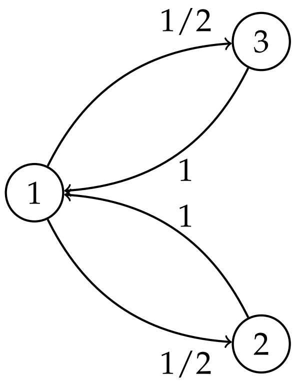


Figure 3.2: Irreducible and periodic communication network

The network is indeed periodic of period 2: $A = A^3 = \ldots$ and $A^2 = A^4 = \ldots$ hence agents will keep swapping their beliefs over time.

In the DeGroot model, the emergence of consensus is characterized by elementary properties of the communication network $\mathcal{N}$. While these properties were also proved in the Bayesian case, although at a higher technical cost, the important benefit from resorting to bounded rationality models is the analytical tractability of the consensus belief. The example in Fig. 3.1 hints at the fact that, given an irreducible and aperiodic communication network $\mathcal{N}$, its limit consensus is analytically determined by the product of the stationary distribution of $\mathcal{N}$ and the vector of initial beliefs. This fact follows directly from Eq. (3.7): if the sequence $(A^t)_t$ admits a limit A^∞ then one immediately has

$$\mu(\infty) = A^\infty \mu(0). \quad (3.11)$$

Note that as $\mathcal{N}$ is assumed to be finite, having both irreducibility and aperiodicity is a sufficient condition to ensure unicity of a stationary distribution π_A, and the consensus belief is also given by

$$\mu(\infty) = \pi'_{adja}\mu(0) \tag{3.12}$$

The DeGroot model gained popularity in the economic literature with Golub and Jackson [74] which connects consensus and learning with the network's adjacency matrix using properties of Markov chains steady-states and provides interpretations in terms of centrality measures and influence. They refer to DeGroot's belief averaging dynamics as *naive learning*. Examples of the use of DeGroot [52] in economic modeling are too numerous to be listed. In the recent literature, Mandel and Venel [112] considers a stochastic game where two misinformers try to influence a population of agents applying naive learning. We refer the reader to the aforementioned surveys Acemoglu and Ozdaglar [5], Mossel and Tamuz [118] and Golub and Sadler [75] for further applications and variations on the model. A more recent reference is Grabisch and Rusinowska [76].

3.5 FROM NAIVE TO REINFORCEMENT LEARNING

3.5.1 Limitations of naive learning

There is a vast body of literature questioning the robustness of learning dynamics.

Criticism on the DeGroot model already featured in Golub and Jackson [74] where authors proved that in the general case, beliefs do not converge for countably infinite player sets. In a recent work, Peretz et al. [125] shows that in the presence of agents with fixed beliefs over time, which they call *bots*, the common limit can converge to any value.

In our view, two major limitations are to be opposed to models based on DeGroot dynamics. First, they have been shown to have limited robustness, in the sense that the repeated averaging overweights initial beliefs while enforcing fast convergence of beliefs. This is particularly the case when some agent i puts some weight on her own belief, i.e. the entry A_{ii} is strictly positive. Consider for instance the following communication network:

$$A = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/10 & 1/10 & 8/10 \end{pmatrix} \tag{3.13}$$

The stationary distribution of the matrix is given by the vector $\pi_A = (\frac{1}{7}, \frac{1}{7}, \frac{5}{7})$. We see here that as agent 3 has in important weight on herself, her final influence on the limit consensus is much stronger than that of agents 1 and 2.

Second, one may question the relevance of a setup where agents directly access and exchange their beliefs on some state. An important consequence of the belief averaging dynamics is that it does not leave room for the occurrence of extreme events like the spread of a marginal belief to a substantial part of a population. As such, using non-Bayesian learning models as a reference in order to build policies countering misinformation may be inefficient. The two core elements of the process, namely the exact access by agents to their beliefs and the averaging equation, yield a deterministic result which can be interpreted as the outcome of Bayesian models, that is, an idealized benchmark. In most small-world communication setups, it seems more realistic to assume that agents do not access the subjective probabilities they put on the possible values of a state, but rather decide to relay some information over another according to the relative probabilities they put on those events. The more they believe in some event, the more likely it is they will transmit that information to their neighbors. In other terms, they do not exchange beliefs but communicate possible states by draws based on their beliefs. In this regard, we introduce and analyze a stochastic variant of DeGroot dynamics where agents behave in this respect.

Introducing some degree of stochasticity strongly changes the perspective on misinformation: in DeGroot [52], slightly modifying the prior beliefs of some agents cannot change drastically the consensus outcome. Yet, most disinformation platforms display some "shadow of doubt" strategy: agents do not transmit false information because they necessarily believe it to be true, but rather because there is some – even small – probability that it may not be false. In other words, they manage to disinform by inducing limited beliefs on their false information. The goal of a disinformer is not to convince but to weave some doubt, as the means of establishing a controversy.

3.5.2 POLYA URN MODELS IN PRACTICE

Our technical approach to this modeling problem is based on the seminal Polya urn model from Eggenberger and Pólya [57] presented in Chapter 2. In this paper, authors consider an urn with balls of several colors and study the convergence of reinforcement dynamics. It is a well known result that proportions in the urn converge to a beta distribution (see Klenke [96] for instance). The strength of Polya's model is its intricate connection with exchangeability. Central papers in the foundation of Bayesian inference like De Finetti [51] and Hewitt and Savage [81] heavily rely on the concept of exchangeability. Polya urn plays a particular role in that Hill, Lane, Sudderth, et al. [82] proved that any exchangeable process of $\{0, 1\}$-valued random variables is either Bernoulli, deterministic or generated by draws from a Polya urn. Our model considers a system of interacting urns in the flavor of Paganoni and Secchi [120] which introduced such systems and first proved convergence when the number of balls in all urns grow at the same speed. Similarly, Dai Pra, Louis, and Minelli [50] shows convergence of proportions in a system where urn reinforcement depends both on proportions in each urn and on the average proportions in the system. Crimaldi, Dai Pra, and Minelli [49] gives further results on convergence and fluctuations around the limit of such system. The model we consider is close yet different, as we consider a system where urns are reinforced at different speeds which correspond to their degree in the communication network. Usual probabilistic tools do not apply as proportions are not martingales and draws are not exchangeable. Instead, we rely on stochastic approximation as introduced by Robbins and Monro [128]. Motivation for the use of stochastic approximation in the study of urn systems can be found in Laruelle, Pages, et al. [103], where authors use this technique in the context of clinical trial modeling.

3.6 CONTRIBUTIONS OF THE CHAPTER

In the model we develop in Chapter 4, we build an opinion formation model where agents communicate by drawing states according to their beliefs instead of directly communicating subjective probabilities. To do so, we model beliefs using reinforcing urns. Studying the evolution of beliefs comes down to characterizing the evolution of urns' compositions. Using stochastic approximation techniques, we show that in such models, under very general conditions, the dynamics of beliefs converge to a rest point. We show that at the steady-state, all agents share the exact same belief on the state of the world. As long as initial beliefs cover the whole state space, the consensus is drawn from a distribution with full support. This strongly contradicts the predictions of DeGroot's and similar models. We then try to characterize this limit distribution using simulations.

4 STOCHASTIC CONSENSUS AND THE SHADOW OF DOUBT

But I think, in general, it's clear that most bad things come from misunderstanding, and communication is generally the way to resolve misunderstandings — and the Web's a form of communications — so it generally should be good.

Tim Berners-Lee
(developerWorks Interviews, 2006)

In this chapter, we present our contribution to the problem raised in Chapter 3. In Section 4.1, we begin by exposing our stochastic extension of non-Bayesian opinion formation models. Our model carries much of the elements already present in the DeGroot model, but transforms the exchange dynamics into a reinforcement learning process. Section 4.2 presents our theoretical contribution. We characterize conditions required to ensure the emergence of a consensus and prove that, when it exists and under mild conditions, the limit belief is a random variable with full support. Section 4.3 provides empirical elements to characterize the distribution of the limit belief based on simulations. Section 4.4 concludes both the chapter and Part i.

4.1 A STOCHASTIC MODEL OF OPINION EXCHANGE

4.1.1 AGENTS AND COMMUNICATION NETWORKS

We consider a finite population $N = \{1, \dots n\}$ of n agents, embedded in a exogeneous, fixed and undirected graph $\mathcal{N} = (N, \mathcal{E})$ with edge set $\mathcal{E}$. For any agent i, we denote by $\mathcal{N}(i)$ the neighborhood of i and by d_i the degree of i. Let $A = (A_{ij})_{i,j \in N}$ be the adjacency matrix of $\mathcal{N}$, with the convention that $A_{ii} = 0$ for all $i \in N$. Unlike in DeGroot [52] and the resulting literature, we consider an unweighted graph and do not require the matrix A to be stochastic. Finally, let D be the diagonal matrix of degrees, that is, the matrix with values d_i on the diagonal and zero elsewhere.

There is a binary state space Θ, e.g. $\Theta = \{0, 1\}$. At the beginning of the game, a state θ is drawn at random from Θ and is unobserved.

The only information agents will access throughout the communication process originates from an initial noisy private signal informing them about the state of the world. We model this signal in the same flavour as Condorcet [46]. Formally, let $\alpha \in [0, 1]$, any player i receives a signal s_i displaying one of the two possible states. With θ the true state,

$$\begin{cases} s_i = \theta \text{ with probability } \alpha \\ s_i = 1 - \theta \text{ with probability } 1 - \alpha \end{cases} \tag{4.1}$$

For a large network, by the law of large numbers, α represents the average proportion of agents initially well-informed. Throughout, we will assume that every player i has at least one neighbor in $\mathcal{N}$. This rules out trivial cases where some agents would be isolated within the communication process, hence their beliefs would remain at their original value $\mu_i(0)$.

4.1.2 BELIEFS AND THE COMMUNICATION PROCESS

Our objective is to extend naive learning models to a setup where agents communicate by signaling states drawn at random according to their relative beliefs. To this end, we model beliefs using urns of infinite capacity with red and blue balls of representing the possible values of the state θ. For any observed signal indicating one of those two events, agents reinforce their urns by balls of the corresponding color. In this sense, at any time $t \geq 1$, the proportions of a given color in agent i's urn then represent her belief over the corresponding event. At time $t = 0$, urns are initialized with a ball corresponding to the agent's signal s_i.

Time is discrete and, at each stage $t \geq 1$, the communication process is as follows:

1. First, each agent draws a ball from her urn with uniform probability and replaces it;
2. Then, each agent observes the draws from her neighbors;
3. For any observed draw, agents reinforce their urns by adding one ball of the corresponding color.

We introduce the following notational elements:

$$\begin{cases} r_i^t & = \text{number of red balls in urn } i \text{ at the end of period } t, \\ b_i^t & = \text{number of blue balls in urn } i \text{ at the end of period } t, \\ s_i^t = r_i^t + b_i^t & = \text{total number of balls in urn } i \text{ at the end of period } t, \\ z_i^t = \frac{b_i^t}{s_i^t} & = \text{proportion of blue balls in urn } i \text{ at the end of period } t, \\ x_i^t & = \text{indicator function of a blue draw from urn } i \text{ at period } t. \end{cases}$$

Let $z^t = (z_1^t, \ldots, z_n^t)$, $x^t = (x_1^t, \ldots, x_n^t)$. Finally, let $\mathcal{F}^t = \sigma\left(\left\{x^k\right\}, k \leq t\right)$ be the sigma-field generated by the sequence of draws and $\mathcal{F}$ be the corresponding filtration. For any $i, j \in N$ and any $t \geq 1$, x_i^t and x_j^t are assumed to be independent. Therefore, for a given value of α, there exists a unique probability measure on the product set Ω of infinite sequences of draws, which we will denote $\mathbb{P}_\alpha$. The expectation with respect to this measure is denoted $\mathbb{E}_\alpha$. The communication process is repeated infinitely.

At any given time, an agent's current belief on θ is given by the proportions of balls in her urn $(z_i^t, 1 - z_i^t)$. The communication protocol defines a stochastic process over an interacting urn system. Our objective is to study the evolution of the urn system and determine whether proportions converge, if a consensus is reached and if so, to characterize it given the network topology and the value of α.

4.1.3 INTRODUCTORY EXAMPLE

Consider three agents connected in line as displayed in Fig. 4.1. Assume that at time $t = 0$, agent 1 and 3 received a signal $s_1 = s_3 = \theta$ and agent 2 received $s_2 = 1 - \theta$. Then at time $t = 0$ urns 1 and 3 will contain a blue ball and urn 2 will contain one red ball. At time $t = 1$, every player draws the only ball their urns contain and display it. Then, they all replace their draw and add a new ball of the corresponding color for every draw they observe. That is, at the end of the first stage, urns 1 and 3 will contain each one blue ball and one red ball, and urn 2 will contain one red ball and two blue balls. Given the initialization procedure detailed above, the first stage of the communication process is always deterministic.

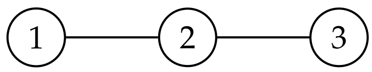


Figure 4.1: Three urns in line.

At time $t = 2$, the communication protocol is repeated: the end urns will draw one blue or one red ball with equal probability and the middle urn will draw one blue ball with probability $2/3$ and a red ball with probability $1/3$.

Table 4.1 details the possible outcomes at time $t = 2$. The left column is the vector of draws from urn 1, 2 and 3 respectively and the right column gives the compositions at the end of the time period in the same order. Left figures correspond to the number of blue balls in the urn and right figures to the number of red balls.

Vector of draws	Probability of occurence	Urns compositions
(B,B,B)	1/6	(2,1)–(4,1)–(2,1)
(B,R,B)	1/12	(1,2)–(4,1)–(1,2)
(R,B,B)	1/6	(2,1)–(3,2)–(2,1)
(R,R,B)	1/12	(1,2)–(3,2)–(1,2)
(B,B,R)	1/6	(2,1)–(3,2)–(2,1)
(B,R,R)	1/12	(1,2)–(3,2)–(1,2)
(R,B,R)	1/6	(2,1)–(1,4)–(2,1)
(R,R,R)	1/12	(1,2)–(1,4)–(1,2)

Table 4.1: Possible outcomes at time $t = 2$.

4.2 BELIEF DYNAMICS AND CONVERGENCE

In this section, we present the analytical results we obtain in terms of convergence of beliefs.

4.2.1 URN DYNAMICS

The sequence of proportions $(z^t)_t$ as defined in Section 4.1 forms a recursive system whose dynamics can be derived as follows:

$$\begin{cases} b_i^{t+1} = b_i^t + \sum_{j\in\mathcal{N}(i)} x_j^{t+1} \\ r_i^{t+1} = r_i^t + d_i - \sum_{j\in\mathcal{N}(i)} x_j^{t+1} \\ s_i^{t+1} = s_i^t + d_i = s_i^0 + d_i(t+1) \end{cases} \tag{4.2}$$

Additionaly, by definition we have that for any $i \in N$ and $t \geq 1$,

$$x_i^t \hookrightarrow \mathcal{B}\left(z_i^{t-1}\right) \tag{4.3}$$

Hence

$$\mathbb{E}\left[b_i^{t+1} - b_i^t | \mathcal{F}_t\right] = \mathbb{E}\left[\sum_{j \in \mathcal{N}(i)} x_j^{t+1} | \mathcal{F}_t\right] = \sum_{j \in \mathcal{N}(i)} z_j^t \tag{4.4}$$

Eq. (4.4) shows that, in expectation, the belief updating process obeys some local averaging property as in canonical naive learning models: the variation of proportion in any urn evolves according to the sum of the proportions in the neighboring urns.

4.2.2 CONVERGENCE OF BELIEFS

We first show that beliefs converge in the sense that color proportions in each urn converge to a stable point. The proof relies on SA techniques, as probabilistic methods do not apply in our case. Indeed, unless the graph $\mathcal{N}$ is regular, neither local nor global proportions are martingales and it is easy to see that the process (z^t) is not exchangeable as the rate at which the number of balls in an urns evolve depends both on the urn's degree and on time. Nevertheless, we are able to write the dynamics as a recursive algorithm for which we can prove convergence.

Theorem 64. *For every $\alpha \in [0,1]$ and any graph $\mathcal{N}$, $\lim_{t \to \infty} z_i^t$ exists $\mathbb{P}_\alpha$-almost-surely for all $i \in N$.*

Proof. From Eq. (4.2) we derive the following recursive formula on z_i^t:

$$z_i^{t+1} - z_i^t = \frac{-d_i z_i^t + \sum_{j \in \mathcal{N}(i)} x_j^{t+1}}{s_i^0 + d_i(t+1)} \tag{4.5}$$

By adding and subtracting the conditional expectation of the number of red draws in neighboring urns to the numerator, we have:

$$\begin{aligned} z_i^{t+1} - z_i^t &= \frac{1}{s_i^0 + d_i(t+1)} \left(-d_i z_i^t + \mathbb{E}\left[\sum_{j \in \mathcal{N}(i)} x_j^{t+1} | \mathcal{F}_t \right] \right) \\ &+ \frac{1}{s_i^0 + d_i(t+1)} \left(\sum_{j \in \mathcal{N}(i)} x_j^{t+1} - \mathbb{E}\left[\sum_{j \in \mathcal{N}(i)} x_j^{t+1} | \mathcal{F}_t \right] \right) \end{aligned} \tag{4.6}$$

Observing that by Eq. (4.3), conditional on $\mathcal{F}_t$, the expected number of blue draws in neighboring urns at time $t+1$ is equal to the sum of their proportions at time t, we obtain:

$$\begin{aligned} z_i^{t+1} - z_i^t &= \frac{1}{s_i^0 + d_i(t+1)} \left(-d_i z_i^t + \sum_{j \in \mathcal{N}(i)} z_j^t \right) \\ &+ \frac{1}{s_i^0 + d_i(t+1)} \left(\sum_{j \in \mathcal{N}(i)} x_j^{t+1} - \mathbb{E}\left[\sum_{j \in \mathcal{N}(i)} x_j^{t+1} | \mathcal{F}_t \right] \right) \end{aligned} \tag{4.7}$$

We now rescale the equation by a factor that is independent of d_i:

$$\begin{aligned} z_i^{t+1} - z_i^t &= \frac{\left(\frac{1 + (t+1)}{s_i^0 + d_i(t+1)} \right)}{1 + (t+1)} \left(-d_i z_i^t + \sum_{j \in \mathcal{N}(i)} z_j^t \right) \\ &+ \frac{\left(\frac{1 + (t+1)}{s_i^0 + d_i(t+1)} \right)}{1 + (t+1)} \left(\sum_{j \in \mathcal{N}(i)} x_j^{t+1} - \mathbb{E}\left[\sum_{j \in \mathcal{N}(i)} x_j^{t+1} | \mathcal{F}_t \right] \right) \end{aligned} \tag{4.8}$$

We simplify this system by introducing the following shorthand notations. Let

$$\begin{aligned} f^t : [0,1]^N &\to [0,1]^N \\ z^t &\mapsto (f_1^t(z^t), \dots, f_N{}^t(z^t)) \\ \text{with } f_i^t(z^t) &= \frac{1+t}{s_i^0 + d_i t} \sum_{j \in \mathcal{N}(i)} z_j^t - d_i z_i^t. \end{aligned} \tag{4.9}$$

The function f_i^t is the deterministic part – or mean-flow – of the stochastic system mapping z^{t-1} to z^t. Let

$$\begin{aligned} u^t &= (u_1^t, \dots, u_N{}^t), \\ \text{with } u_i^t &= \frac{1+t}{1+d_i t} \sum_{j \in \mathcal{N}(i)} x_j^{t+1} - \mathbb{E}\left[\sum_{j \in \mathcal{N}(i)} x_j^{t+1} | \mathcal{F}_t \right] \end{aligned} \tag{4.10}$$

The elements u_i^t are random variables (r.v.s) measuring how the realized proportions z_i^t move away from their mean-flow value. Finally let

$$\gamma^t = \frac{1}{1 + (t+1)} \tag{4.11}$$

The sequence (γ_t) may be seen as a sequence of step sizes at which the system evolves.

We can rewrite Eq. (4.8) in matrix form, which yields the much nicer system:

$$z^{t+1} - z^t = \gamma^t \left[f^t(z^t) + u^t \right] \tag{4.12}$$

The system Eq. (4.12) forms a s.a.a. in the sense of Robbins and Monro [128]. It describes the evolution of a discrete-time stochastic process, z^t, as the γ^t-weighted sum of a deterministic continuous-time map f^t and a random residual u^t. SA theory aims at establishing conditions that ensure trajectories of the discrete-time stochastic differential system Eq. (4.12) will approach those of a suitable continuous-time o.d.e. in the limit.

We now study the asymptotic behavior of Eq. (4.12). In order to ensure convergence of the stochastic system, we first make the following observations.

Observation 1. The sequence $(\gamma_t)_t$ satisfies

$$\begin{cases} \sum_{t=1}^{\infty} \gamma^t = \infty \\ \sum_{t=1}^{\infty} \left(\gamma^t\right)^2 < \infty \end{cases} \tag{4.13}$$

Observation 1 is central in any s.a.a. *à la* Robbins and Monro [128] with deterministic weights. These weights serve as the increments of time discretization. In that perspective, the first point implies that the algorithm will cover the entire time interval. The second point involves, jointly with the next observation, the disappearing of noises $\sum_t u_i^t$ in the limit. As γ^t is of the order of $1/t$, Observation 1 is immediate.

Observation 2. For every i in N, the sequence (u_i^t) is a martingale difference noise relative to $\mathcal{F}_t$.

Observation 2, when combined with the second point in Observation 1, ensures that the cumulative error due to the discretization noise is negligible almost-surely, as the noise variance will vanish asymptotically. Observation 2 holds as, for any $i \in N$, the sequence (u_i^t) is a sequence of bounded r.v.s with zero mean.

Observation 3. The maps f_i^t are Lipschitz continuous and measurable with respect to $\mathcal{F}_t$ and uniformly continuous in t for $t \geq 1$.

SA ensures that a discrete-time stochastic process evolves along the trajectories of a continuous-time o.d.e.. In that respect, Observation 3 ensures that the o.d.e. is well defined and has a unique solution.

Finally, although the maps $(f_i^t)_{i,t}$ in Eq. (4.12) depend on time, for any $i \in N$, the sequence of maps (f_i^t) converge to a time-independent limit as time goes to infinity. Indeed, for any $i \in N$ and any $z \in [0,1]^n$, let $\bar{f}_i(z) = \frac{1}{d_i}\sum_{j\in\mathcal{N}(i)} z_j - d_i z_i$ and $\bar{f} : z \mapsto (\bar{f}_1(z), \dots, \bar{f}_n(z))$.

Observation 4. For any $z \in [0,1]^N$ and any $k \in \mathbb{N}^*$,

$$\lim_{s\to\infty}\left|\sum_{t=s}^{s+k}\gamma^t\left[f_i^t(z) - \bar{f}_i(z)\right]\right| \to 0$$

Observation 4 holds immediately as, for any $i \in N$, $f_i^t \to \bar{f}_i$ as $t \to \infty$.

Based on Observations 1–4, we can apply Theorem 2.3 from Kushner and Yin [100, chapter 5].

Theorem 65 (Kushner and Yin [100, chapter 5]). *If Observations 1–4 hold and (z^t) is bounded with probability one, then for $\mathbb{P}_\alpha$-almost all $\omega \in \Omega$, the limits $\bar{z}(\omega)$ of convergent subsequences of $(z^t(\omega))$ are trajectories of*

$$\dot{z^t} = \bar{f}(z^t) \tag{4.14}$$

in some bounded invariant set and $(z^t(\omega))$ converges to this invariant set.

This result ensures that the system Eq. (4.12) evolves almost-surely along trajectories of Eq. (4.14) and converges to some invariant set of the o.d.e. system, that is, a set $M \subseteq [0,1]^n$ such that $\bar{f}(M) = M$ where $\bar{f}(M)$ denotes the image of the set M by the map $\bar{f}$.

To further characterize convergence properties of Eq. (4.8), we turn to the study of limit points of the o.d.e. Eq. (4.14). Observe that Eq. (4.14) is a autonomous linear system with constant coefficients. For any $z \in [0,1]^n$, we can rewrite the system as follows:

$$\dot{z} = \bar{f}(z) \tag{4.15}$$

$$= -\left(\frac{1}{d_1}, \ldots, \frac{1}{d_n}\right)' Lz \tag{4.16}$$

Where L is the Laplacian matrix of $\mathcal{N}$ as introduced in Chapter 2. The behavior of trajectories in the vector field defined by *Eq. (4.14)* is captured by spectral properties of L. More precisely, equilibrium points of *Eq. (4.14)* belong to the nullspace of $-L$, as $\frac{1}{d_i} > 0$ for every $i \in N$.

Going further requires a bit of additional notation. Let $\{1, \ldots, K\}$ be the connected components of $\mathcal{N}$, i.e. maximal subsets N_k of agents such that for any two $i, j \in N_k$ there exists a path from i to j. For any connected component k, let n_k be its number of nodes.

By permuting the agents according to their belonging to a same component, let

$$\mathcal{L}_{\mathcal{N}} = \left\{ \left(\underbrace{a_1, \ldots, a_1}_{n_1 \text{ times}}, \overbrace{a_2, \ldots, a_2}^{n_2 \text{ times}}, \ldots, \underbrace{a_k, \ldots, a_k}_{n_k \text{ times}} \right) | a_1, \ldots, a_k \in [0,1] \right\} \tag{4.17}$$

That is, $\mathcal{L}_{\mathcal{N}}$ is the set of vectors in $[0,1]^n$ with identical entry for all agents belonging to a same component. We now state our main convergence result on Eq. (4.14):

Lemma 66. *Every trajectory of Eq. (4.14) converges to the set $\mathcal{L}_{\mathcal{N}}$.*

Proof of Lemma 66. The proof of Lemma 66 follows entirely from arguments on the properties of the matrix L. First, we show that $\mathcal{L}_{\mathcal{N}}$ is the eigenspace of L associated with the eigenvalue 0. Recall that for any $i, j \in N$,

$$L_{i,j} = \begin{cases} d_i \text{ if } i = j; \\ -1 \text{ if } i \neq j \text{ and } i \text{ and } j \text{ are neighbors;} \\ 0 \text{ elsewhere.} \end{cases} \tag{4.18}$$

The Laplacian matrix of a graph captures crucial information about the topology of the graph. We recall its elemental properties:

Property 67. *For any undirected network $\mathcal{N}$ with non-negative weights,*

1. *L has only real eigenvalues;*
2. *L is positive semidefinite;*
3. *The smallest eigenvalue of L is $\lambda_1 = 0$ with corresponding eigenvector $(1, \ldots, 1)'$. The multiplicity of λ_1 is equal to the number K of connected components of $\mathcal{N}$.*

Positive semidefiniteness follows immediately from the expression of the inner product $\langle Lz, z \rangle = \sum_{i \sim j}(z_i - z_j)^2$. The last property is a bit more difficult to prove. Let $\mathcal{N}$ be a graph consisting in one single connected component. By definition of L, the sum of entries on each row is equal to zero, hence zero is an eigenvalue of $\mathcal{N}$ with associated eigenvector $1_n = (1, \ldots, 1)'$. To prove that the eigenspace associated with zero is of dimension one, let $y \neq 1_n$ be another eigenvector associated to zero. Then, by definition,

$$y'Ly = \sum_{i \sim j}(y_i - y_j)^2 = 0 \tag{4.19}$$

therefore $y \in Vect\,(1_n)$ and the multiplicity of the eigenvalue zero is one. Now, if a graph $\mathcal{N}$ consists in $K > 1$ connected components, then its Laplacian matrix L can be written as a block diagonal matrix where blocks correspond to the Laplacian matrices $L_1, \ldots, L_K$ of its components. By what precedes, each components has an eigenvalue 0 with multiplicity 1, therefore L has the eigenvalue 0 with multiplicity K. Corresponding K eigenvectors are vectors with entry 1 for agents belonging to component k and 0 else, for each component k. This proves that for any graph $\mathcal{N}$, the set $\mathcal{L}_{\mathcal{N}}$ is indeed the nullspace of its Laplacian matrix. For further properties and results on Laplacian matrices, we refer the reader to the survey Mohar [115].

In turn, this implies that the set of equilibrium points – i.e. points z such that $\dot{z} = 0$ – of Eq. (4.14) is equal to $\mathcal{L}_{\mathcal{N}}$. Note that $\mathcal{L}_{\mathcal{N}}$ is trivially an invariant set of Eq. (4.14). From Property 67, we know that L is positive semidefinite, hence $-L$ is negative semidefinite. As a consequence, the set $\mathcal{L}_{\mathcal{N}}$ is exponentially stable. In other terms, once a trajectory of Eq. (4.14) approaches the set $\mathcal{L}_{\mathcal{N}}$, not only does it it remain close to $\mathcal{L}_{\mathcal{N}}$ forever, but it also converges to $\mathcal{L}_{\mathcal{N}}$ at exponential speed. Once it reaches any point p in the set, as $-Lp = 0$, it remains at this point forever.

Finally, it remains to show that $\mathcal{L}_{\mathcal{N}}$ is the only invariant set to which the process z^t converges. This fact follows immediately from the fact that $-L$ is negative semidefinite.

□

Having established Lemma 66, Theorem 64 follows directly since Theorem 65 guarantees that for any $\omega \in \Omega$, $z^t(\omega)$ converges along trajectories of Eq. (4.14). As the latter converge to stable points in $\mathcal{L}_{\mathcal{N}}$, so does $z^t(\omega)$, which concludes the proof.

□

This first result ensures that for any graph structure $\mathcal{N}$ and any initial condition on the urns, proportions converge almost-surely to a stable point. In particular, convergence is independent of the initial signal structure and applies for any alternative initialization of the system. By the construction of the set $\mathcal{L}_{\mathcal{N}}$, we recover the result from Theorem 61 in an undirected context. The next result details the conditions required for a consensus to emerge.

4.3 EMERGENCE OF CONSENSUS

In this section, we will show that in our model, a consensus emerges under the sole condition that the network $\mathcal{N}$ is connected. This condition strongly resembles what the DeGroot model requires to achieve a similar outcome – see Section 3.4.

Theorem 68. *For every $\alpha \in [0,1]$, suppose that the graph G is connected. Then for any $i, j \in N$, $\lim_{t\to\infty} z_i^t = \lim_{t\to\infty} z_j^t$ $\mathbb{P}_\alpha$-almost-surely.*

Proof. From Theorem 64, we know that proportions converge $\mathbb{P}_\alpha$-almost-surely along the trajectories of Eq. (4.14), that is:

$$\dot{z}_i^t = \frac{1}{d_i} \sum_{j\in\mathcal{N}(i)} z_j^t - d_i z_i^t \tag{4.20}$$

As shown in the proof of Theorem 64, limit values of Eq. (4.20) belong to limit values of the following autonomous linear system:

$$\dot{z}^t = -Lz^t \tag{4.21}$$

Where L is the Laplacian matrix of the graph $\mathcal{N}$. From Property 67, we know that if $\mathcal{N}$ is connected, then the nullspace of $-L$ is of dimension 1 with associated eigenvector $(1,\dots,1)'$. By the same arguments than the proof of Theorem 64, it follow if $\mathcal{N}$ is connected, any trajectory of the process z^t converges to the set

$$\mathcal{L}_{\mathcal{N}} = \{(a,\dots,a)\,|a \in [0,1]\} \tag{4.22}$$

which corresponds exactly to the set of beliefs where all agents agree on one consensual belief. □

Throughout the rest of the chapter, we will assume $\mathcal{N}$ is connected hence beliefs converge to a limit consensus. So far, we obtained results very similar in substance to those obtained in the DeGroot model, namely convergence of beliefs in a general setup and emergence of a consensus if the communication network is connected. This confirms the interpretation of our model as a stochastic extension of DeGroot [52]. Yet, while a consensus emerges in both models, its value strongly differs. Recall from Section 3.4 that in naive learning, with a irreducible and aperiodic communication network, the value of the limit belief is given by the product of the stationary distribution of the matrix A and the vector of prior beliefs. In other terms, given a communication network and a vector of initial beliefs, there exists a unique consensus which is analytically tractable *ex-ante*. From this fact originate both the strength and main failure of the DeGroot model: namely an extreme tractability easing the use of the framework in modeling, but not leaving room to realistic phenomena. In that respect, our model features an opposite behavior: it provides realistic predictions in terms of the spread of competing beliefs within a population, at the cost of limited tractability. The next theorem states our main result: that in our stochastic extension of DeGroot [52], the limit consensus is no longer a value but a full-support r.v..

Theorem 69. *For any connected graph $\mathcal{N}$ and any $\alpha \in (0,1)$, the limit belief $\bar{z}$ is a non-trivial r.v. with full support on $[0,1]$.*

Proof. The proof is based on the concept of *attainability* from Benaïm [21].

Definition 70. A point $p \in \mathbb{R}^N$ is *attainable* by z if for every $t > 0$ and every open neighborhood U of p,

$$\mathbb{P}_\alpha\left(\exists s \geq t : z^s \in U\right) > 0.$$

In other terms, a point p is attainable if, from any vector of proportions, there is a strictly positive probability that z^t becomes arbitrarily close to p in finite time.

Lemma 71. *Any point p in the interior of $\mathcal{L}_\mathcal{N}$ is attainable.*

To prove Lemma 71, simply observe that, from Eq. (4.2), $z_i^{t+1} - z_i^t$ is of the order of $\frac{1}{t}$. If $\alpha \in (0,1)$, $\mathbb{P}\left(z_i^t \in (0,1)\right) > 0$ for every i and $t \geq 0$.

We showed that any point in $\mathcal{L}_{\mathcal{N}}$ is attainable. To complete the proof of Theorem 64, it remains to show that any attainable point in $\mathcal{L}_{\mathcal{N}}$ belongs to the support of $\bar{z}$. This follows trivially as, by definition, the support of a r.v. is the smallest closed set with full probability mass. Then, all points in $\mathcal{L}_{\mathcal{N}}$ being both attainable and equilibria is sufficient to conclude.

□

The limit cases where either $\alpha = 0$ or $\alpha = 1$ are trivially solvable, as urns in the system display only one color hence beliefs will remain at their original value forever.

Theorem 69 marks a major difference between our model and the original model of DeGroot [52]. The limit belief being a r.v. with full support as long as $\alpha \in (0,1)$ means that if the probability that an agent is misinformed is strictly positive then, *ex-ante*, the limit belief may take any possible value in $[0,1]$. In particular, occurrences where one single agent spreads some misinformation and overturns the opinion of the entire network may happen. Such phenomenon only appears in extensions of DeGroot [52] where heterogeneous agents are considered.

These results lead naturally to the study of the distribution of the limit belief. The strength of DeGroot [52] is to connect analytically the consensus to premisses of the model, and seeking a similar result is a logical extension of our work. Unfortunately, the characterization of closed-form formulas for the distribution of limit values of a s.a.a. is known to be a seldom solvable problem. For details on the general intractability of limits of interacting urn systems, we refer the reader to Paganoni and Secchi [120] and Crimaldi, Dai Pra, and Minelli [49] where authors detail state of the art methods to obtain partial fluctuation results. This is why we resorted to using simulations in order to obtain further results in this direction. Interestingly, the data obtained from these simulations highlighted a tighter relation between our model and classical instances of reinforcement learning as Eggenberger and Pólya [57]. The next section details the empirical evidence we obtained on the limit distribution of beliefs.

4.4 LIMIT DISTRIBUTION

While our efforts in characterizing the limit distribution of the consensus $\bar{z}$ as a function of $\mathcal{N}$ and α failed, large scale simulations provide some useful evidence. We simulated the learning dynamics on three network structures: stars, regular graphs with varying degree and complete networks, each of size up to $n = 100$. The values of the limit belief were simulated for different values of α. The value of the limit belief was collected after 3000 iterations of the communication process. Main elements of the code used for the simulation feature in Appendix a.1.

4.4.1 REINFORCEMENT LEARNING AND THE BETA DISTRIBUTION

In the classical model from Eggenberger and Pólya [57], an urn is initialized at time $t = 0$ with $\alpha \geq 1$ blue balls and $\beta \geq 1$ red balls. Then, at each discrete time step, a ball is drawn from the urn and replaced with $m \geq 1$ additional balls of the same color. It is widespread that the proportion of blue balls converges in distribution to a beta distribution $\mathcal{B}(\alpha/m, \beta/m)$. For a proof of the result, see for instance Mahmoud [111].

For any two reals $a, b > 0$, the beta distribution $\mathcal{B}(a, b)$ has a density function

$$p(x, a, b) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)} \mathbb{1}_{\{x \in [0,1]\}} \tag{4.23}$$

where $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and Γ is the Gamma function defined by

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt \tag{4.24}$$

for any $z > 0$.

Although we consider a system of interacting urns rather than a single urn, the beta distribution stands as a strong candidate for the limit distribution.

Conjecture 72. *The distribution of $\bar{z}$ follows a beta distribution $\mathcal{B}(a,b)$ for some $a, b, > 0$ which depend only on α and $\mathcal{N}$.*

A more refined intuition that would support Conjecture 72 relies on Hill, Lane, Sudderth, et al. [82]. In this seminal paper, authors show that in a general class of urn models, if a $\{0,1\}$-valued process of the outcomes of draws is *exchangeable*, then this process is a sequence of i.i.d. Bernoulli draws, deterministic, or generated by a Polya urn.

Definition 73. A finite sequence $X_1, \ldots, X_n$ of r.v.s is said to be *exchangeable* if its joint probability is invariant under any finite permutation. An infinite sequence $X_1, X_2, \ldots$ is exchangeable is its finite subsequences $X_1, \ldots, X_n$ are exchangeable for any $n \in \mathbb{N}$.

Exchangeability is known to be connected to i.i.d. r.v.s since De Finetti [51], where it is proved that if an infinite sequence of $\{0,1\}$-valued r.v.s is exchangeable, then the joint distribution of any subsequence is a mixture of independent Bernoulli distributions. It is also a well-known result that, given a Polya urn scheme, the random sequence defined by the indicator variable of draws being of a fixed color is an exchangeable sequence. What Hill, Lane, Sudderth, et al. [82] achieves is to show that in a class of generalized Polya urn processes, any exchangeable $\{0,1\}$-valued random sequence based on the outcome of draws must either be deterministic, a sequence of i.i.d. Bernoulli draws or generated by a Polya urn.

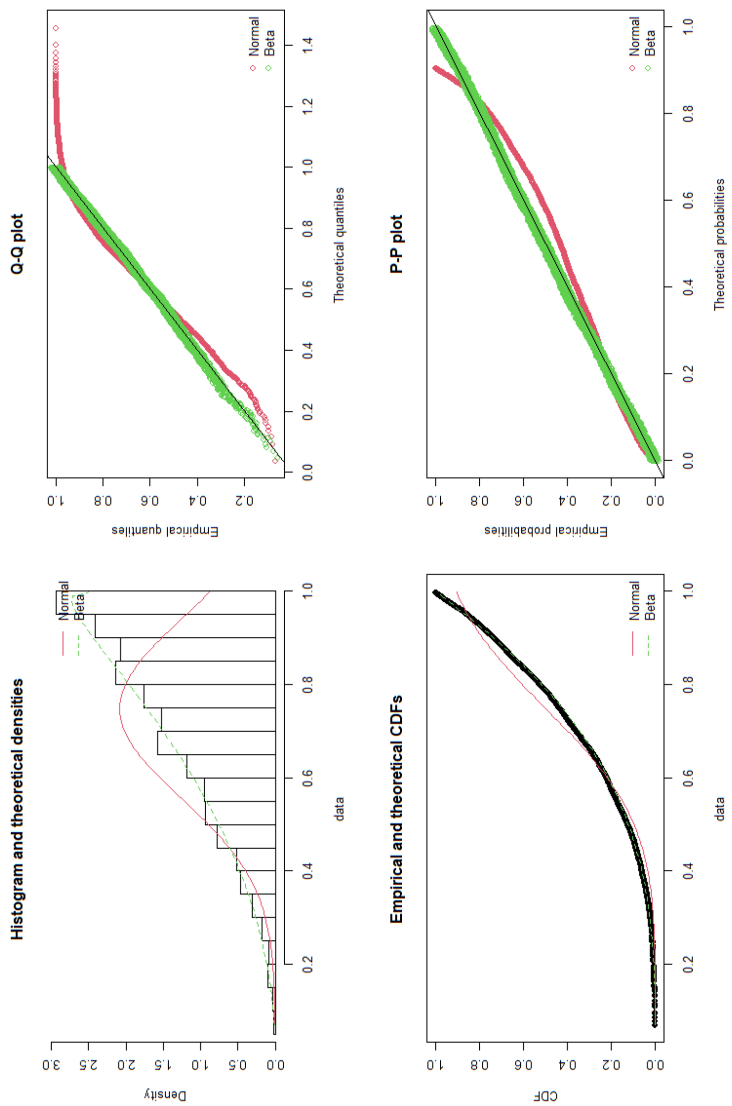


Figure 4.2: Fitness measures for beta and normal distributions on a star graph with $\alpha = 0.75$, $n = 5000$ observations.

Our framework strongly differs from those considered in Hill, Lane, Sudderth, et al. [82], both in its setup – as we consider a system of interacting urns instead of a single urn – and its properties as sequences of draws $x_1, x_2, \ldots$ are usually not exchangeable. Yet, since Theorem 68 ensures that proportions converge to a limit random value, the belief process obeys a looser definition of exchangeability.

Definition 74. A sequence $(X_1, X_2, \dots)$ is said to be asymptotically exchangeable if there exists an exchangeable sequence $(Y_1, Y_2, \dots)$ such that

$$(X_{k+1}, X_{k+2}, \dots) \xrightarrow[k\to\infty]{\mathcal{D}} (Y_1, Y_2, \dots) \tag{4.25}$$

In our model, the sequence of draws $(x^1, x^2, \dots)$ is clearly asymptotically exchangeable since proportions converge to a common fixed value. This implies the following proposition:

Proposition 75. *For any $i \in N$, the sequence of draws $x_i^1, x_i^2, \dots$ is asymptotically exchangeable.*

This result follows directly from the almost-sure convergence of the vector of proportions to a stable vector with identical entries. This implies that the sequence of draws converges to a sequence of i.i.d. Bernoulli variables with equal parameter, hence for each urn i we have asymptotic exchangeability.

To our knowledge, there is no result connecting immediately exchangeability in the limit to Polya urns and the beta distribution. Yet, we believe that there is a potential for an analytical proof of Conjecture 72 when combined by the local properties of the communication protocol in the limit. The exploration of this idea is still ongoing. In the meantime, to support our statement, we relied on large-scale simulations.

We collected the values of the limit belief for star, k-regular and complete graphs of fixed size $N = 100$ and fixed values of α. As all the simulations were run independently, for any given graph structure and value of α, the set of values of the limit beliefs is an i.i.d. sample. A beta distribution fitting was computed by maximum likelihood estimation. Other distributions were fitted in order to assess goodness-of-fit using usual criteria. Fig. 4.2 compares the fitted distributions assuming respectively a normal distribution and a beta distribution. The graph displays empirical and theoretical densities, quantile-quantile plots, cumulative distribution functions and probability plots. The beta distribution clearly appears as well fitted to the sample. Additional plots feature in the appendix for different network structures and values of α. In all the aforementioned cases, fitness measures yielded similar results, where the beta distribution clearly appears as more adapted to describe the data.

4.4.2 ESTIMATION OF THE PARAMETERS

Assuming the limit belief does follow a beta distribution, we are able to estimate its parameters using maximum likelihood estimation for various networks and values of α. Our first conjecture concerns the average of the limit distribution. All the simulations we ran conducted to a strong belief in that its value is α. In other terms, the expected proportion of red balls in the limit is equal to the expected number of misinformed agents *ex-ante*.

Conjecture 76. *For any $\alpha \in [0,1]$ and any $i \in N$, $\mathbb{E}_\alpha[\bar{z}] = \alpha$.*

While this stands as a conjecture in the general case, it can be proved for regular networks:

Theorem 77. *For any natural $1 \leq r \leq n$, assume $\mathcal{N}$ is a r-regular network. Then for any $\alpha \in [0,1]$ and any $i \in N$, $\mathbb{E}_\alpha[\bar{z}] = \alpha$.*

Proof. Let $M^t = \sum_{i \in N} z_i^t$.

$$M_t - \mathbb{E}\left[M_{t+1}|\mathcal{F}_t\right] = \mathbb{E}\left[\sum_{i=1}^{N} z_i^{t+1} - z_i^t|\mathcal{F}_t\right] \tag{4.26}$$

$$= \sum_{i=1}^{N} \frac{-d_i z_i^t + \sum_{j \in \mathcal{N}(i)} z_j^t}{d_i(t+1)+1} \tag{4.27}$$

$$= \sum_{i=1}^{N} z_i^t \left[\sum_{j \in \mathcal{N}(i)} \frac{1}{d_j(t+1)+1} - \frac{1}{d_i(t+1)+1}\right] \tag{4.28}$$

If $\mathcal{N}$ is r-regular, $d_i = r$ for every $i \in N$, hence

$$M_t - \mathbb{E}\left[M_{t+1}|\mathcal{F}_t\right] = \sum_{i=1}^{N} z_i^t \left[\sum_{j \in \mathcal{N}(i)} \frac{1}{r(t+1)+1} - \frac{1}{r(t+1)+1}\right] = 0 \tag{4.29}$$

This proves that the process (M^t) is a martingale. In particular, we have that $\mathbb{E}\left[M_t|z^0\right] = n\alpha$ for every t. Additionally, by Theorem 68,

$$\lim_{t\to\infty} M^t = \lim_{t\to\infty} \sum_{i\in N} z_i^t = \sum_{i\in N} \lim_{t\to\infty} z_i^t = n\bar{z} \tag{4.30}$$

hence we conclude that $\mathbb{E}_\alpha\left[\bar{z}\right] = \alpha$.

□

To support this conjecture in the case of non-regular graphs, we simulated the communication dynamics on a star graph of size $n = 50$ for increasing values of α. Table 4.2 provides estimates of the sample means.

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Av.	0.097	0.20	0.29	0.40	0.49	0.60	0.69	0.79	0.89

Table 4.2: Empirical mean in a star network of size $N = 50$ (7500 obs.).

According to Theorem 77 and Conjecture 76, the network structure does not influence the average limit belief, which remains close to the initial average belief on average. This leaves room for further study of the fluctuations of $\bar{z}$ around its expectation. In terms of applications of the model, Conjecture 76 implies that a risk-neutral disinformer who would try to maximize the spread of a false belief in the network is indifferent regarding who to influence. While this seems very intuitive in the case of regular graphs, as no player should have more influence than the other give they all share the same degree, this is quite surprising for non-regular graphs. In particular, it would imply than when trying to disinform a star network, on expectation targeting the central agent yields the same result than targeting any other node. Finally, note that while Theorem 77 and Conjecture 76 state that the network topology does not influence the average value of the limit belief, it is not true for higher order moments. On this point, simulations on different graph structures with identical number of agents and signal accuracy α yield different parameter estimates when fitting a beta distribution. The network structure also influences the speed of convergence, which happens at a faster rate in star networks than in circles for instance.

4.5 CONCLUSION

In this chapter, we have provided an alternative approach to modeling non-Bayesian opinion exchange using a stochastic framework. We assimilate the network of communication agents to an interacting urn system that evolves according to reinforcement dynamics. Based on stochastic approximation methods, we are able to prove the convergence of our model and recover the central properties of the DeGroot model with respect to the emergence of consensus.

Beyond the existence of this limit belief, the two models diverge in their results. In our case, we obtain a full-support random variable under the condition that the initial signals received by agents put a non-zero probability on the two possible states of the world. This answers our critiques regarding naive learning and its deterministic built-in convergence. By contrast, a random consensus leaves room to propagation phenomena that are observed in reality, as unlikely as they may be. Both Conjecture 72 and Conjecture 76 provide information on the likelihood of extreme events.

In their current state, Chapter 3 and Chapter 4 form a working paper which is still under progress. Several points still deserve additional efforts, in particular when it comes to characterizing the limit distribution and generalizing our results to finite state spaces and general urn initializations. While an analytical characterization of the limit belief is unlikely, we have strong hints that suggest this limit distribution is a beta distribution.

This confirms the intuition that, as a whole, the system acts as a global Polya urn. Further exploration of exchangeability properties and their connection to the beta distribution might help in supporting this intuition and provide some tools for a full proof of Conjecture 72. To our knowledge, no paper has been able to achieve such formal results on interacting urn systems.

A better understanding of the limit distribution would provide a better applicability of our results, in particular in designing a model of strategic disinformation with disinformants being parts of the network. That application was the initial motivation of the paper and remains its main objective. In that respect, proving Conjecture 76 would strengthen our results in an interesting direction. If it does stand for any graph topology, it would imply that the degree of an agent plays no role in the process of information spreading.

Part II

LEARNING IN DYNAMIC NONATOMIC ROUTING GAMES WITH INCOMPLETE INFORMATION

This part is based on the paper *Social Learning in Nonatomic Routing Games* in collaboration with Tristan Tomala (HEC Paris) and Marco Scarsini (LUISS). We consider a discrete-time nonatomic routing game with variable demand and uncertain costs. Given a routing network with single origin and destination, the cost function of each edge depends on some uncertain persistent state parameter. At every period, a random traffic demand is routed through the network according to a Wardrop equilibrium. The realized costs are publicly observed and the public Bayesian belief about the state parameter is updated. We say that there is *strong learning* when beliefs converge to the truth and *weak learning* when the equilibrium flow converges to the complete-information flow. We characterize the networks for which learning occurs. We prove that these networks have a series-parallel structure and provide a counterexample to show that learning may fail in non-series-parallel networks.

Chapter Chapter 5 introduces routing games and rational learning, the two central elements of the model. It proposes a brief overview of the standard models and major results. Chapter Chapter 6 details the model and presents the main results.

The paper has been presented at several conferences and published as an abstract in Chen et al. [40]. It is currently in *R&R* status at *Games and Economic Behavior*. Since we started this project, we greatly benefited from fruitful comments from participants in seminars and conferences where we got the chance to present it.

5

SOCIAL LEARNING AND CONGESTION GAMES

99 second hand smartphones are transported in a handcart to generate virtual traffic jam in Google Maps. Through this activity, it is possible to turn a green street red which has an impact in the physical world by navigating cars on another route to avoid being stuck in traffic.

Simon Weckert, 2020
(Google Maps Hack Performance, Berlin)

This chapter presents the problem of social learning in routing games studied in Chapter 6 through an overview of the major related threads of literature. Section 5.1 presents the non-theoretical motivations that lead to our main problem. Section 5.2 details the main elements in the informational framework at play in our model. Section 5.3 introduces the classes of congestion, potential and routing games. It exposes the main modeling elements and results from the algorithmic game theoretical literature. Section 5.4 presents the existing models on the topic of routing games with incomplete information and explains the main differences with what is done in Chapter 6. In Section 5.5, we discuss the key differences between continuous and finite player sets. Finally, Section 5.6 gives an overview of our main contributions on the topic.

5.1 NAVIGATION SYSTEMS AS LEARNING SYSTEMS

The market for navigation systems has settled as one of the biggest among the information and communication technologies, both in its number of users and its revenues. According to GlobalWebIndex data, 54% of cellphones used GoogleMaps in the year 2013, making it the most popular app, with nearly $3 billion in revenue for the year 2019. With numerous operators, the market for navigation systems is highly competitive, with the competition between providers being driven by their respective reliability as perceived by users. A majority of current navigation systems rely on real-time estimations to propose the most efficient path. To do so, they combine users' data streams to compute the level of congestion and corresponding travel times on each road in the routing network. This naturally raises the question of the reliability of their available data. In a notable performance, German artist Simon Weckert connected 99 phones to GoogleMaps and wandered in Berlin to simulate what the app recognized as a traffic jam. Through this performance, he intended to prove that not only is the data inflow of navigation systems manipulable, but that is also has consequences on other drivers, as path recommendations changed to avoid his artificial jam, rerouting actual drivers. This, of course, is an extreme instance of data manipulation and, although it deserves to be acknowledged as a potential threat to such systems, there are reasons to believe that, in general, GPS crowd-sourced data is trustworthy. But even in the sole presence of what could be called *truthful* drivers, basing recommendations on users data is highly problematic *per se* as all observations are equilibrium outcomes. Indeed, although this may be a sensible scheme when it comes to estimating real-time congestion on heavily used roads, it is unclear whether decentralized equilibrium dynamics provide a sufficient level of exploration of the whole network or not. Two mechanics are at play in this setup. First, the entangled relationship between equilibrium actions and information in-flow, characterized in Kremer, Mansour, and Perry [98] as the *feedback effect*: agents' choices generate new information, which will influence equilibrium behavior thereafter. Second, routing frameworks usually aim at modeling a vast number of agents that is, belong to the class of large games. Arbitrarily large player sets are approximated by continua of agents which are know to show myopic behavior in repeated settings. The conjunction of these two effects – endogeneous information acquisition and myopic behavior – may prevent agents from exploring their whole action sets and leave optimal paths unidentified.

There are two complementary approaches to this problem, which mostly depends on how the player set is modeled. On the one hand, drivers may be thought of as a set of agents sequentially arriving and making their decisions. In other terms, the navigation system, which performs the role of a principal, faces one driver at a time. It provides some information to the agent, which then chooses a path to go from her origin to her destination. It is then possible for the platform to reverse-engineer equilibrium behavior with respect to the offered information and determine the optimal information provision in order to achieve some social optimum. Of course, this requires the navigation system to have a sufficient knowledge of the routes, hence it has to incentivize some exploration of the routing network first. To do so, the platform has to provide – from time to time – some information on the unknown roads in order to push users to explore them and provide better information to later agents. This generates an exploration/exploitation trade-off in a sequential information design approach, which was studied in Kremer, Mansour, and Perry [98].

On the other hand, instead of considering a sequence of single players acting independently, one may consider instead a sequence of player sets. If player sets are continuous, hence all agents are negligible, the platform is limited to a non-strategic role, only acting as an information aggregator and public broadcaster. The central mechanics there becomes the interplay between the sequence of equilibria and the dynamics of beliefs, *à la* sequential learning. This is where players' myopia hinders accurate information aggregation, as there may exist some paths which would not be take along the sequence of equilibria, while being optimal if their cost functions were known. In other terms, even if information is publicly and completely broadcast – i.e. all drivers have perfect information about traffic – the amount of gathered information might not be socially efficient, as some routes may never be taken at any equilibrium of the game. This is the problem we tackle in Chapter 6.

Of course, there exists some middle-ground between the two models. One may consider a sequence of finite player sets for instance. They are not under the scope of our paper, but Section 5.5 provide some elements on their behavior.

5.2 RATIONAL LEARNING IN GAMES

5.2.1 Bayesian learning dynamics

The understanding of learning in the present part strongly differs from Part i. In the previous chapters, we referred to learning in the sense of belief exchange within a population that has no access to further information. Every signal was present *ex ante* and the main focus was on how they transit across the agent network, given some communication protocol. In this part, the information is obtained gradually through feedbacks controlled by action profiles. Those feedbacks may take many forms, from observing the choice of actions of other players to state-dependent payoffs. The central problem becomes whether the equilibrium dynamics lead players to learn efficiently, in the sense that the sequence of equilibria converges to the set of Nash equilibria of the game with full information.

The interest of game theorists in processes converging to Nash equilibria can be traced back to early works by Robinson [129] on the convergence of fictitious play. Most of these early models consider a non-strategic process or boundedly-rational players. The inclusion of a full-fledged Bayesian framework in the study of learning dynamics originated through two concurrent lines of work. The first line studies the concept of *rational learning* and transposes results in statistical learning into a game theoretical language. Introduced in Kalai and Lehrer [90], the phrasing *rational learning* refers to a framework where agents are assumed to be subjectively rational in the sense that prior to the game they hold private prior beliefs on their opponents payoffs, which they adapt over the course of the game *via* Bayes Rule. The model accounts for the fact that players know that their opponents are also engaged in the process of learning and operate in the long-term. The main result from Kalai and Lehrer [90] show that, for finite player and action sets, if agents' prior beliefs on their opponents' strategies are absolutely continuous with respect to the truth, then the sequence of Nash equilibria of the game with incomplete information converges to a point arbitrarily close to the set of Nash equilibria of the game with full information almost-surely. The absolute continuity condition, also referred to as *grain of truth* condition, follows from the seminal result from Blackwell and Dubins [28] on merging of stochastic processes. Importantly, authors in Kalai and Lehrer [90] distinguish two outcomes of the learning process: in the long-run, and under the conditions of their main result, as the sequence of equilibria converges to the set of equilibria of the full-information game, players may not learn their opponents strategies accurately. We carry this distinction throughout Chapter 6.

The second line of work which introduced Bayesian learning dynamics in repeated games originated with Banerjee [17] and is referred to as *sequential social learning*. These models consider short-lived agents sequentially facing a decision problem. Agents differ only in their initial private information. Prior to making their decision, they observe a subset of actions selected by past players. The core mechanics at play is the interaction of the two sources of information: private signals and observed plays. Banerjee [17] and Bikhchandani, Hirshleifer, and Welch [27] considered models where agents enter a market sequentially, update their beliefs by taking into consideration their private signals and the actions chosen by the previous agents, and make their optimal decisions accordingly. They show that social learning may fail, that is, it is possible that in equilibrium all agents choose a suboptimal action. Such phenomenon occurs when, after some time threshold, players ignore their private signals. This pattern is know as an *informational cascade*. Smith and Sørensen [144] showed that cascades are due to the hypothesis that the private signals are bounded, so, from some point on, no private signal can overcome the observations' strength. When signals are unbounded, social learning occurs with probability one. A very general version of this model was recently studied by Arieli and Mueller-Frank [9]. A thorough overview of the subject can be found in the survey Golub and Sadler [75]. From this part of the literature, we borrow the idea of a sequence of short-lived agent sets. In the general case, this is done at the cost of a strong reduction of the set of equilibria of the repeated game, as short-lived agents do not optimize intertemporally. In our case, though, there is no loss of generality. Details on this specific point feature in Section 5.5.

5.2.2 PARTIAL LEARNING IN GAMES

There is an important body of literature on partial learning and the interaction between equilibrium and beliefs dynamics. As exposed above, Kalai and Lehrer [91] defined rational learning for an infinitely repeated game where Bayesian players with heterogeneous beliefs maximize their expected utilities. They showed that, if agents' prior beliefs and the truth satisfy an "absolute continuity" condition, then the sequence of plays converges to an outcome arbitrarily close to a Nash equilibrium. Yet, in general, even if agents perfectly observe action profiles, errors in prior beliefs may persist over the course of the game. Fudenberg and Levine [63] introduced the concept of self-confirming equilibrium, where players' beliefs about other players' actions are required to be correct only on the equilibrium path. As agents play best reponses to their beliefs, these may differ substantially from the truth, as long as nothing contradicts the beliefs. Hence, self-confirming equilibria and Nash equilibria of a game need not be the same. Given the sequential nature of the repeated game, self-confirming equilibria may allow players to hold beliefs on other agents that may be inconsistent with rational behavior. Battigalli and Guaitoli [19] refined self-confirming equilibrium by proposing the concept of conjectural equilibrium for extensive form games of incomplete information without a common prior. At a conjectural equilibrium, players are assumed to behave rationally given the information they have about the game parameters. Rubinstein and Wolinsky [137] pushed this idea further through the concept of rationalizable conjectural equilibrium, which requires that agents' rationality be common knowledge. In this paper, we adopt a similar point of view. Namely, we study the steady states of social learning dynamics where players myopically best-respond to the information obtained by previous generations. We show that this information is correct for the edges of the network that have been explored along the equilibrium path, but may remain incorrect for other edges. Thus, some paths that would be used under full information may remain unused forever.

5.3 INTRODUCING ROUTING GAMES

Routing games form a popular family of models at the intersection of economics and computer science. Their popularity stems both from the practicality of the problem they pose – how does traffic route in a decentralized network – and the strength of their structure. As they form the main building element in our model, we propose a brief overview of the various existing models, their properties and the essential questions studied in the literature.

5.3.1 CONGESTION GAMES AND POTENTIAL GAMES

Routing games belong to the broader class of *congestion games*. In the most general sense, they consist in games where players aim at minimizing a cost generated by congestion externalities. In the formal sense, a congestion game if defined as a tuple $(N, \mathcal{E}, \mathcal{R}, \{c_e\}_{e \in \mathcal{E}})$ containing the following:

- N is the player set;
- $\mathcal{E}$ is a finite set of resources or facilities;
- $\mathcal{R} \in 2^{\mathcal{E}}$ is the set of pure strategies consisting in subsets of the resource set $\mathcal{E}$;
- For each $e \in \mathcal{E}$, c_e is a continuous, positive and increasing function mapping the mass of players using facility e to the cost they incur by doing so.

The congestion externalities are captured by the cost functions: the more players use some resource e, the higher the cost they will have to pay. Two different models are defined depending on the nature of the player set N:

- A congestion game is *atomic* if the set N is countable;
- A congestion game is *nonatomic* if the set $\mathcal{E}$ is a continuum.

In congestion games, players are characterized by the mass of resource, or unit of flow, they require. In nonatomic games it corresponds to the Lebesgue measure d of the player set N. In an atomic congestion game each player $i \in N$ is associated to her mass d_i expressed in units of flows. The total mass d is then the sum of individual masses d_i. Two important variants of atomic games exist depending on the nature of available actions:

- An atomic congestion game is said to be *non-splittable* if players only have access to pure strategies;
- An atomic congestion game is said to be *splittable* if players may play mixed strategies.

In other terms, if a game is splittable, players can split their mass on the different subsets of resources. For any pure strategy $r \in \mathcal{R}$, $y_r \in \mathbb{R}_+$ denotes the mass of players choosing pure strategy r. We refer to y as a *flow*. For each resource $e \in \mathcal{E}$, the *load* x_e of e is defined as

$$x_e := \sum_{r \ni e} y_r. \tag{5.1}$$

The symbols $\boldsymbol{x} = \{x_e\}_{e \in \mathcal{E}}$ and $\boldsymbol{y} = \{y_r\}_{r \in \mathcal{R}}$ denote the load vector and the flow vector, respectively. Notice that $\boldsymbol{x}$ is uniquely determined by $\boldsymbol{y}$, but not vice versa.

The cost of using edge e is $c_e(x_e)$, with an abuse of notation the cost of using path r is denoted by

$$c_r(\boldsymbol{y}) := \sum_{e \in r} c_e(x_e). \tag{5.2}$$

The cost vector $(c_e(x_e))_{e \in \mathcal{E}}$ induced by the load vector $\boldsymbol{x}$ is denoted by $\boldsymbol{c}(\boldsymbol{x})$. When considering atomic games, we denote $c^i(\boldsymbol{y})$ the cost paid by player $i \in N$ under the flow $\boldsymbol{y}$.

A flow vector $\boldsymbol{y}$ is *feasible* if it satisfies the demand, i.e.,

$$\sum_{r \in \mathcal{R}} y_r = d, \tag{5.3}$$

The set of *feasible flows* is denoted by $\mathcal{Y}$.

We now define the equilibrium concepts.

Definition 78. A feasible flow $\boldsymbol{y}^* \in \mathcal{Y}$ is a *Nash equilibrium* of an atomic game G if, for every $i \in N$ and action $r \in \mathcal{R}$, we have

$$c^i(\boldsymbol{y}^*) \leq c^i(r, \boldsymbol{y}^*_{-i}). \tag{5.4}$$

For nonatomic games, the standard solution concept is *Wardrop equilibria*, due to Wardrop [154].

Definition 79. A feasible flow $\boldsymbol{y}^* \in \mathcal{Y}$ is a *Wardrop equilibrium* (WE) of a nonatomic game G if, for all $r, r' \in \mathcal{R}$ with $y^*_r > 0$, we have

$$c_r(\boldsymbol{y}^*) \leq c_{r'}(\boldsymbol{y}^*). \tag{5.5}$$

An important consequence of Definition 79 is that, at equilibrium, all the paths receiving a non-zero mass of users have the same cost. Further, monotonicity of the cost functions imply uniqueness of this equilibrium cost.

Haurie and Marcotte [80] studies the connection between the two equilibrium concepts. The details of the asymptotic relation between atomic and nonatomic congestion games have been examined by Cominetti et al. [43].

Atomic congestion games were introduced by Rosenthal [131], who proved that any congestion game admits a potential function, whose local minimizers are the pure Nash equilibria of the game.

Definition 80. A game $G = (N, \mathcal{S}, u)$ is a *potential game* if there exists a function $\Psi : \mathcal{S} \to \mathbb{R}$ such that, for all $i \in N$, all $s_{-i} \in \mathcal{S}_{-i}$ and $s_i, s'_i \in \mathcal{S}_i$,

$$u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i}) = \Psi_i(s_i, s_{-i}) - \Psi_i(s'_i, s_{-i}) \tag{5.6}$$

For any atomic non-splittable game, the Rosenthal potential is defined as

$$\Psi(\boldsymbol{y}) = \sum_{e \in \mathcal{E}} \sum_{i=1}^{x_e} c_e(i) \tag{5.7}$$

If, instead, the game is atomic splittable or nonatomic, then the Rosenthal potential is

$$\Psi(\boldsymbol{y}) = \sum_{e \in \mathcal{E}} \int_{t=0}^{x_e} c_e(t) dt \tag{5.8}$$

The existence of a potential implies the following theorem, which holds both for atomic and nonatomic games with the corresponding equilibrium concept.

Theorem 81. *Every congestion game has at least one equilibrium flow.*

Note that for atomic non-splittable games with finitely many players, this immediately involves the convergence of the myopic best-response dynamics in finite time. An important example of congestion games which is featured in Rosenthal [131] is the Cournot oligopoly.

Monderer and Shapley [116] studied potential games and some of their generalizations and proved that for any potential game there exists a congestion game with the same potential function. Congestion games with a continuum of players and their relation to potential games were studied by Sandholm [140]. This isomorphic connection between the classes of congestion and potential games make the latter a natural framework for the study of online optimization algorithms. In particular, in nonatomic games, under strict monotonicity of the cost functions, the potential is a strictly convex function, which ensures uniqueness of the equilibrium flow. Coucheney, Gaujal, and Mertikopoulos [48] provides a regularized algorithm whose trajectories converge to Nash equilibria for both continuous and discrete-time potential games. Foster et al. [59] and Lykouris, Syrgkanis, and Tardos [110] provide fast-converging algorithms for regret minimization under *smoothness* assumption. Algorithmic game theory is a vivid field of research with a high frequency of publications and large range of applications. Giving an exhaustive account falls far beyond the scope of the present thesis. We refer the interested reader to the recent edition of Roughgarden [134].

5.3.2 routing networks and braess' paradox

Traffic routing games exist as a subclass of congestion games. They study the problem of decentralized agents who choose one path on a network from their origin to their destination, with the goal to minimize their cost, identified with the traveling time. This traveling time depends on the choice of all players, since the cost of an edge increases with the number of agents who use it. This family of models have applications in transportation, data traffic, patient flows in healthcare, etc. The major addition with respect to congestion games is a routing network structuring the set of pure strategies. From an abstract perspective, routing games add a combinatorial structure on the set of pure strategies. A *routing network* is an multigraph $\mathcal{N} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the vertex set and $\mathcal{E}$ is the edge set, endowed with an origin/destination pair $\mathsf{O}, \mathsf{D} \in \mathcal{V}$. In the vast majority of the literature, $\mathcal{N}$ is a directed network. Given $v, u \in \mathcal{V}$, a *path* from v to u is an ordered set of edges $e_1, \dots e_k$ such that the tail of e_1 is v, the head of e_k is u, and, for each $i \in \{1, \dots, k-1\}$, the head of e_i is the tail of e_{i+1}. The set $\mathcal{R}$ indicates the set of paths from O to D. To avoid trivialities, we assume that each edge is part of a path in $\mathcal{R}$. Each $e \in \mathcal{E}$ is endowed with a continuous increasing function $c_e \colon [0, \gamma_e) \to \mathbb{R}_+$ that represents the cost of using edge e as a function of its load. Every routing game being a congestion games, all the definitions and results aforementioned apply.

The earliest example of a traffic congestion game can be found in Pigou [126] through a simple example. There are two nodes, the origin O and the destination D, connected by two parallel directed edges. Costs are defined as in Fig. 5.1: the upper edge has a linear cost $c_1(x) = x$ while the lower edge has a constant cost $c_2(x) = 1$.

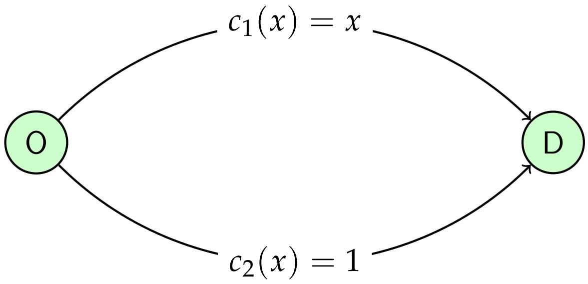


Figure 5.1: Pigou's Routing Network.

Assume there is a demand mass $d = 1$. Then, there is a unique flow $\boldsymbol{y}^*$ where all users transit through the upper road, with total cost at equilibrium $\boldsymbol{y}_1^* \dot{c}_1(\boldsymbol{y}^*) = 1$.

Some properties of this class were then studied by Beckmann, McGuire, and Winsten [20], in particular connections to linear programming and queueing theory.

Starting with Roughgarden and Tardos [135], an important body of work has been devoted to the study of equilibrium efficiency. In most applications of routing games, network performance is a crucial issue which is difficult to address by simple regulatory means. Roughgarden and Tardos [135] and followers worked on quantifying the degradation in performance due to pure decentralized behavior. Quantification is based on the ratio of total equilibrium and optimal costs. The worst-case value of such ratio is coined *price of anarchy* (P.o.A.) in Papadimitriou [122]. Its best-case counter-part is named *price of stability* (P.o.S.) in Anshelevich et al. [8]. In Pigou's example, it is immediate to see that the unique equilibrium flow $\boldsymbol{y}^*$ is not efficient. An equal splitting of the demand between the two paths would yield a total cost of $\left(\frac{1}{2}\right)^2 + \frac{1}{2} = \frac{3}{4}$. This value corresponds to the lowest achievable cost, hence in this example $\mathsf{PoA} = \mathsf{PoS} = \frac{4}{3}$.

More generally, Roughgarden and Tardos [135] show that the P.o.A. is bounded above by $\frac{3}{4}$ for affine cost functions. Roughgarden [132] shows that this inefficiency is not a property of the network in the sense that in a Pigou network, it is possible to obtain an unbounded P.o.A. using appropriate cost functions. Different bounds are computed for other classes of cost functions and we refer the reader to Roughgarden [134] for further information on the subject.

Proving that it is possible to achieve an arbitrarily high P.o.A. in a Pigou network is not proving that the network topology also has no impact on equilibrium efficiency. On the contrary, there is a strong connection between the routing network and the occurrence of a phenomenon known as the *Braess paradox*, after Braess [30]. We borrow the following definition from Milchtaich [114].

Definition 82. Let $G := (d, \mathcal{N}, \{c_e\}_{e \in \mathcal{E}})$ be a routing game. Braess paradox occurs in G if there exist an edge $e \in \mathcal{E}$ and two cost functions $\hat{c}$ and $\tilde{c}$ such that $\hat{c}(x) \leq \tilde{c}(x)$ for any load, and such that for any equilibrium flows $\hat{\boldsymbol{y}}$ when $c_e = \hat{c}$ and $\tilde{\boldsymbol{y}}$ when $c_e = \tilde{c}$ with equal demand, the total equilibrium costs $c(\hat{\boldsymbol{y}})$ and $c(\tilde{\boldsymbol{y}})$ satisfy $c(\hat{\boldsymbol{y}}) > c(\tilde{\boldsymbol{y}})$.

In other terms, Braess paradox occurs in a network $\mathcal{N}$ when raising the cost on an edge strictly lowers the total equilibrium cost. A simple instance of this phenomenon occurs in the following. Let $\mathcal{N}$ be the network depicted in Fig. 6.1.

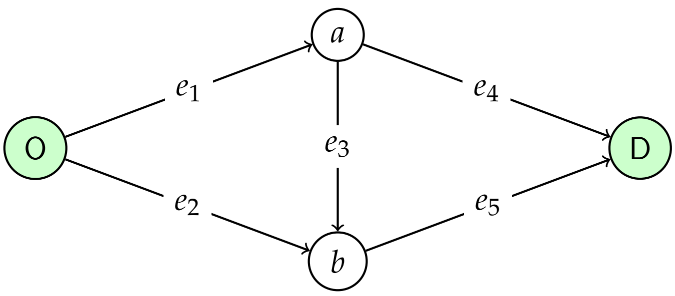


Figure 5.2: Braess Paradox.

Let c_1, c_2, c_3, c_4, c_5 be respectively the cost functions on edges e_1, e_2, e_3, e_4 and e_5 and assume for any $x \in \mathbb{R}$

$$c_1(x) = c_5(x) = x \tag{5.9}$$

$$c_2(x) = c_4(x) = 1 \tag{5.10}$$

$$\text{and } c_3(x) = 0. \tag{5.11}$$

Assume that the demand mass is 1. Then, there is a unique equilibrium flow with a mass 1 routing on the path $e_1 - e_3 - e_5$ and the total equilibrium cost is equal to 2. If instead we have $c_3(x) = \infty$ for any $x \in \mathbb{R}$, then there exists a unique equilibrium flow with a mass $\frac{1}{2}$ of users choosing the path $e_1 - e_4$ and a mass $\frac{1}{2}$ using the path $e_2 - e_5$. The total equilibrium cost becomes $\frac{3}{2}$ which is strictly lower: by removing edge e_3, the equilibrium cost improves for every player.

Braess paradox has been shown to emerge from the topology of the routing network. Holzman and Law-yone [84] and Milchtaich [114] have characterized respectively the directed and undirected network classes that are immune to this phenomenon in routing games with a single origin-destination pair. Chen, Diao, and Hu [39] extends both papers to networks with a finite number of origin-destination pairs. In those papers, authors show that the class of networks where Braess paradox does not happen corresponds to the class of series-parallel networks defined in Duffin [56] and Riordan and Shannon [127]. They consist in a class of graphs that can be constructed by an iterative process of series and parallel operations.

Definition 83. A network $\mathcal{N}$ is called *series-parallel* (S.P.) if it can be defined sequentially as follows:

(a) Either $\mathcal{N}$ has a single edge.

(b) Or $\mathcal{N}$ consists of two S.P. networks connected in series, by merging the destination of the first with the origin of the second.

(c) Or $\mathcal{N}$ consists of two S.P. networks connected in parallel, by merging the origin of the first with the origin of the second and the destination of the first with the destination of the second.

Fig. 5.3 provides examples of S.P. and non-S.P. networks.

Intuitively, S.P. aim at preventing the embedding of subnetworks similar to Fig. 5.2. We refer the reader to Duffin [56],

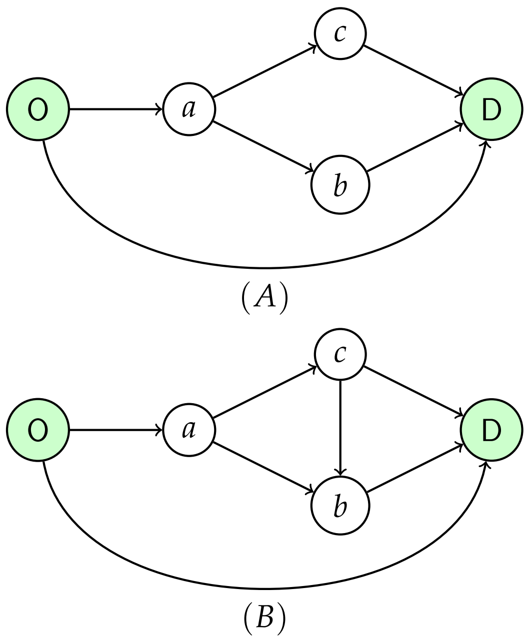


Figure 5.3: Network (A) is S.P.. Network (B) is not S.P., due to the edge from node c to node b.

Holzman and Law-yone [84], Milchtaich [114], and Riordan and Shannon [127] for further information on S.P. networks.

5.4 ROUTING GAMES WITH INCOMPLETE INFORMATION

In most of the existing literature on routing games, costs are assumed to be known, but in reality they are affected by unpredictable circumstances: car accidents, downed servers and so on. Uncertainty relative to these cost functions may strongly impact the equilibrium behavior, attracting agents away from potentially optimal actions. Thus the analysis of routing games of incomplete information is an important object of study.

As being a playground for economists and computer scientists, much has been said on the efficiency of algorithmic learning in routing games. This is less true when it comes to more traditional economic approaches to learning, although there exists some body of literature. As is shown in Acemoglu et al. [4] as well as Chapter 6, informational issues in routing games are strongly connected to the topological issues posed by Braess paradox. We review papers that are closely related to our work. Due to the diversity of these papers and the relative relevance of their results regarding our own, we will only describe their main focal points briefly.

5.4.1 Informational issues and equilibrium efficiency

Informational issues in static and dynamic routing games have been considered by several authors, some of them dealing with atomic games. For instance, Gairing, Monien, and Tiemann [66] considered atomic routing games with incomplete information where the type of a player is her traffic, which is private information. They proved existence of Bayesian pure equilibria and they studied their complexity. Then, they studied properties of completely mixed Nash equilibria and finally they provided bounds for the price of anarchy. Gairing [65] studied atomic congestion games where each player can be of two types, rational or malicious. He proved that pure Bayesian equilibria may fail to exist and studied the price of malice. Ashlagi, Monderer, and Tennenholtz [10] studied symmetric congestion games where the number of active players is unknown and there is no known prior distribution over the number of active players. Berenbrink and Schulte [24] considered evolutionary stable strategies for Bayesian routing games with parallel links. Fotakis et al. [60] studied how the inefficiency of equilibria in congestion games is affected by what they call social ignorance, that is, the lack of information about the presence of other players. Scarsini and Tomala [142] studied repeated versions of routing games. Syrgkanis [149] and Roughgarden [133] proved that known bounds for the price of anarchy in smooth games extend to their incomplete information version when players' private preferences are drawn independently. Roughgarden [133] showed that the extension does not hold for correlated preferences. Cominetti et al. [42] studied the behavior of the price of anarchy in atomic congestion games where each player i takes part in the game with some probability p_i and the participations are independent. Gaitonde and Tardos [67] and Gaitonde and Tardos [68] examined discrete-time queueing models where routers compete for servers and learn using strategies that satisfy a no-regret condition. The key element of their model is the explicit consideration of carryover effect from one period to the other.

5.4.2 ROUTING GAMES WITH RANDOM DEMAND

Our model adopts parts of the sequential social learning framework, mainly in the form of a sequence of demands of random size. The following references deal with uncertainty in the traffic demand. Wang, Doan, and Chen [153] examined nonatomic routing games with random demand and studied the dependence of the price of anarchy on the variability of the demand and on the network structure. Correa, Hoeksma, and Schröder [47] studied a class of nonatomic routing games where different sets of players take part in the game with some probability; they considered the Bayes Nash equilibrium of these games and showed that the bounds for price of anarchy do not deteriorate with respect to the complete information version of the game. Bhaskar et al. [26] studied non-atomic routing games where cost functions are unknown. Their goal was to determine edge tolls in order to achieve specific equilibria. They showed that this can be done under mild conditions through the use of an oracle computing equilibrium flows given a set of tolls. In particular, they computed tight complexity bounds for series-parallel routing networks.

5.4.3 INFORMATIONAL BRAESS PARADOX

The closest thread of research to our work studies the connection between equilibrium inefficiency, the topology of the network and the information structure.

Acemoglu et al. [4] dealt with nonatomic routing games where different types of agents have different information sets and each agent can only use paths in her own information set. They considered non-oriented routing networks and defined the concepts of information-constrained Wardrop equilibrium and of *informational Braess paradox*, that is, a situation where, if agents get more information, they experience a higher cost in equilibrium. They showed that such a paradox cannot happen if and only if the network is *series of linearly independent* (SLI). The class of SLI networks is a subclass of S.P. networks that prevents the embedding of subnetworks in the form of the network depicted in Fig. 5.4.

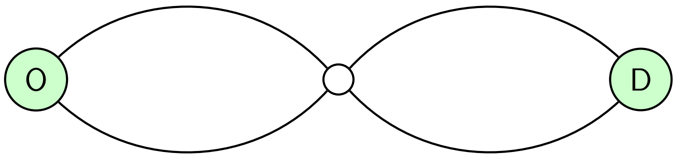


Figure 5.4: An undirected, non-SLI S.P. network.

Wu, Liu, and Amin [160] studied a routing game where agents subscribe to one of two traffic information systems and characterized equilibria under two information structures whose difference is the assumption of the common prior in one and not in the other. In the routing game studied by Wu, Amin, and Ozdaglar [159] the population of users is divided into groups and each group subscribes to one specific traffic information system. The cost functions on each edge are state dependent and each traffic information system sends a noisy signal only to its subscribers. The solution concept that they adopted is the Bayesian Wardrop equilibrium. They studied the sensitivity of the equilibria with respect to changes in the size of the groups. The paper by Wu and Amin [158] is the closest in spirit to our own. They analyze nonatomic routing games with unknown costs where Bayesian public beliefs are updated over time. At equilibrium, only the used edges provide some information about the realized costs. The difference between their model and ours is that they consider constant demand and noisy costs with Gaussian noise, possibly correlated across different edges. While close in flavor, the two different sets of assumptions produce different learning outcomes. In our model, costs are deterministic functions of a random state of the world which is fixed *ex-ante*. This implies that learning revolves around sampling the cost functions at sufficiently many load levels in order to accurately distinguish the different possible states. This allow us to find conditions on distribution of demand and network topology that ensure learning. The differences between the two models are detailed in Section 6.5 where we provide instances of routing games where noisy costs cannot induce learning while a random demand does.

5.5 ON FOLK AND ANTIFOLK THEOREMS

One of the reasons why sequential social learning models consider sequences of short-lived agents is that this enforces myopic behavior. An agent living only one period has no incentives to play some dominated strategy in the hope to provide information to future generations. This shrinks the set of equilibria of the repeated game, as they reduce to a sequence of Nash equilibria of the stage games. In our paper, we adopt a similar modeling hypothesis as we consider a sequence of demand masses of short-lived agents. The reason we require this is to allow for randomness on the demand size, which would make little sense if we considered the same agent set at every period. Regarding the nature of equilibria in the dynamic game, this assumption is not as strong as it may seem. We even claim that it is done at very little loss of generality.

5.5.1 MYOPIC BEHAVIOR OF CONTINUOUS PLAYER SETS

While it may seem counter intuitive that long-lived agents may only behave myopically in a repeated game, it is mostly a consequence of the nonatomic nature of the games we consider. When dealing with continua of players, folk theorems usually cease to apply and leave room for their opposite counterpart, *anti-folk theorems*. This line of results originated with Dubey and Kaneko [55], Green [77], and Kaneko [95]. In these papers, authors expose the limits of folk theorems with respect to the observability of actions. They show that the set of equilibria of a repeated game shrinks when players have limited monitoring of their opponents. Green [77] restricts his analysis to the implementability of trigger strategies in a context of public monitoring and anonymous actions. He shows that in a repeated game where players are not informed of their opponents' actions, any equilibrium sustained by a trigger strategy is a sequence of ε- equilibria of its stage games. Dubey and Kaneko [55] shows that in discounted games with public monitoring, if actions are anonymous, i.e. individual deviators cannot be identified, then the set of Nash equilibria of the repeated game is equal to the set of sequences of Nash equilibria of the stage games. A particular case where this result applies is games with continuous player sets: if payoffs only depend on the aggregate distribution of actions, then a zero-measure deviation is not detectable hence players are anonymous and the anti-folk theorem applies. With a finite number of players, conditions are more demanding as whenever a player deviates, there is a change in the aggregate choice of actions. Sabourian [138] provides proofs on the tightness of the anonymity assumption for games with large but finite player sets and extends the results from Green [77] to general strategy sets. In Al-Najjar and Smorodinsky [6], authors further extend these results to games where payoffs are not necessarily anonymous. More recently, Pai, Roth, and Ullman [121] extended the analysis to games with private monitoring.

In our model, the player sets are continuous and cost functions are continuous and strictly increasing functions of the mass of agents using the edges of the network. Building on these theorems, it is equivalent to assume that realized costs are unobserved on edges receiving a zero-measure of agents and to assume that the demand consists in a sequence of short-lived agent sets. In Chapter 6, due to this consequential non-strategic behavior, we will avoid referring to players and player sets and mostly consider them as a demand inflow.

5.5.2 ATOMIC ROUTING GAMES WITH UNKNOWN COSTS

As we stressed the strong differences between atomic and nonatomic models, it is natural to wonder what would have happened of our model in the case of a finite player set. Fortunately, in the case of agents with non-splittable demand, an answer has already been provided by Wiseman [157]. In this paper, Wiseman aims at establishing a folk theorem result for games where payoffs depend on an unknown state parameter. There is finitely many players, states and actions. There is a common prior belief μ on the state and the history of the game is public. Agents share a common discount rate. His main result is as follows: fix an $\varepsilon > 0$ and a vector $(v(\theta_1), \ldots, v(\theta_k))$of payoffs in the interior of the set of feasible and individually rational payoffs in each state and assume that the common prior belief μ puts a strictly positive weight on each state, then there exists a discount threshold $\bar{\delta}(\mu) < 1$ such that for all $\delta > \bar{\delta}(\mu)$ there is a sequential equilibrium such that when the realized state is θ, the expected payoff vector is ε-close to $v(\theta)$ and with probability at least $1-\varepsilon$, the belief on θ is ε-close to 1 in finite time.

As our model considers a finite routing network $\mathcal{N}$, a finite state space Θ and a public history, were we to consider an atomic agent set with non-splittable demands, we would fall under the scope of this partial folk theorem. We would jointly have a characterization of payoffs that can be sustained in the repeated game as well as learning in finite time.

There remains the case of an atomic player set with splittable demand. We do not have an immediate answer to this question, which is non-trivial. We leave the door open for future exploration.

5.6 CONTRIBUTIONS OF THE CHAPTER

We consider a repeated symmetric nonatomic routing game (NRG) where the cost functions of each edge depend on the load of the edge and on an unknown state parameter that is invariant over time. The set of states is finite and endowed with a common prior. At each period of time, a short-lived generation of users with a given total demand plays the game and realizes a Wardrop equilibrium with respect to the expected costs on edges: each path that receives positive load has the least expected cost. For every used edge, its load and the corresponding realized cost become public information for the following generations. There is perfect recall, so each generation knows the entire past history of the game and updates its beliefs in a Bayesian way. The sequence of different generations' demands is assumed to be random, i.i.d..

We consider two concepts of social learning: under strong learning, players eventually learn the true state of the world; under weak learning they learn to play the game as if the true state of the world was known. We show that weak learning is a strictly weaker concept than strong learning and that the conditions to achieve either of them depend on the topology of the network and on the support of the random demand. Our main theorem proves that weak learning occurs if the routing network is series-parallel and both the cost functions and the support of the demand are unbounded. Further, we show that strong learning is achieved under the same prerequisites and the additional condition that the demand has full support over $\mathbb{R}_+$. The intuition behind this result is the following: when the demand is stochastic, equilibrium flows vary. This generates observations of the cost functions for different values of loads. Based on results from Cominetti, Dose, and Scarsini [44] on the variation of equilibrium flows with respect to the demand, we prove that in a series-parallel network, as the demand goes high, all edges are used in equilibrium and equilibrium loads are unbounded. This implies that the cost functions will be observed at levels which allow distinguishing between the cost-relevant states with probability one. Finally, we prove that the condition on the network topology is necessary: for typical networks that do not satisfy it, we show that there exists an assignment of cost functions and capacities such that weak learning fails for any distribution of the demand.

The intuition is that a network which is not series-parallel contains a Wheatstone sub-network, and the topology of Wheatstone network is such that some edges are not used in equilibrium when the demand is high.

6

SOCIAL LEARNING IN NONATOMIC ROUTING GAMES

Y en la antena de la radio flotaba locamente la bandera con la cruz roja, y se corría a ochenta kilómetros por hora hacia las luces que crecían poco a poco, sin que ya se supiera bien por qué tanto apuro, por qué esa carrera en la noche entre autos desconocidos donde nadie sabía nada de los otros, donde todo el mundo miraba fijamente hacia adelante, exclusivamente hacia adelante.

Julio Cortázar, *La Autopista del Sur*
Todos los fuegos el fuego (1966)

The main question raised in Chapter 5 may be expressed as follows: if a routing game where cost functions are not known is repeated over time and beliefs of players are updated taking into account observations of previous players, will the costs functions be eventually learned? We identified two opposite effects arising naturally: on one hand, agents aim at minimizing the costs they incur immediately. On the other hand, socially efficient behavior requires agents to explore the routing network in order to learn the actual costs. We consider generations of short-lived players who play the game only once and are being replaced every period by a new set of players. This artifact from sequential models of social learning eases the exposition while being down at a relatively small loss of generality. Some amount of social learning may be achieved as players of one generation update their beliefs based on the behavior of the previous generations. Thus, selfish behavior may provide public information for the next generations of players. One challenge is to analyze the amount of such public information provision. If the game parameters are stationary, given that each generation has no incentive to be forward looking, potentially informative behavior may be off equilibrium path and thus social learning may fail. Yet, when there is variability in the circumstances in which the game is played, current equilibrium behavior may provide useful information to the subsequent generations of players.

The chapter is structured as follows. Section 6.1 presents the static model of routing games with unknown costs. Then Section 6.2 presents the dynamic version of the game and our definitions of weak and strong learning. In Section 6.3, our main results are stated along with examples that illustrate the importance of the conditions they require. In Section 6.4, we provide the complete proofs of our main theorems. Section 6.5 details the major differences between our model and its closest counterpart, Wu and Amin [158]. Finally, Section 6.6 concludes the chapter and Part ii.

6.1 ROUTING GAMES WITH UNKNOWN NETWORK STATE

6.1.1 CAPACITATED ROUTING GAMES

Our modeling is based on the baseline routing game detailed in Section 5.3. Let $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ be an oriented multigraph endowed with an origin/destination pair $\mathsf{O}, \mathsf{D} \in \mathcal{V}$. The set $\mathcal{R}$ indicates the set of paths from O to D.

We consider capacitated edges: every edge in the network allows for a given maximal mass of agents. This framing encompasses the model presented in Section 5.3 as capacities can be infinite.

Formally, each $e \in \mathcal{E}$ is endowed with a *capacity* $\gamma_e \in (0, +\infty]$ and with a continuous strictly increasing function $c_e \colon [0, \gamma_e) \to \mathbb{R}_+$ that represents the cost of using edge e as a function of its load. The traffic demand is denoted by d. The tuple $G := (d, \mathcal{N}, \{\gamma_e\}_{e\in\mathcal{E}}, \{c_e\}_{e\in\mathcal{E}})$ defines a *nonatomic routing game*.

A *cut* $\mathcal{C}$ of the network is a subset of $\mathcal{E}$ such that there does not exist a path from O to D the uses only edges in $\mathcal{E} \setminus \mathcal{C}$. The capacity of $\mathcal{C}$ is the sum of the capacities of its edges $\gamma_{\mathcal{C}} := \sum_{e\in\mathcal{C}} \gamma_e$. The *capacity* γ of the network $\mathcal{N}$ is the smallest capacity among all possible cuts; it corresponds to the maximum traffic that can flow from origin to destination [see 58]. Throughout the paper, we assume $d \in [0, \gamma)$, that is the demand satisfies the capacity constraints. If $\gamma = \infty$, then any positive demand can be satisfied.

For each path $r \in \mathcal{R}$, $y_r \in \mathbb{R}_+$ denotes the *flow* over path r. For each edge $e \in \mathcal{E}$, the *load* x_e of e, load vector $\boldsymbol{x} = \{x_e\}_{e\in\mathcal{E}}$ and the flow vector $\boldsymbol{y} = \{y_r\}_{r\in\mathcal{R}}$ are defined as in Section 5.3.

A flow vector $\boldsymbol{y}$ is *feasible* if it satisfies the demand and obeys the capacity constraints, i.e.,

$$\sum_{r\in\mathcal{R}} y_r = d, \tag{6.1}$$

$$x_e < \gamma_e, \quad \text{for all } e \in \mathcal{E}. \tag{6.2}$$

The set of *feasible flows* is denoted by $\mathcal{Y}$. Notice that $\boldsymbol{x}$ is uniquely determined by $\boldsymbol{y}$, but not vice versa.

This model encompasses the classical case where there is no limit of capacity: $\gamma_e = +\infty$ for each e. It also covers M/M/1 queuing models where the cost (or waiting time) tends to infinity as the demand approaches the capacity. An alternative approach would be to extend the cost functions to the whole $\mathbb{R}_+$ and allow values in $[0, +\infty]$; this would have no impact on our results.

Since cost functions are strictly increasing, there exists a unique load $\boldsymbol{x}^*$ associated to any equilibrium flow $\boldsymbol{y}^*$ (this unique load is the minimizer of a strictly convex potential function).

6.1.2 Bayesian Wardrop equilibrium

We now introduce uncertainty, represented by a finite probability space $(\Theta, 2^{\Theta}, \mu)$ called the *state space*. A *nonatomic routing game with unknown network state* is a tuple $G_{\mu} := (G, (\Theta, 2^{\Theta}, \mu))$, where G is a nonatomic routing game as before and, for each $e \in \mathcal{E}$ and each $\theta \in \Theta$, the cost function $x \mapsto c_e(x, \theta)$ is continuous and strictly increasing on $x \in [0, \gamma_e)$. To guarantee identifiability of the states, we assume that for every pair $\theta, \theta' \in \Theta$, there exists an edge $e \in \mathcal{E}$ such that $c_e(\,\cdot\,, \theta) \neq c_e(\,\cdot\,, \theta')$.

Given a prior distribution $\mu \in \Delta(\Theta)$, with a common abuse of notation, the expected costs are denoted by

$$c_e(x, \mu) := \int_{\Theta} c_e(x, \theta) \, \mathrm{d}\mu(\theta) \quad \text{and} \quad c_r(\boldsymbol{y}, \mu) := \int_{\Theta} c_r(\boldsymbol{y}, \theta) \, \mathrm{d}\mu(\theta). \tag{6.3}$$

We extend the concept of the WE to incomplete information by defining the *Bayes-Wardrop equilibrium* (BWE) as the Wardrop equilibrium flow of the game played with the expected costs. A similar definition appeared in Wu, Amin, and Ozdaglar [159].

Definition 84. A flow $\boldsymbol{y}^* \in \mathcal{Y}$ is a BWE of G_{μ} if, for all $r, r' \in \mathcal{R}$ with $y_r^* > 0$, we have

$$c_r(\boldsymbol{y}^*, \mu) \leq c_{r'}(\boldsymbol{y}^*, \mu). \tag{6.4}$$

Note that strict monotonicity of the cost functions implies uniqueness of the equilibrium load. The BWE load of the game G_{μ} with demand d is denoted by $\boldsymbol{x}^*(d, \mu)$.

6.2 DYNAMIC ROUTING GAMES WITH UNKNOWN STATE

6.2.1 SEQUENCES OF RANDOM DEMANDS

We now consider a discrete-time model of social learning where a routing game with unknown network state is played over time and with a population that changes at every period. The goal is to find conditions under which social learning is achieved, that is, each generation learns from the behavior of the previous generations and the public beliefs about the state of nature converge to the true value.

If the traffic demand were constant over time, then the same BWE would end up being played every period. As a consequence, learning does not occur whenever some equilibrium paths of the complete-information games are not used along the sequence of equilibria. Therefore, to achieve learning, we will assume that demands are given by a sequence $(D^t)_{t\in\mathbb{N}}$ of i.i.d. nonnegative random variables with common marginal distribution, denoted by F. The symbol D denotes a generic element of the sequence and $\operatorname{supp}(D)$ denotes its support. We assume independence between the demands and the state of nature. Therefore, by Kolmogorov extension theorem, F and μ induce a unique product measure $\mathbb{P}$ on the measurable product space $\big([0,\gamma)^\infty \times \Theta, \mathcal{B}([0,\gamma)^\infty) \otimes 2^\Theta\big)$.

At every period t, a demand D^t is realized and observed, the WE is played, the equilibrium load profile $\boldsymbol{x}^{*t}$ and the equilibrium costs $\boldsymbol{c}(\boldsymbol{x}^*,\theta) = (c_e(x^{*t},\theta))_{e\in\mathcal{E}}$ are observed. Therefore, for every $t = 1,2,\dots,$ the history at period t is

$$\begin{aligned} h^t &:= \Big(D^1, \boldsymbol{x}^{*1}(D^1), \boldsymbol{c}\Big(\boldsymbol{x}^{*1}(D^1),\Theta\Big),\dots,D^{t-1},\boldsymbol{x}^{*t-1}(D^{t-1}),\boldsymbol{c}\Big(\boldsymbol{x}^{*t-1}(D^{t-1}),\Theta\Big),D^t\Big) \\ &= \Big(\widetilde{h}^{t-1}, D^t\Big), \end{aligned} \tag{6.5}$$

where Θ is the random state.

The distribution μ on the state space is updated according to Bayes rule and μ^t denotes the posterior distribution $\mu(\,\cdot \mid h^t)$, that is,

$$\mu^t(\theta) := \mathbb{P}\big(\Theta = \theta \mid h^t\big). \tag{6.6}$$

The pair $(G, \mathbb{P})$ defines a *dynamic nonatomic routing game with unknown network state*, which will be denoted by Γ.

The sequence of posterior beliefs is a bounded martingale. Thus, by the martingale convergence theorem, there exists a random variable μ^∞ such that $\mu^t \to \mu^\infty$ almost surely. Moreover, since the set of states is finite and costs are deterministic functions of states and loads, there exists a random time τ such that almost surely, $\mu^t = \mu^\infty, \forall t \geq \tau$. Indeed in every period, either $\mu^t = \mu^{t+1}$ or the support of μ^{t+1} is a strict subset of the support of μ^t.

6.2.2 DEFINING LEARNING

As mentioned before, we will find conditions under which social learning is achieved. We now define two concepts of learning. To this end, the Dirac measure on θ will be denoted by δ_θ.

Definition 85. Consider a dynamic nonatomic routing game with unknown network state Γ. We say that:

(a) *strong learning* is achieved if

$$\mu^\infty = \delta_\Theta \quad \mathbb{P}\text{-a.s.}. \tag{6.7}$$

(b) *weak learning* is achieved if

$$\boldsymbol{x}^*(\,\cdot\,, \mu^\infty) = \boldsymbol{x}^*(\,\cdot\,, \delta_\Theta) \quad \mathbb{P}\text{-a.s.}. \tag{6.8}$$

The idea of Definition 85 is the following. If strong learning is achieved, asymptotically the true state of the world is discovered. Actually, given the fact that the state space Θ is finite, strong learning implies that there exists a random time τ such that $\mu^t = \delta_\Theta$ for all $t \geq \tau$. Under weak learning the true state is not necessarily discovered, but, asymptotically, in equilibrium, traffic is routed as if the state were known. This distinction follows Kalai and Lehrer [92], who studied the convergence of beliefs in repeated games and showed that players may learn to predict their opponents' actions even though they do not learn their payoff matrices. Similarly, in our model, players may not identify the true state of the world, yet they may learn to play optimally conditional on this state. It is not difficult to see that strong learning implies weak learning, but the converse implication is false. We will show this with the following example.

6.2.3 STRONG LEARNING IMPLIES WEAK LEARNING

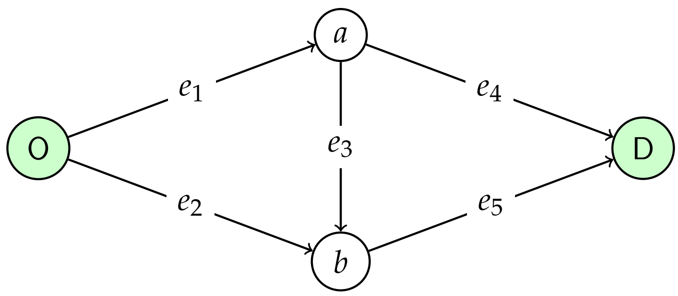


Figure 6.1: Wheatstone network.

Example 86. Consider the network in Fig. 6.1 where each edge has infinite capacity and

$$c_1(x,\theta) = c_5(x,\theta) = x, \tag{6.9}$$
$$c_2(x,\theta) = c_4(x,\theta) = 1 + \varepsilon x, \tag{6.10}$$
$$c_3(x,\theta) = \begin{cases} \varepsilon x & \text{for } \theta = \theta_G, \\ 10 + \varepsilon x & \text{for } \theta = \theta_B \end{cases} \tag{6.11}$$

with $\mu(\theta_B) = \mu(\theta_G) = 1/2$ and $\varepsilon < 1$. Let the demand D^t have a distribution with support $[20, \infty)$.

In this setup, for every value of the demand in the support, edge e_3 is not used. To see this, consider the paths

$$r_1 = (e_1, e_4), \quad r_2 = (e_2, e_5), \quad r_3 = (e_1, e_3, e_5), \tag{6.12}$$

and let y_1, y_2 and y_3 be their respective equilibrium flows under μ. Observe that for any value of D^t, $y_1 = y_2$ by symmetry. Then, given a realization d of D^1, there is a positive flow on r_3 if and only if $c_{r_3}(0,\mu) \leq c_{r_1}(d/2,\theta)$, that is,

$$d+5 \leq \frac{d}{2}(1+\varepsilon)+1, \quad \text{i.e.,} \quad \frac{d/2+4}{d/2} \leq \varepsilon,$$

which is impossible. Thus, no demand is routed on r_3, edge e_3 remains unexplored, and the state is not identified at time 1. At time 2 the situations is the same, so, by induction, $\mu^t = \mu$ for every $t \in \mathbb{N}$. On the other hand, observe that even under full information, e_3 would not be used given that $D^t \geq 20$. This shows that weak learning is trivially achieved, but strong learning is not.

6.3 SOCIAL LEARNING IN THE DYNAMIC GAME

6.3.1 MAIN RESULTS

Our convergence result requires assumptions on the cost functions, on the sequence of demands, and on the structure of the network. As is shown later, these assumptions are necessary. The next theorem provides conditions for weak and strong learning.

Theorem 87. *Let Γ be a dynamic nonatomic routing game with unknown network state such that, the network $\mathcal{N}$ is S.P. and, for each edge $e \in \mathcal{E}$ and each $\theta \in \Theta$, we have $\lim_{x \to \gamma_e} c_e(x,\theta) = +\infty$.*

(a) *If $\operatorname{supp}(D) = [0,\gamma)$, then strong learning occurs.*[1]

(b) *If for every $\varepsilon > 0$, there exists $d \in (\gamma - \varepsilon, \gamma)$ such that $d \in \operatorname{supp}(D)$, then weak learning occurs.*

The following theorem shows that the assumption that the network is S.P. is necessary in the sense that, if it does not hold, it is possible to construct an assignment of cost functions and capacities that satisfies the other hypotheses of Theorem 87 and for which weak learning fails for any distribution of the demand.

1 By definition, the support of a random variable is a closed set. Here it is closed relative to the space of feasible demands $[0,\gamma)$.

Theorem 88. *If the network $\mathcal{N}$ is not S.P., then there exist capacities and uncertain unbounded cost functions such that weak learning fails for every distribution of the demand.*

6.3.2 LEARNING FAILURE

Learning may fail if any of the assumptions of Theorem 87 does not hold. We present below a series of examples that show why learning may fail in those cases.

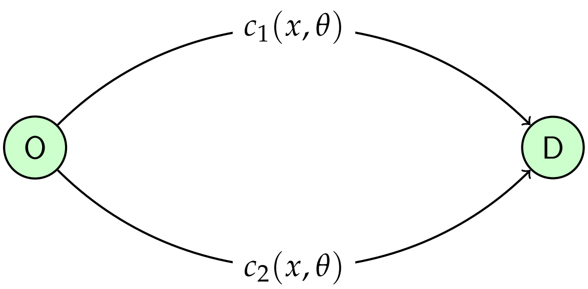


Figure 6.2: Parallel edge network.

Example 89 (Bounded costs). Consider the network in Fig. 6.2 with infinite capacity on each edge and

$$c_1(x,\theta) = 1 - \mathrm{e}^{-x}, \qquad c_2(x,\theta) = \begin{cases} x & \text{for } \theta = \theta_{\mathsf{G}}, \\ x + 10 & \text{for } \theta = \theta_{\mathsf{B}}, \end{cases} \tag{6.13}$$

with $\mu(\theta_{\mathsf{B}}) = \mu(\theta_{\mathsf{G}}) = 1/2$.

In the unique equilibrium of the game all the demand uses edge e_1 at any period t, for any possible value of D^t. This is due to the fact that the cost $c_1(\cdot,\theta)$ is bounded above by 1 and $c_2(x,\mu) = x + 5$. The lower path is dominated in expectation for every possible value of D^t; hence, no positive mass ever uses it at any equilibrium and the public belief remains equal to the prior. As a consequence, weak learning does not occur.

Example 90 (Bounded demand). Consider the network in Fig. 6.2 with infinite capacity on each edge and

$$c_1(x,\theta) = x, \qquad c_2(x,\theta) = \begin{cases} x & \text{for } \theta = \theta_G, \\ x + a & \text{for } \theta = \theta_B, \end{cases} \tag{6.14}$$

with $a > 0$ and $\mu(\theta_B) = \mu(\theta_G) = 1/2$.

The expected cost of edge e_2 is

$$c_2(x,\mu) = x + \frac{1}{2}a. \tag{6.15}$$

Therefore, if $D^t < a/2$ with probability 1, then edge e_2 is never used and even weak learning fails. In this example, learning fails because the lower path is dominated in expectation for low values of D^t. Under complete information, edge e_2 is used in equilibrium when the state is θ_G. When states are unknown, the presence of a fixed cost a in state θ_B deters the use of edge e_2 for low values of the demand. Hence in equilibrium the public belief remains equal to the prior.

Example 91 (Non-SP network). Consider the costs and network of Example 86 with infinite capacity on each edge. The expected cost of edge e_3 is

$$c_3(x,\mu) = \varepsilon x + 5. \tag{6.16}$$

If the demand D^t has an exponential distribution with parameter 1, then edge e_3 is never used, no matter the realization of D^t. Yet, it would be used for small values of the demand under θ_G. This shows that weak learning fails.

Since the network of this example is not series-parallel, we were able to find costs such that edge e_3 is used only under complete knowledge of state θ_G and for low values of the demand. Due to the fixed cost in state θ_B, no demand will use this edge, hence no learning occurs.

6.4 PROOFS OF THE MAIN RESULTS

In this section we provide the proofs of our main results.

6.4.1 PROOF OF THEOREM 87

Proof of Theorem 87. Given a prior μ^t, we define $L(\mu^t)$ the set of possible demands d such that, under the equilibrium load profile $\boldsymbol{x}^{*t}$, we have

$$\mu^t \neq \mu^{t+1} \tag{6.17}$$

if $D^t = d$. Notice that if $\mu^t \neq \mu^\infty$, the set $L(\mu^t)$ is nonempty.

Whenever $\mu^t \neq \delta_\Theta$, there exist θ_1, θ_2 such that

$$0 < \mu^t(\theta_1) < 1 \quad \text{and} \quad 0 < \mu^t(\theta_2) < 1. \tag{6.18}$$

Eq. (6.18) implies that there exists an edge $e \in \mathcal{E}$ such that, for the above states θ_1, θ_2,

$$c_e(\,\cdot\,, \theta_1) \neq c_e(\,\cdot\,, \theta_2). \tag{6.19}$$

Let $\bar{x}_e$ be such that

$$c_e(\bar{x}_e, \theta_1) \neq c_e(\bar{x}_e, \theta_2). \tag{6.20}$$

We want to prove that there exists a value d of the demand for which the equilibrium load on edge e is $\bar{x}_e$. To do this we use the following lemmata:

Lemma 92 (Cominetti, Dose, and Scarsini [45, Proposition 3.12])**.** *Let $\mathcal{N}$ be a finite S.P. network. Then there exists an equilibrium load profile function $d \mapsto \boldsymbol{x}^*(d)$ whose components $x_e^*(d)$ are nondecreasing functions of the demand d.*

The equilibrium edge costs are continuous in the demand [see, e.g., 45, Proposition 3.1]. Unboundedness, continuity and monotonicity of the equilibrium edge costs imply the following lemma.

Lemma 93. *In a nonatomic routing game played on a SP network, for every $e \in \mathcal{E}$, the equilibrium load map x_e^* is unbounded.*

Proof of Lemma 93. By Lemma 92 and strict monotonicity of the cost functions, the equilibrium load profile x_e^* is weakly increasing. Since cost functions are unbounded, continuous and monotonic, equilibrium costs are unbounded as the demand tends to infinity. Therefore, for large enough demand, all routes are used and have the same equilibrium cost, which is also unbounded. It follows that equilibrium flows on routes are unbounded. Therefore equilibrium loads on edges are unbounded as well. □

Continuity of $c_e(\,\cdot\,,\theta)$ for every $\theta \in \Theta$ and Eq. (6.20) imply that $c_e(x_e,\theta_1) \neq c_e(x_e,\theta_2)$ for every x_e in some neighborhood $\mathcal{I}$ of $\bar{x}_e$. Moreover, there exists a demand interval $\mathcal{D}$ such that, for every $d^t \in \mathcal{D}$, the expected equilibrium cost of edge e is $c_e(x_e^*,\mu^t)$, with $x_e^* \in \mathcal{I}$. Therefore, when a demand $d^t \in \mathcal{D}$ occurs, learning is achieved because one of the two costs $c_e(x_e^*,\theta_1)$ or $c_e(x_e^*,\theta_2)$ is realized, so that either

$$\mu^t(\theta_1) = 0, \quad \text{or} \quad \mu^t(\theta_2) = 0. \tag{6.21}$$

(a) The assumption that $\operatorname{supp}(D) = [0,\gamma)$ implies that the event $D^t \in \mathcal{D}$ has positive probability. Therefore,

$$\mathbb{P}\big(D^t \in \mathcal{D}, \text{ for some } t \in \mathbb{N}\big) = 1, \tag{6.22}$$

which concludes the proof.

(b) If $\mu^\infty = \delta_\Theta$, then strong learning is achieved. This implies that weak learning is achieved. If $\mu^\infty \neq \delta_\Theta$, then there exist $\theta_1,\theta_2 \in \Theta$ and $\bar{t}$ such that

$$0 < \mu^t(\theta_1) < 1 \quad \text{and} \quad 0 < \mu^t(\theta_2) < 1, \quad \text{for all } t \geq \bar{t}. \tag{6.23}$$

From the previous arguments, for any such θ_1,θ_2, if a value d of the demand is such that, for $h^t = \left(\widetilde{h}^{t-1}, d\right)$, either

$$\mu^t(\theta_1) = 0 \quad \text{or} \quad \mu^t(\theta_2) = 0, \tag{6.24}$$

then $d \notin \text{supp}(D)$. This shows that there is no value in the support of D such that an edge with unknown cost is used. Therefore the split of the flow is the same as it would be under perfect information.

□

6.4.2 PROOF OF THEOREM 88

The following lemma will be used in the proof of Theorem 88.

Lemma 94 (Chen, Diao, and Hu [39, Theorem 3.5])**.** *If a network $\mathcal{N}$ is not S.P., then it contains an* O-D *paradox, i.e., a subgraph $\mathcal{G} = r_1 \cup \widetilde{r}_2 \cup \widetilde{r}_3$ such that*

(i) *r_1 is a path from* O *to* D *that meets in this order the vertices* O$, a, u, v, b,$ D*,*

(ii) *$\widetilde{r}_2$ is a path from a to v whose only vertices in common with r_1 are a and v,*

(iii) *$\widetilde{r}_3$ is a path from u to b whose only vertices in common with r_1 are u and b,*

(iv) *$\widetilde{r}_2$ and $\widetilde{r}_3$ have no common vertices.*

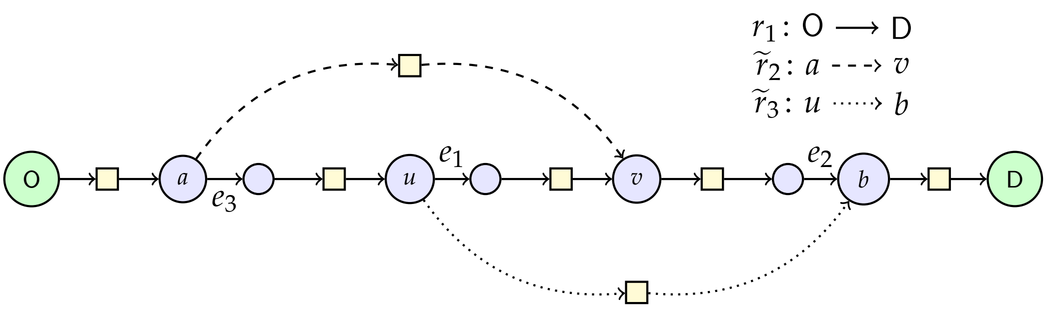


Figure 6.3: O-D paradox. The yellow squares represent finite sequences of nodes connected in series.

Proof of Theorem 88. Let $\mathcal{N}$ be a network that is not S.P.. Then, by Lemma 94, it contains an O-D paradox $\mathcal{N}'$ as in Fig. 6.3. The idea of the construction is to assign capacities and cost functions to edges in such a way that, the cost function c_5 is never learned, whatever the feasible demand.

Call r_2 the path that coincides with r_1 from O to a, with $\widetilde{r}_2$ from a to v, and with r_1 from v to D. Call r_3 the path that coincides with r_1 from O to u, with $\widetilde{r}_3$ from u to b, and with r_1 from b to D.

Let

$$\begin{aligned}
k_{\mathsf{O}a} &= \text{number of edges on } r_1 \text{ between O and } a,\\
k_{au} &= \text{number of edges on } r_1 \text{ between } a \text{ and } u,\\
k_{uv} &= \text{number of edges on } r_1 \text{ between } u \text{ and } v,\\
k_{vb} &= \text{number of edges on } r_1 \text{ between } v \text{ and } b,\\
k_{b\mathsf{D}} &= \text{number of edges on } r_1 \text{ between } b \text{ and D},\\
k_2 &= \text{number of edges on } \widetilde{r}_2,\\
k_3 &= \text{number of edges on } \widetilde{r}_3.
\end{aligned}$$

Let A and ε be two positive real numbers such that

$$\varepsilon < \frac{1}{3} \quad \text{and} \quad 3 < A. \tag{6.25}$$

All the edges that appear in Fig. 6.3 have infinite capacity. Let $\Theta = \{\theta_{\mathsf{G}}, \theta_{\mathsf{B}}\}$ with $\mu(\theta_{\mathsf{G}}) = \mu(\theta_{\mathsf{B}}) = 1/2$ and let the costs on the edges of the network be as follows:

$$c_2(x,\theta) = \left(A + \frac{\varepsilon}{k_{vb}}\right)x, \quad \text{for all } \theta \in \Theta \tag{6.26}$$

$$c_3(x,\theta) = \left(A + \frac{\varepsilon}{k_{au}}\right)x, \quad \text{for all } \theta \in \Theta \tag{6.27}$$

$$c_1(x,\theta) = \begin{cases} \left(A + \dfrac{\varepsilon}{k_{uv}}\right)x & \text{for } x \le 1, \quad \text{for all } \theta \in \Theta \\ \left(A + \dfrac{\varepsilon}{k_{uv}}\right) + \varepsilon^2(x-1) & \text{for } x > 1 \text{ and } \theta = \theta_{\mathsf{G}}, \\ \left(A + \dfrac{\varepsilon}{k_{uv}}\right) + \left(2A + \dfrac{2\varepsilon}{k_{uv}} - \varepsilon^2\right)(x-1) & \text{for } x > 1 \text{ and } \theta = \theta_{\mathsf{B}}. \end{cases} \tag{6.28}$$

For every other edge e appearing in Fig. 6.3 the cost functions are as follows:

$$
\begin{aligned}
c_e(x,\theta) &= \frac{\varepsilon}{k_{\mathrm{O}a}} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is between O and } a,\\
c_e(x,\theta) &= \frac{\varepsilon}{k_{au}} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is between } a \text{ and } u,\\
c_e(x,\theta) &= \frac{\varepsilon}{k_{uv}} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is between } u \text{ and } v,\\
c_e(x,\theta) &= \frac{\varepsilon}{k_{vb}} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is between } v \text{ and } b,\\
c_e(x,\theta) &= \frac{\varepsilon}{k_{b\mathrm{D}}} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is between } b \text{ and D},\\
c_e(x,\theta) &= \frac{\varepsilon}{k_2} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is on } \widetilde{r}_2,\\
c_e(x,\theta) &= \frac{\varepsilon}{k_3} x, \quad \text{for all } \theta \in \Theta, \quad \text{if } e \text{ is on } \widetilde{r}_3.
\end{aligned}
\tag{6.29}
$$

All the edges e that do not appear in Fig. 6.3 have a capacity

$$\gamma_e = \frac{\kappa}{|\mathcal{R}|}, \tag{6.30}$$

where $|\mathcal{R}|$ is the cardinality of $\mathcal{R}$ and

$$\kappa < \frac{1}{2}. \tag{6.31}$$

Moreover, for these edges

$$c_e(x_e) = \frac{1}{\gamma_e - x_e}, \quad \text{for } x_e \in [0, \gamma_e). \tag{6.32}$$

Eq. (6.30) implies that the load on edge e_1 coming from flows of paths different from r_1 is smaller than 1.

We prove now that, in equilibrium, the total load on edge e_1 is smaller than 1. Let $\boldsymbol{y}$ be a feasible flow vector and let y_1, y_2, y_3 be the corresponding flows on r_1, r_2, r_3, respectively. The expected costs given μ satisfy the following inequalities

$$
\begin{aligned}
c_{r_1}(\boldsymbol{y},\mu) &\geq \varepsilon(y_1+y_2+y_3)+(\varepsilon+A)(y_1+y_3)+(\varepsilon+A)y_1+(\varepsilon+A)(y_1+y_2) \\
c_{r_2}(\boldsymbol{y},\mu) &\leq \varepsilon(y_1+y_2+y_3+\kappa)+\varepsilon(y_2+1)+(\varepsilon+A)(y_1+y_2+\kappa) \\
c_{r_3}(\boldsymbol{y},\mu) &\leq \varepsilon(y_1+y_2+y_3+\kappa)+(\varepsilon+A)(y_1+y_3+\kappa)+\varepsilon(y_3+\kappa)
\end{aligned}
\tag{6.33}
$$

The path r_1 has a positive flow in equilibrium if and only if

$$c_{r_1}(\boldsymbol{y},\mu) \leq c_{r_2}(\boldsymbol{y},\mu) \tag{6.34}$$

and

$$c_{r_1}(\boldsymbol{y},\mu) \leq c_{r_3}(\boldsymbol{y},\mu). \tag{6.35}$$

The inequalities in Eqs. (6.33) and (6.34) imply

$$(\varepsilon+A)(2y_1+y_3) \leq \varepsilon(y_2+\kappa)+(2\varepsilon+A)\kappa. \tag{6.36}$$

Similarly, from Eqs. (6.33) and (6.35), we obtain

$$(\varepsilon+A)(2y_1+y_2) \leq \varepsilon(y_3+\kappa)+(2\varepsilon+A)\kappa. \tag{6.37}$$

Summing Eqs. (6.36) and (6.37), we obtain

$$(\varepsilon+A)(4y_1+y_3+y_2) \leq (2A+6\varepsilon)\kappa+\varepsilon(y_2+y_3), \tag{6.38}$$

that is,

$$y_1 \leq \frac{(2A+6\varepsilon)\kappa-A(y_3+y_2)}{4(\varepsilon+A)} \leq \frac{(A+3\varepsilon)\kappa}{2(\varepsilon+A)} \leq \frac{(A+3\varepsilon)\kappa}{2A} \leq \kappa. \tag{6.39}$$

Therefore, because of Eq. (6.31), the load on e_1 is at most $\kappa+\kappa \leq 1$. This implies that the cost function c_1 is not learned, for any value of the demand. On the other hand, if the true state were known to be θ_G, the equilibrium flow would be different than the one under uncertainty. This shows that weak learning is not achieved. □

6.5 RANDOM DEMAND OR NOISY OBSERVATIONS

In our model, the demand is random and realized costs are observed with certainty. Wu and Amin [158] also considered dynamic nonatomic routing games with unknown states where realized costs are observed. Unlike what we do in our paper, they assumed constant demand and noisy costs, that is, in their model realized costs depend on the unknown state and on multivariate normally distributed noises. The following example shows that these different sources of randomness lead to different learning outcomes.

Consider the network in Fig. 6.2 with infinite capacities and

$$c_1(x,\theta) = x, \qquad c_2(x,\theta) = \begin{cases} x & \text{for } \theta = \theta_{\mathsf{G}}, \\ x + a & \text{for } \theta = \theta_{\mathsf{B}}, \end{cases} \tag{6.40}$$

with $a > 0$ and $\mu(\theta_{\mathsf{B}}) = \mu(\theta_{\mathsf{G}}) = 1/2$. As shown before, the expected cost of edge e_2 is

$$c_2(x,\mu) = x + \frac{1}{2}a. \tag{6.41}$$

Assume that D^t has an exponential distribution with mean $a/4$. As $\operatorname{supp}(D^t) = \mathbb{R}_+$, we have

$$\mathbb{P}\big(D^t > a/2, \text{ for some } t \geq 1\big) = 1, \tag{6.42}$$

which implies that, almost surely, edge e_2 is used at some point and, in our model, strong learning occurs.

Consider now the same network game with the information model of Wu and Amin [158] with the demand $d^t = a/4$, i.e., equal to the expected demand of our model. Let the observed costs at period t be realizations of the random variables

$$\tilde{c}_e^t(x) := c_e(x) + \varepsilon^t, \tag{6.43}$$

for $e = 1, 2$, with ε^t normally distributed with mean 0 and variance σ^2. No matter the realization of the random cost, at any period t, the expected cost of edge 1 is lower than the expected cost of edge 2, so edge 2 is never used and weak learning fails.

In this example, there is learning only when high demand forces exploration of the edge with unknown cost. This shows that, despite their similarities, from the perspective of learning, the model with random demand and the one with noisy costs are different.

6.6 CONCLUSION

In this paper, we showed that the occurrence of social learning in nonatomic routing games is a complex phenomenon, which involves the topology of the routing network, properties of the cost functions and high volatility in the demand. Nonetheless, we are able to characterize necessary and sufficient conditions that ensure that either strong or weak learning will happen in the long-run. These conditions may appear as demanding. Requiring that the demand may reach values arbitrarily close to the capacity of the network may make sense when thinking about large cities whose networks are often saturated, but much less when considering routing networks at large scale. Having latencies approaching infinity may also seem extreme. And, last but not least, series-parallel networks are seldom observed in actual routing environments. But what we propose here is a theoretical artifact and, in the terms of a wise reviewer, "it is as unrealistic as the assumptions of a continuum of players and infinite-horizon games".

Chapter 5 and Chapter 6 form together an article, *Social Learning in Nonatomic Routing Games* which was revised and resubmitted to *Games and Economic Behavior* in September. The paper was previously accepted and published as an abstract in Chen et al. [40] and has been presented in talks and poster sessions at various conferences since its first version in the summer of 2019.

Over its two years of existence, our model has drastically evolved yet many things remain to be said on the subject. Although the example in Section 6.5 shows that models with noisy observation of realized costs and models with variable demand yield different properties, an interesting open question is how these two sources of randomness interact when combined. One particular case of interest is a model where the variance of the noise on a given edge is proportional to either its equilibrium load or the total demand. Another direction of extension is to provide bounds on the speed of learning for specific classes of cost functions. Little can be said on convergence speed without restrictions on the class cost functions. Indeed, two functions may differ by an arbitrarily small margin, or on a set of arbitrarily small probability. In line with the literature, restricting to a specific class of costs may yield computable bounds.

Part III

REVISION GAMES

This part is devoted to an ongoing project studying the impact of the frequency at which players choose and revise their actions on the level of commitment to environmental policy-making. We consider a 2-player stage-game where two countries decide on their levels of emissions. We assume that choosing to reduce emissions yields lower instantaneous returns than keeping the *status quo*, so that the latter is strictly dominant in the one-shot game, in the flavor of a prisoner's dilemma. The general game is as follows: there is a continuous and bounded time interval and, at times determined by a Poisson clock, the stage-game is played. A state of the world captures the overall frequencies at which each action has been chosen for the duration of the game. When time reaches the end of the interval, a terminal payoff is realized as a function of the state of the world. Our aim is to study the influence of the rate of the Poisson clock on the trade-off between flow and terminal payoffs. We aim at characterizing symmetric strategies that support cooperative behavior.

This project is still at an early stage of development and has neither been presented nor submitted yet.

7

REVISION GAMES WITH FLOW PAYOFFS

In the past, as we have sought new energy sources, we have too often damaged or despoiled our land. Actions to avoid such damage will probably aggravate our energy problems to some extent and may lead to higher prices. But all development and use of energy sources carry environmental risks, and we must find ways to minimize those risks while also providing adequate supplies of energy.
Richard Nixon, *State of the Union Message* to the Congress on Natural Resources and the Environment, February 15, 1973

In this chapter, we study the influence of time preferences on the existence of cooperative strategies in a model where two competing countries choose their level of emissions repeatedly. Their decisions influence the environmental state of the world that captures a cumulative effect of emissions throughout the duration of the game. Decisions are made at the ticks of a Poisson clock on a bounded time interval. Thus, the probability to act once more decreases as time advances, which is equivalent to players discounting their payoffs with a decreasing discount factor. When time reaches the end of the interval, players receive a terminal payoff that depends only on the state of the world. Our main interest is to study both the influence of this decreasing discount factor on the optimal emission trajectories and the existence of equilibria sustaining cooperative strategies.

This chapter is structured as follows. In Section 7.1, we present the applied motivation and the main theoretical problem at play in our model. Then in Section 7.2 we review the general framework of revision games as introduced originally in Kamada and Kandori [93]. In Section 7.3, we expose our model in detail. Section 7.4 presents our research agenda.

7.1 ENVIRONMENTAL DECISION MAKING AND TIME PREFERENCES

As its consequences are becoming more and more tangible, climate change has slowly risen as one of the top preoccupation of citizens worldwide. The scientific consensus on the human impact on the environment is facing a lesser opposition and populations as well as their leaders are internalizing the fact that drastic measures need to be taken to reduce the global level of emissions. But it has also become clear that state-level actions cannot achieve these goals on their own, and that international cooperation is necessary to establish clear goals, enforce them and reduce any source of misalignment between those goals and the countries taking part in those decisions.

7.1.1 THE LIMITED EFFICIENCY OF CLIMATE AGREEMENTS

International cooperation started with the early IPCC works 32 years ago, which led to the Rio agreements of 1992 and the organization of 27 United Nations Climate Change Conferences since 1995. This fast-paced collective decision system has led to two major agreements, the Kyoto agreements in 1997 and the Paris agreements in 2015. With the benefit of hindsight, it is now clear that the goals set in those agreements will not be achieved. While the United Nations Environment Program stated that instead of the 2°C target increase, average temperatures will likely increase by at least 3.2°C, with harsh economic consequences. Several explanations have been put forward to explain this inefficiency. First, it has been argued – see e.g. Calvo, Rubio, et al. [34] and Harstad [79] – that the non-binding aspect along with the incomplete nature of climate agreements disincentivizes participation. This creates a coalitional problem where agreements have to be designed by accommodating duration, punishments and each country's participation constraints, as in Battaglini and Harstad [18]. While providing important explanatory factors, this approach is not self-sufficient. Indeed, even with a binding structure and punishment options, climate coordinated action is still difficult to implement. On this point, one may consider the European action on climate: while having better collective performance within the European Union, as highlighted by the European Green Deal negotiations, the overall action remains insufficient to fulfill the the UNCCC objectives. Another complementary explanation focuses on technological investments and spillovers – see Acemoglu et al. [1] for instance – and studies models where states optimize a mix of clean and dirty technologies over time. Quite notably, these topics were absent from the Paris agreements. Such absence can be explained both by the limited control states have on an uncertain research outcome, which limits technological investment policies, and the strong competition on the field of green technologies which limits the existence of technological spillovers. It appears that long-term incentives for cooperation do not out-balance the strong drive for immediate returns, which is particularly salient since the short-term economic returns from renewable energies remain generally lower than fossil fuels. This is mostly due to costs per MWh remaining higher for renewable sources although public subsidies along with technological improvements reduced the gap over the recent years. Still, there remains a short-term opportunity cost in converting energy production from fossil to renewable sources. Meanwhile, the economic cost of long-term consequences keeps growing as we approach most of the IPCC deadlines. This intertemporal tradeoff justified a third approach to the problem, which revived a game theoretic debate in arguing the role played by discounting and time inconsistency.

7.1.2 TIME-PREFERENCES AND COOPERATION

Naturally, in taking decisions with strong long-term impact, the discount factor is of utmost importance; yet its nature and possible values are subject to strong debate (see for instance the Nordhaus-Stern debate following the publication of Stern [147]). Such discussions were already present in Samuelson [139] where exponential discounting was introduced. Samuelson expressed reservations as "*It is completely arbitrary to assume that the individual behaves so as to maximize an integral of (this) form*". To address these concerns, many authors have proposed alternative discounting schemes like hyperbolic discounting in Strotz [148] – backed by empirical evidence in Frederick, Loewenstein, and O'donoghue [61]. Uncertainty relative to the discounting factor and resulting controversies are addressed in Weitzman [155] and Gollier and Weitzman [73].

Ecological transition – e.g. conversion of energy production to renewable sources – requires to sustain instantaneous costs to the benefit of the global environment, assuming a constant technological level. Existence and sustainability of cooperative strategies strongly depend on the time preferences of the players. This effect has been proved in the lab in Jackson and Yariv [87]. Jackson and Yariv [88] further shows that aggregating heterogeneous time preferences yields time-inconsistency. Heterogeneous and/or non-constant discounting has been a difficult technical problem in the context of repeated games. Lehrer and Pauzner [105] shows that usual cooperative results like Folk Theorems do not hold when players have different discounting factors. For infinitely repeated games, Chen and Takahashi [38] extends results from Lehrer and Pauzner [105] and establish conditions to ensure cooperation. Similarly, Kochov and Song [97] shows that cooperative behavior emerges in an infinitely repeated Prisoner's dilemma with endogenous discounting. But in the context of environmental policy making, the strong presence of a time horizon, either fixed in agreements or as an endogenous deadline imposed by natural factors, prevents the application of such results. Managing non-standard time preferences in the presence of a time-horizon requires a stronger modeling structure. The seminal model of revision games introduced in Kamada and Kandori [93], later published as Kamada and Kandori [94], allows for both heterogeneous time-varying discounting and a fixed deadline. In such games, players act over a finite time interval. At random times determined by a – possibly individual – Poisson clock, they are offered the opportunity to revise their decision. As time approaches the end of the interval, the probability of revising a decision vanishes, which is equivalent to a discount factor converging to zero. This makes revision games an important framework in which agents discount their utilities at en endogenous rate that decreases as time increases. Kamada and Kandori [94] studies how it influences cooperative strategies in generalized prisoner's dilemma. This interplay between time preferences and the sustainability of cooperative strategies in environmental policy-making is the core question that this chapter proposes to explore. We try to answer this problem by modeling the trade-off between short-term revenue maximization and long-term control of a common state of the world.

7.1.3 CONTRIBUTION OF THE CHAPTER

In this part, our objective is two-fold. From a theoretical perspective, we intend to build a model of revision games with flow payoffs and a cumulative state to structure the trade-off faced by states involved in environmental transition. Two players act over a finite time interval and are offered at stochastic dates to revise their decision. This yields a flow of payoffs and determines a common state of the world that captures the share of time each player has played the non-cooperative action. This state is a proxy to model the impact of countries emissions over the course of the game. When time reaches the end of the interval, a terminal payoff is determined according to the state of the world.

This model remains an early-state project and most of our efforts have been spent in identifying the adequate model. Nevertheless, we provide a detailed agenda of the results we are working on. In practical terms, we first intend to prove the existence of an optimal symmetric strategy profile in the form of a threshold strategy: prior to a time threshold that negatively depends on the expected frequency of revision times, both players play the Nash equilibrium of the one-shot game, then shift to a cooperative action profile. In that respect, the higher the frequency at which players revise their strategies, the shortest the duration of the cooperative regime. This relationship between decision frequency and cooperative behavior is already observed in revision games without flow payments.

7.2 REVISION GAMES

Revision games is a class of games originally introduced in the mimeo Kamada and Kandori [93], later published in Kamada and Kandori [94]. They consist in games played over a finite time interval where players get to change – *revise* – their actions at random times. This is not the first model where agents are allowed to revise their choice of actions along the course of the game, which appears in Caruana and Einav [35] with switching costs for instance, nor is it the first that introduces a decreasing replay-probability as an endogenous discount factor, as was done in Vives [151] with continua of agents and infinite time horizon, but it features a comprehensive formal framework that best fits our modeling needs.

7.2.1 GENERAL FRAMEWORK

Formally, a revision game Γ as exposed in Kamada and Kandori [94] is described as follows. Consider a 2-player normal form game $G = (N, (\mathcal{S}_i)_i, (u_i)_i)$ called the *component game* with $N = \{1,2\}$. Time is continuous over an interval $[-T, 0]$ with $T > 0$. At time $t = -T$, each player chooses an action. Then, at times determined by a Poisson process, players can revise their choice of actions and observe their opponent's choice and revisions. When time reaches the *deadline*, i. e. $t = 0$, each player i receives her payoff $u_i(a)$ according to the action profile chosen at the last revision opportunity.

It is important to distinguish between two subsequent models: on the one hand, revision times may be determined by one single Poisson clock with rate λ, in which case the game is a *synchronous revision game*. On the other hand, each player i may receive revision opportunities by her own independent clock with rate λ_i in which case the game is an *asynchronous revision game*. Each model yields slightly different results, with the asynchronous case being more technically involving. In their paper, the authors focus on the synchronous case. The component game is assumed to be a generalized prisoner's dilemma: action spaces $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S}$ are convex subsets of $\mathbb{R}$ and payoffs $u_1(s, s') = u_2(s', s)$ satisfy the following assumptions:

- The component game G admits a unique pure symmetric N.e. profile (s^N, s^N) with payoff u^N and a unique best symmetric action s^* that maximizes $u(s) = u_1(s,s)$ for $s \in \mathcal{S}$;

- If $s^N < s^*$, $u(s)$ is strictly increasing for $s < s^*$ and symmetrically if $s^* < s^N$;

- $u_1(s_1, s_2)$ is continuous and $\max_{s_1} u_1(s_1, s_2)$ exists for all s_2 so the *gain from deviation* at a symmetric profile (s,s)

$$d(s) \equiv \max_{s_1} u_1(s_1, s) - u_1(s,s) \tag{7.1}$$

  is well-defined;

- If $s^N < s^*$, $d(s)$ is increasing on $[s^N, s^*]$ and non-decreasing for $s > s^*$ and symmetrically if $s^* < s^N$.

In this framework, *cooperation* refers to a choice of actions that provides a payoff superior to u^N to both players.

7.2.2 STRATEGIES AND OPTIMAL PLANS

Revision games are defined in continuous time but as revision times are countable, the games behaves admits a tree structure. At the n-th revision time t_n, the history of the game is given by

$$h_{t_n} = \left(s_T, t_1, s_{t_1}, \ldots, t_{n-1}, s_{t_{n-1}}, t_n\right). \tag{7.2}$$

The set of histories of size n is then $H_n = (\mathcal{S}_1 \times \mathcal{S}_2) \times ([0,T] \times (\mathcal{S}_1 \times \mathcal{S}_2))^n$. Then, a (behavior) strategy σ_i of player i in the game Γ is a measurable mapping from $\cup_{n \geq 0}(H_n \times [0,T])$ to the set $\Delta(\mathcal{S}_i)$ of probability distributions over $\mathcal{S}_i$. Note that in this definition, the dependence on time is with respect to the remaining time until the deadline.

In Kamada and Kandori [94], the analysis is restricted to *trigger strategies* where players coordinate on a symmetric cooperative action and punish deviations by playing the N.e. action

s^N. A symmetric trigger strategy is characterized by a *revision plan* $x : [0, T] \to \mathcal{S}$ that maps the amount of time remaining until the deadline to actions. At time $t = -T$, players start the game with action $x(T)$ and then play according to the revision plan. If a deviation occurs, then the Nash action s^N is played at all subsequent revision opportunities. Given a revision plan x, the expected payoff is

$$V(x) = u(x(T))e^{-\lambda T} + \int_0^T u(x(t))\lambda e^{-\lambda t} dt. \tag{7.3}$$

The first element in the right-hand side corresponds to the case where no revision opportunity occurs during the game. The second element is the expected payoff with respect to the distribution of the last revision time. Then, under the additional assumption that the gain from deviation is differentiable on $\left(s^N, s^*\right]$ with positive derivative if $a^N < a^*$ (and symmetrically), the prove their main result.

Theorem 95 (Kamada and Kandori [94])**.** *There exists a revision plan $\bar{x}$ that achieves the maximum trigger strategy payoff. It is continuous in t with $\bar{x}(0) = s^N$ and satisfies*

$$\frac{dx}{dt} = f(\bar{x}(t)) \tag{7.4}$$

if $\bar{x}(t) \neq a^*$ and $d'(\bar{x}(t)) \neq 0$, where

$$f(x) = \frac{\lambda(d(x) + u(x) - u^N)}{d'(x)} \tag{7.5}$$

Moreover, $\bar{x}(t) \in \left[s^N, s^*\right]$ for all $t \in [0, T]$ and $f(x) > 0$ on $\left(s^N, s^*\right)$.

They further characterize the existence of cooperative strategies under the assumption that the mapping f above is Lipschitz continuous on $\left[s^N + \varepsilon, s^*\right]$ for any $\varepsilon \in \left(0, s^* - s^N\right]$ when $s^N < s^*$ and symmetrically when $s^N > s^*$.

Theorem 96 (Kamada and Kandori [94]). *The optimal trigger strategy equilibrium plan $\bar{x}$ sustains cooperation, i. e. $\bar{x}(t) \in \left(s^N, s^*\right]$ for some t, if and only if*

$$\lim_{s \downarrow s^N} \int_s^{s^*} \frac{1}{f(x)} dx < \infty. \tag{7.6}$$

The condition in Eq. (7.6) requires that the solution to the differential equation in Theorem 95 travels from s^* to s^N in finite time. Thus, they characterize a necessary and sufficient condition for the existence of an optimal cooperative plan.

In the general case, a strategy σ_i is *Markovian* if it depends only on the remaining time and the opponent's action, i. e. σ_i is a measurable mapping from $[0, T] \times \mathcal{S}_{-i}$ to $\Delta(\mathcal{S}_i)$. A pair $\sigma = (\sigma_1, \sigma_2)$ of strategies induces a unique probability distribution $\mathbb{P}_\sigma$ over the set of histories of finite length. The relevant equilibrium concept is the following:

Definition 97. A strategy profile σ is a *Markov Perfect equilibrium* (M.P.e.) of Γ if both σ_1 and σ_2 are Markovian and σ is a S.P.e. of Γ.

In Kamada and Kandori [94], authors explain that the core element at play in Theorem 96 is the relative speed of convergence of gains to deviation and benefits from cooperation as the symmetric action profile converges to the N.e. profile of the component game. Namely, if the gains from deviation converge to zero faster than $u(s) - u^N$ does, then there exists an equilibrium which sustains cooperation. If the condition is not satisfied, then there is no cooperation in any equilibrium of the revision game.

A similar model was introduced in Calcagno and Lovo [33] to study preopening of financial markets. In Calcagno et al. [32], authors study the connection between players revision rates and equilibrium selection. Their results depend on the cooperative or adversarial nature of the game. Gensbittel et al. [71] studies asynchronous revision games with zero-sum payoffs.

7.2.3 STOCHASTIC REVISION GAMES

Closer to what we aim at modeling, Lovo and Tomala [109] introduces a model of stochastic revision games where players control a payoff-relevant state of the world through their actions and proves the existence of M.P.e.. Their model consists in the following elements:

- A finite set N of n players and a finite state space Θ;
- For each $i \in N$ and $\theta \in \Theta$, $\mathcal{S}_i(\theta)$ is the set of actions available to player i in state θ;
- The payoff function u maps states to $\mathbb{R}^n$;
- States evolve according to the transition probability $q : \Theta \times \prod_{i \in N} \mathcal{S}_i \to \Delta(\Theta)$;
- Each state θ is associated with a positive rate λ_θ;
- A time interval $[-T, 0]$.

The timing of the game is as follows: at time $t = -T$, an initial state θ_0 is drawn and the the game starts. A random time τ is drawn according to the exponential distribution $\mathcal{E}(\lambda_{\theta_0})$. If $\tau - T > 0$ then the initial action profile is fixed and the game terminates in state θ_0 with payoff $u(\theta_0)$. If $\tau - T \leq 0$, then a revision happens at time $\tau - T$. Each player i selects an action in $\mathcal{S}_i(\theta_0)$. With s the chosen action profile, a new state θ' is drawn with probability $q(\theta'|\theta_0, s)$. Then the subgame with remaining time $T - \tau$ and state θ' starts. Actions, states and revision times are perfectly observed.

The main results of Lovo and Tomala [109] prove the existence of M.P.e. with and without correlation. These results closely relate to the general scope of continuous-time stochastic games, studied in Levy [107], where general existence of equilibria in Markovian public correlated strategies is proved. Indeed, although payoffs are received only when time reaches the deadline, any revision game can be shown to be strategically equivalent, in the sense that strategy spaces and expected payoffs are equal, to a stochastic game as in Shapley [143]. The equivalent stochastic game is as follows: time is discrete and each period

t_n corresponds to the n-th revision opportunity of the revision game. A state variable θ_{t_n} is defined as the remaining time in the associated revision game. At each period, with chosen action profile (s_1, s_2), received payoffs are given by $u_i(s_1, s_2)e^{-\lambda\theta_{t_n}}$. The discount factor $e^{-\lambda\theta_{t_n}}$ can be seen as the probability that the game continues on the next period, a interpretation which was originally detailed in Shapley [143]. Importantly, Kamada and Kandori [94] points out the fact that the subclass of stochastic games equivalent to revision games does not belong to the set of stochastic games for which Folk theorems have been proved.

7.3 MAIN MODEL

In this section, we detail our modeling choices and the subsequent retained model. We begin by introducing the component game G.

7.3.1 COMPONENT GAME

The model we build is adapted from revision games. We define the normal-form component game G as follows:

- There is a set $N = \{1, 2\}$ of two players;
- $\mathcal{S}_1$ and $\mathcal{S}_2$ are the two finite action sets of players 1 and 2;
- A payoff function u maps pairs of actions to $\mathbb{R}^2$.

We denote by u_1 and u_2 respectively the payoffs of player 1 and 2. We assume that u is symmetric: for every pair $(s, s') \in s_1 \times \mathcal{S}_2$, $u_1(s, s') = u_2(s', s)$. Similar to Kamada and Kandori [94], we consider a version of the prisoner's dilemma and make the following assumptions:

1. $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S} = \{s\prime s^*\}$;
2. The game G admits a unique N.e. profile (s^N, s^N) with payoff $u_1(s^N, s^N) = u_2(s^N, s^N) = u^N$ and a unique best sym-

metric profile (s^*, s^*) that maximizes the sum $u_1(s, s') + u_2(s, s')$ for $s, s' \in \mathcal{S}$;

The action s^N represents the *status quo* action where the player choosing it maintains its emissions level, while the action s^* represents emissions reduction. In the one-shot game G, the status quo is then a dominant strategy. Modeling using variants of the prisoner's dilemma is frequently met in the environmental literature. Here, it captures both the fact that countries outputs are positively correlated with their individual level of emissions and the existence of some level of competition between the two players.

7.3.2 REVISION GAME AND CUMULATIVE STATE

We now define our revision game Γ. Time is continuous over the interval $[0, T]$, with $T > 0$. On this interval, the two countries play a synchronous revision game with rate $\lambda > 0$. At time $t = 0$, they both choose an initial action and then revise it at random times determined by the ticks of the Poisson clock. We denote by $(\tau_0, \tau_1, \dots)$ the sequence of revision times, with $\tau_0 = 0$. At any point in time, player i receives a payoff stream of $u_i(a)dt$ with a the action profile chosen at the last revision time.

We assume that revision times and subsequent choices are perfectly observable. At any point in time t, we write n_t the number of revision times occurring prior to t. The history of the game is given by

$$h_t = \left(s_{\tau_0}, \tau_1, s_{\tau_1}, \dots, \tau_{n_t}, s_{\tau_{n_t}}, t \right). \tag{7.7}$$

To capture the cumulative effect of the level of emissions, we introduce a state parameter that contains, through the game, the share of time players maintained the status quo. Namely, for any player i at any time t with history h_t, we define player i's cumulated emission level as

$$\theta_i(t) = \frac{1}{T-t}\left(\sum_{k=0}^{n_t-1}(\tau_{k+1}-\tau_k)\mathbb{1}_{s_{\tau_k,i}=s^N} + (t-\tau_{n_t})\mathbb{1}_{s_{\tau_{n_t},i}=s^N}\right) \tag{7.8}$$

with $s_{\tau_k,i}$ the action chosen by player i at the k-th revision time. We assume both countries emit in equal fashion, hence we define the overall cumulated emission level as

$$\theta_t = \theta_1(t) + \theta_2(t). \tag{7.9}$$

When time reaches 0, $\theta(h_T) \in [0, 2]$ corresponds to the total share of time countries kept the status quo over the course of the game. A terminal payoff $v(\theta(h_T))$ is received by both players, with $v : [0, 2] \to \mathbb{R}^-$. We assume that v is continuous and decreasing, with $v(0) = 0$. The terminal payoff v can be interpreted as the impact of emission trough the duration of the revision game on the economic value of the world at the deadline. This choice of modeling emissions *via* a cumulative state is close to the literature on games with frequency-dependent payoffs as introduced in Brenner and Witt [31]. These games correspond to stochastic games with a deterministic state modeled, in the case of two players, as a matrix whose entries converge to the empirical distribution of action profiles over the course of the game. Such games are further studied in Joosten, Brenner, and Witt [89] where authors prove a Folk theorem.

7.3.3 STRATEGIES AND EQUILIBRIA

A (behaviour) strategy for any player i is then a measurable mapping from the set H of histories to $\Delta(\mathcal{S}_i)$. A pair σ of strategies and a rate $\lambda > 0$ induce a unique probability distribution $\mathbb{P}_{\sigma,\lambda}$ on the set of histories of finite length, that uniquely extends to all histories by Kolmogorov extension. Then, for any profile σ of strategies at revision time $t = 0$, we can write player i expected payoff as

$$\begin{aligned} V_i(\sigma) = &\left(T u_i(\sigma) + \mathbb{E}_\sigma\left[v(\theta_T)\right]\right) e^{-\lambda T} \\ &+ \int_0^T \lambda e^{-\lambda(T-t)} \left(u_i(\sigma) + \mathbb{E}_\sigma\left[v(\theta_T)|h_t\right]\right) dt \end{aligned} \tag{7.10}$$

A pair of strategies $\sigma = (\sigma_i, \sigma_{-i})$ is a N.e. of the game Γ if, for every player i and every strategy σ_i',

$$V_i(\sigma_i, \sigma_{-i}) \geq V_i(\sigma_i', \sigma_{-i}) \tag{7.11}$$

We are particularly interested in the existence of M.P.e., that is, S.P.e. in Markovian strategies.

7.4 FUTURE DIRECTIONS

This research project is ongoing and although some results were obtained on previous versions of the model, they do not fit its current form. Based on previous work and the existing literature, we have established the following research agenda.

1. Prove the existence of M.P.e. in the general case: there is no *a priori* reason to believe M.P.e. would not exist under mild conditions. We believe that the reasoning displayed in Lovo and Tomala [109] transposes to our framework at a small cost.

2. Prove the existence of an optimal share of time τ^* that depends on λ, u and v during which players cooperate, which is equivalent to the existence of an optimal global emission level θ^*.

3. Prove the existence of an optimal symmetric strategy that works in the following way: at any revision time prior to

$T - \tau^*$, both players play the N.e. action s^N and switch actions to s^* at any revision time happening after. Such behavior was obtained in previous variations of the model. The intuition is as follows: if players could revise their strategies at any time, there could not be an equilibrium strategy such that players first cooperate then defect, as, granted the existence of aforementioned threshold τ^*, for any revision time happening in $[T - 2\tau^*, T - \tau^*]$, countries have a strict incentive to deviate to the N.e. action s^N unilaterally as long as, if they do so, their opponent would support the action s^*.

Establishing these results is a first step in obtaining comparative statics on τ^* with respect to λ. Preliminary work on simplified versions of the model provided results sustaining the idea that, for a fixed terminal payoff function v, the higher λ is, the shortest the amount of cooperation should be.

RÉSUMÉ EN LANGUE FRANÇAISE

8 RÉSUMÉ DE LA THÈSE

8.1 CADRE GÉNÉRAL DE LA THÈSE

8.1.1 INTRODUCTION

La théorie des jeux émerge en tant que champ de recherche indépendant dans les années 1940, à la suite de la publication de Von Neumann and Morgenstern [152], à l'interface des mathématiques et des sciences sociales. Depuis, le cadre de ses applications s'est étendu à de nombreux autres domaines parmi lesquels la science politique, l'informatique et la biologie. Près de 80 ans après la publication de son texte fondateur, la théorie des jeux se nourrit d'une variété de problèmes nouveaux et s'enrichit aussi bien en tant que discipline que comme outil. Un tel succès s'explique en ce qu'elle propose un cadre de pensée simple et transposable pour lequel toute situation d'interaction stratégique se modélise en un *jeu* à partir duquel il est possible de dériver une variété de résultats prédictifs. Ces prédictions se fondent sur le concept d'*équilibre*, un état où les comportements individuals s'équilibrent de sorte qu'aucun agent ne peut améliorer sa situation en changeant unilatéralement sa décision. Il est démontré que ces équilibres existent sous des conditions usuelles et, le cas échéant, ils offrent un point de référence pour les décideurs cherchant à appréhender l'issue la plus probable à une interaction entre des individus rationnels et stratégiques. La théorie des jeux est un objet dual, à la fois champ de recherche en-soi et ensemble de méthodes aux applications interdisciplinaires. Cette dualité est essentielle pour comprendre comment la théorie des jeux a évolué et intégré aussi bien les problèmes que les méthodes des domaines périphériques.

Cette thèse subsume les travaux effectués pendant nos études doctorales sur trois sujets différents, qui nous ont intéressé en ce qu'ils permettaient d'éclairer des problématiques réelles tout en ayat un attrait méthodologique. Les trois études qui en résultent sont essentiellement indépendantes, aussi bien dans les problèmes abordés que dans les méthodes employées pour les résoudre. Elles figurent donc en tant que trois parties distinctes, présentées dans l'ordre chronologique de leur conception.

8.1.2 DÉSINFORMATION ET APPRENTISSAGE DANS LES JEUX

La première partie de cette thèse trouve son origine dans les phénomènes de désinformation à large échelle observés au début des années 2010. La mise en oeuvre fréquente de fausses informations a créé une brèche dans la confiance du public envers les agences de presse, les institutions publiques mais aussi les acteurs du monde scientifique. En a résulté un état de doute et de méfiance aux conséquences économiques et institutionnelles certaines. Il convient alors de s'intéresser aux mécaniques de la désinformation pour comprendre comment fonctionnent les stratégies mises en oeuvre pour désinformer et, surtout, comment s'en prémunir.

Il existe une importante littérature étudiant l'information dans les jeux. Les asymétries et externalités d'information occupent une place essentielle dans le raisonnement économique, et la théorie des jeux a apporté des éclairages sur le sujet depuis la publication des travaux pionniers de R. Aumann et M. Maschler sur les jeux à information incomplète, réédités dans Aumann, Maschler, and Stearns [13]. L'obtention et l'exploitation de l'information dans les jeux est un phénomène généralement bien compris lorsqu'il se produit dans les limités du cadre bayésien. Ces modèles analysent les interactions entre des agents traitant l'information de façon rationnelle en ce sens qu'ils forment des croyances probabilistes sur les éléments incertains et mettent à jour ces croyances en appliquant la formule de Bayes. Ce courant de littérature a démontré que les extensions informationelles des modèles usuels – e.g. autoriser un agent à connaître un paramètre inconnu des autres agents – change fortement l'ensemble des équilibres d'un jeu. Pourtant, la modélisation bayésienne peine à analyser des grandes populations: elle repose sur une hypothèse forte sur les aptitudes calculatoires des agents et propose des solutions dont la complexité explose lorsque le nombre d'agents considérés augmente, sauf dans des structures de jeux précises comme dans Smith and Sørensen [144]. Pour ces raisons, les modèles bayésiens sont efficaces dès lors qu'il s'agit d'expliquer les comportements d'un faible nombre d'agents formant des croyances et exploitant de l'information, mais offrent une perspective limitée sur les phénomènes de masse comme la désinformation et la propagation de *fake news*.

Une approche alternative, dite *non-bayésienne* ou à *rationalité limitée*, comprend l'ensemble des modèles qui rejettent – partiellement ou entièrement – le cadre bayésien. Ces travaux reposent sur l'idée que les agents ont une capacité calculatoire et/ou d'observation limitée qui restreint leur traitement de l'information. Ce principe général abrite une variété de modèles et d'approches mais, en matière de formation d'opinion et d'échange de croyances à large échelle, un de ces modèles fait figure de référence canonique: le modèle de DeGroot. Ce modèle considère une population d'agents structurée en un réseau où, itérativement, chaque joueur remplace sa croyance par la moyenne de celles de ses voisins. Ce processus converge, sous des conditions raisonnables, vers un consensus qu'il est facile de déterminer à partir de la structure du réseau de communication et des croyances initiales. De nombreuses variations autour de ce modèle ont été proposées, la plupart reposant sur ce principe de moyenne locale. Bien que la simplicité des calculs qu'il propose soit un avantage comparatif certain du modèle de DeGroot par rapport aux modèles bayésiens, il n'en est pas moins sans reproche. En particulier, ce modèle ne permet pas l'existence de phénomènes extrêmes où une croyance marginale se propage à une part substantielle de la population. De notre opinion, ce fait est principalement dû à la nature du processus d'échange des croyances.

Pour ces raisons, nous proposons dans la première partie de cette thèse une approche alternative au modèle de DeGroot où les agents communiquent en effectuant des tirages selon leur croyance plutôt qu'en diffusant directement celle-ci. Ce modèle étend l'analyse de DeGroot et, dans le même temps, pose la question de la robustesse de ses prédictions.

8.1.3 RÉSEAUX DE ROUTAGE ET EXTERNALITÉS DE CONGESTION

La seconde partie de cette thèse est également liée à la littérature sur l'apprentissage dans les jeux mais avec une perspective inverse. Alors que la première partie se concentre sur l'échange et la diffusion d'opinions exogènes, cette deuxième partie étudie un modèle où l'information est obtenue de façon endogène, résultant des équilibres d'un jeu dynamique. Ce travail trouve son origine dans la littérature à l'interface entre la théorie des jeux et l'informatique. Ces deux champs présentent un intérêt fort pour les problématiques liées à l'apprentissage, mais les abordent de deux façons radicalement différentes. L'informatique est le domaine naturel d'étude des algorithmes d'apprentissage comme la méthode des poids multiplicatifs ou les équations de réplication. Ces approches non-comportementales de l'apprentissage ont prouvé leur efficacité dans plusieurs classes de jeux, parmi lesquelles celle des jeux de congestion est une des plus étudiées. Ces modèles de jeux étudient les situations où émergent des externalités de congestion, c'est-à-dire des problèmes d'allocation de ressources où plus le nombre d'agents utilisant une même ressource est élevé, plus le coût d'utilisation de cette ressource est élevé. Plusiers propriétés expliquent l'attrait particulier de cette classe de jeux: les équilibres en stratégies pures existent presque toujours, l'unicité de l'équilibre s'obtient à la simple hypothèse que les coûts sont strictement monotones, et il existe un lien fort entre ces jeux et l'optimisation convexe. Néanmoins, en dépit de ces propriétés structurelles fortes, et même si les algorithmes d'apprentissage y convergent en espérance, ces derniers sont parfois sujets à des comportements chaotiques. Il a été montré dans Chotibut et al. [41] que ce phénomène se produit dans des cas simples et non-dégénérés de jeux de routage.

Les jeux de routage forment une sous-classe des jeux de congestion. Ils modélisent les déplacements d'un flux d'agents sur un réseau de routage – e.g. un réseau routier ou un réseau informatique – chacun cherchant à minimiser son temps de trajet d'un noeud origine à un noeud destination. Chotibut et al. [41] démontre en particulier que dans un réseau comportant un seul noeud origine et un seul noeud destination, avec deux chemins parallèles les connectant, si les temps de trajet des deux chemins sont linéaires dans la demande, alors il existe un ensemble de paramètres pour lesquels la méthode des poids multiplicatifs se comporte de façon chaotique au sens de Li-Yorke. Si ce type de comportement peut se produire au sein du monde algorithmique, il est impossible qu'une croyance bayésienne se comporte de la sorte, puisqu'elle converge presque-sûrement en tant que martingale. Ainsi, alors qu'en première partie nous proposions une solution non-bayésienne aux limites rencontrées par les modèles bayésiens, dans cette seconde partie nous proposons l'inverse. Nous étudions un modèle de jeux de routage dynamiques en information incomplète et explorons les conditions permettant la convergence d'une croyance publique bayésienne vers la vérité.

Notre problématique principale peut être formulée comme suit. Considérons un système de navigation non-stratégique qui agrège et diffuse les choix de déplacement et les temps de trajet afférents à une population renouvelée chaque jour. Si les fonctions de coût qui font correspondre les temps de trajet à la demande sont inconnues, peuvent-elles être efficacement identifiées par le système de navigation? Le cas échéant, lui suffit-il d'un nombre fini d'observations? Au-delà, comment la dynamique de l'apprentissage est-elle influencée par la structure du réseau de routage? Nous apportons une réponse à l'ensemble de ces questions en utilisant une modélisation bayésienne.

8.1.4 Théorie des jeux et modélisation environementale

La troisième partie de cette thèse quitte les questions d'apprentissage et de réseaux pour explorer des approches alternatives à la modélisation des préférences pour le temps et leurs conséquences et termes de décision environnementale. Le point de départ de cette étude est le constant de l'échec d'un grand nombre d'états à tenir les objectifs du protocole de Kyoto et leur progressif désengagement de l'accord à mesure que sa date d'expiration approchait. Le protocole de Kyoto a été négocié en décembre 1997 en étendant la convention-cadre des Nations unies sur les changements climatiques, et est devenu effectif en février 2005. Une première période d'engagement s'est terminée en décembre 2012, suivie par une prolongation du traité pour 8 ans. En décembre 2020, la seconde période d'engagement s'est achevée sur un bilan mitigé. Le protocole de Kyoto est un accord complexe qui a reçu de nombreuse critiques, celles-ci ayant mené à la rédaction des accords de Paris, et nombreuses sont les raisons invoquées pour expliquer l'insuffisance de ses résultats.

La théorie des jeux propose un ensemble d'explications à ce problème. En premier lieu, la nature incomplète du contrat entre les états explique l'émergence de problèmes de *hold-up*. En deuxième lieu, on constate le manque de pouvoir exécutif dans le dispositif proposé: seuls les pays figurant dans l'annexe B s'étaient engagés à des objectifs sous risque de sanctions. Au demeurant, ces sanctions restaient limitées et n'ont pas dissuadé les déviations individuelles. Enfin, il semble que dans l'étude d'accords de long-terme, le modèle usuel de préférences pour le temps, i. e. un escompte exponentiel à taux constant, ne correspond pas aux comportements observés de façon expérimentale. Au contraire, les états semblent montrer des incohérences temporelles en ce sens qu'un décideur aujourd'hui serait en désaccord avec ses choix du futur car, à deux dates différentes, ils ne pondèrent pas le futur de façon égale. Plusieurs modèles existent sur ce sujet et l'escompte à taux non-constant est fréquemment rencontré en économie de l'environnement. Mais ces modèles intègrent rarement le concept de date butoir, alors même qu'elles semblent jouer un rôle détermiant aussi bien dans la conception des accords internationaux que dans la communication scientifique sur le réchauffement climatique.

Dans cette partie, nous proposons une approche à ce problème fondée sur un modèle où deux états sont en concurrence sur un intervalle de temps borné. Ils jouent à des dates déterminées par un processus de Poisson et contrôlent un état du monde commun via leurs actions. De cette façon, nous endogénéisons l'existence d'un taux d'escompte décroissant: plus les états approchent de la limite de l'intervalle de temps, moins il est probable qu'ils puissent rejouer ensuite. L'objectif de cette étude est de déterminer les stratégies optimales et d'étudier le rôle joué par la fréquence du processus de Poisson sur l'existence de comportements coopératifs.

8.1.5 considérations générales

Bien que chaque partie de cette thèse aborde un problème indépendant, plusieurs thèmes sous-jacents les relient. Chacun des modèles considérés repose sur une forme d'incertain: dans les deux premières parties, un état du monde est inconnu et, dans la dernière partie, la séquence des dates auxquelles les agents jouent est aléatoire. En conséquence, l'analyse de chacun de ces modèles repose sur l'utilisation de méthodes stochastiques. Les deux premières parties sont thématiquement réliées: chacune étudie un modèle d'apprentissage, bien que reposant sur des heuristiques différentes, et s'intéressent à la formation et à la nature d'une croyance publique. La présence d'externalités de réseau lie également les deux parties, puisque la diffusion des croyances se fait dans un réseau de communication en première partie, et que la structure combinatoire des ensembles d'actions est caractérisée par un réseau de routage en deuxième partie. Enfin, la deuxième et la troisième partie s'intéressent toutes deux à l'existence de comportements coopératifs dans des situations où les Folk theorems ne s'appliquent pas.

8.2 RÉSUMÉ DES PRINCIPALES CONTRIBUTIONS

Nous résumons ci-dessous les principales contributions des différentes parties de la thèse.

8.2.1 première partie

Dans la première partie de cette thèse nous proposons une extension stochastique aux modèles non-bayésiens d'échange d'opinion. Nous contruisons un modèle de formation d'opinion où les agents communiquent et tirant les valeurs possibles de l'état du monde selon leur croyance plutôt que de communiquer directement leur croyance. Pour cela, nous modélisons les croyances des agents par un système d'urnes en interaction. Des boules de différentes couleurs représentent les valeurs possibles de l'état du monde. A chaque itération du processus de communication, chaque agent tire une boule de son urne et présente le résultat à ses voisins. Puis, chacun ajoute à son urne une boule correspondant à chaque tirage observé. A un instant donné, la croyance d'un agent en l'une des valeurs possibles de l'état du monde est donnée par la proportion de boules de la couleur associée dans son urne. Nous assimilons ainsi l'évolution des coyances au cours du processus de communication à l'évolution des proportions des différentes couleurs dans les urnes.

Dans ce cadre, nous prouvons que la dynamique des croyances converge vers un point stable sous des conditions générales. Notre preuve s'appuie sur les méthodes de l'approximation stochastique. Nous démontrons ensuite qu'à la limite, tous les agents d'un même sous-graphe connecté partagent la même croyance sur les états du monde: un consensus émerge. Ce résultat découle des propriétés algébriques des graphes. Finalement, nous montrons que, dès lors que les croyances initiales couvrent l'espace des états dans son ensemble, le consensus est une variable aléatoire tirée selon à distribution à support complet. Ce résultat contredit fortement les prédictions du modèle de DeGroot. Nous nous consacrons ensuite à la caractérisation de cette croyance limite sur la base de simulations. Nous établissons un ensemble de conjectures à propos de ce consensus. D'abord, il apparaît que la croyance limite suit une loi beta. Cette conjecture est appuyée par la proximité entre notre modèle et le modèle d'urne de Polya. Ensuite, nous supposons que la valeur moyenne de la croyance limite sur chacun des états du monde est égale à leurs probabilités respectives dans le processus d'initialisation des urnes. En d'autres termes, à la limite, la croyance moyenne sur la mauvaise valeur de l'état du monde est égale à la probabilité initiale d'être mal informé. Ce résultat implique que, bien que n'état pas une martingale, le vecteur des proportions se comporte de façon similaire. Ce résultat est démontré pour les graphes réguliers, et demeure à l'état de conjecture dans le cas général.

Cette partie contribue à la littérature portant sur les modèles d'échange d'opinion et sur l'émergence d'un consensus en étendant l'un des modèles canoniques dans un cadre stochastique et en montrant en quoi les prédictions sont fortement changées. Des travaux préexistant critiquaient déjà la robustesse du modèle de DeGroot. Nous apportons des éléments nouveaux à la controverse, tout en proposant un modèle où des phénomènes extrêmes, observés dans la réalité mais absents des prédictions de DeGroot, peuvent se produire avec une probabilité positive. L'objectif initial de cette étude était de proposer une métrique sur les croyances permettant de construire un modèle de désinformation stratégique. Bien que la nature exacte du consensus limite demeure à l'état de conjecture, nous avons bon espoir de pouvoir démontrer ces résultats et de les mettre en oeuvre dans un modèle plus large.

8.2.2 deuxième partie

Dans la deuxième partie de cette thèse, nous considérons un jeu de routage non atomique répété où les fonctions de coût de chaque arête du réseau de routage dépendent conjointement de la demande sur cette arête et d'un paramètre d'état inconnu et ne variant pas au cours du temps. L'ensemble des états est fini et doté d'une croyance *a priori*. À chaque période de temps, une génération éphémère d'utilisateurs ayant une demande totale donnée joue le jeu et réalise un équilibre de Wardrop par rapport aux coûts espérés sur les arêtes : chaque chemin recevant une charge positive d'agents minimise le coût espéré. Pour chaque arête utilisée, sa charge et le coût réalisé correspondant deviennent des informations publiques pour les générations suivantes. On suppose que chaque génération connaît toute l'histoire passée du jeu et met à jour ses croyances de manière bayésienne. La séquence des demandes des différentes générations est supposée être aléatoire, i.i.d..

Nous considérons deux concepts différents d'apprentissage social : dans le cas d'un apprentissage fort, les joueurs finissent par apprendre le véritable état du monde ; dans le cadre d'un apprentissage faible, ils apprennent à jouer le jeu comme si le véritable état du monde était connu. Nous montrons que l'apprentissage faible est un concept strictement plus faible que l'apprentissage fort et que les conditions pour atteindre l'un ou l'autre dépendent de la topologie du réseau et du support de la demande aléatoire. Notre théorème principal prouve que l'apprentissage faible se produit si le réseau de routage est série-parallèle et si les fonctions de coût et le support de la demande sont non-bornés. De plus, nous montrons que l'apprentissage fort est obtenu avec les mêmes conditions préalables et la condition supplémentaire que la demande ait un support complet sur $\mathbb{R}_+$. L'intuition derrière ce résultat est la suivante : lorsque la demande est stochastique, les flux d'équilibre varient. Cela génère des observations des fonctions de coût pour différentes valeurs de charges. En se basant sur les résultats de Cominetti, Dose, and Scarsini [44] sur la variation des flux d'équilibre par rapport à la demande, nous prouvons que dans un réseau série-parallèle, lorsque la demande augmente, toutes les arêtes sont utilisées à l'équilibre et les charges d'équilibre sont non-bornées. Ceci implique que les fonctions de coût seront observées à des niveaux qui permettent de distinguer les états pertinents presque-sûrement. Enfin, nous prouvons que la condition sur la topologie du réseau est nécessaire : pour les réseaux qui ne la satisfont pas, nous montrons qu'il existe une affectation des fonctions de coût et des capacités telle que l'apprentissage faible échoue pour toute distribution de la demande.

Cette partie contribue à la littérature sur l'apprentissage social en offrant un exemple de jeu *large* où des séquences d'ensembles continus de joueurs permettent l'apprentissage. Nous contribuons également à la littérature sur les jeux de routage en la reconnectant à une approche plus traditionnelle sur l'apprentissage dans les jeux. Ainsi, nous offrons un modèle d'apprentissage qui est immunisé contre le chaos et où les croyances convergent en temps fini.

8.2.3 TROISIÈME PARTIE

Dans la troisième partie de cette thèse, notre objectif est double. D'un point de vue théorique, nous construisons un modèle de jeux de révision avec des flux de paiement et un état cumulatif pour modéliser l'arbitrage intertemporel auquel font face les Etats impliqués dans la transition environnementale. Deux joueurs agissent sur un intervalle de temps fini et se voient proposer à des dates stochastiques de réviser leur décision. Cela produit un flux de paiements et détermine un état commun du monde qui capture la part de temps pendant laquelle chaque joueur a joué l'action non-coopérative. Cet état est une approximation de l'impact des émissions des joueurs au cours du jeu. Lorsque le temps atteint la fin de l'intervalle, un gain final est déterminé en fonction de l'état du monde.

Ce projet reste à un stade précoce de développement et l'essentiel de nos efforts a été consacré à l'identification du modèle adéquat. Néanmoins, nous fournissons un programme détaillé des résultats sur lesquels nous travaillons. En termes pratiques, nous avons d'abord l'intention de prouver l'existence d'un profil symétrique de stratégies optimales sous la forme de stratégies de seuil : avant un seuil de temps qui dépend négativement de la fréquence attendue des temps de révision, les deux joueurs jouent l'équilibre de Nash du jeu à un coup, puis passent à un profil d'action coopératif. À cet égard, plus la fréquence à laquelle les joueurs révisent leurs stratégies est élevée, plus la durée du régime coopératif est courte. Cette relation entre la fréquence de décision et le comportement coopératif est déjà observée dans les jeux de révision sans flux de paiement.

Ce projet contribue à la littérature récente sur les jeux de révision en proposant une version alternative du modèle où les joueurs arbitrent entre un flux de paiements instantanés et un gain terminal. Il contribue également à la littérature sur la prise de décision environnementale en reliant théoriquement la fréquence à laquelle les politiques sont révisées avec l'émergence d'un comportement coopératif. Enfin, tout en étant équivalente aux jeux stochastiques, la classe des jeux de révision ne permet pasl'application des Folk theorems existants. Prouver l'existence d'équilibres parfaits en sous-jeu soutenant une forme de coopération est donc un apport à ce qui est connu sur la coopération dans les jeux dynamiques non-coopératifs.

Part IV

APPENDIX

a

SIMULATIONS USED IN PART I

In this appendix, we provide elements on the simulations used to support conjectures made in Chapter 4. Appendix a.1 provides elements of the code used to simulate the communication process. Appendix a.2 gives additional visual outputs from those simulations that further support our conjectures.

A.1 PYTHON CODE

We provide the main elements coded to simulate the communication process. The code is annotated with details on the nature of the different functions and a description of the parameters. This code is a simplified version of the main code we used. The main difference between the two version is that the final code was optimized to reduce its run time and adapted to a multiprocessing environment.

```
###################################
###0: Packages required
###################################

import networkx as nx
import matplotlib.pyplot as plt
import random
from scipy.stats import bernoulli
import csv
import numpy as np



###################################
###1: Functions definitions
###################################


## Star Graph generator
```

```
def star_gen(n):
    g=nx.Graph()
    for i in range(1,n+1):
        g.add_node(i)
        if i>1:
            g.add_edge(1,i)
    return g


### 1.1 Urns initialization
###################################

## IID Bernoulli initialization (g=graph, error=alpha, theta=
    true value)

def init_urns(g,error,theta):
    urns={}
    for i in g.nodes():
        own={}
        gen=random.uniform(0,1)
        if gen>error:
            own['W']= theta
        else:
            own['W']=1-theta
        own['B']=1-float(own['W'])
        urns[i]=own
    nx.set_node_attributes(g,urns,'urn')



## Uniform initialization (1W 1B, g=graph)

def init_unif(g):
    urns={}
    for i in g.nodes():
        own={}
        own['W']=1
        own['B']=1
        urns[i]=own
    nx.set_node_attributes(g,urns,'urn')




### 1.2 Updating processes
###################################

## Draw step (g=graph, state=vector of proportions)

def draw_step(g,state):
    draw={}
```

```
    cb=0
cw=0    for j in g.nodes():
        urn = nx.get_node_attributes(g,'urn')[j]
        rat=float(urn['W'])/(float(urn['B'])+float(urn['W']))
        state[j].append(rat)
        #print rat
        gen=random.uniform(0,1)
        if gen>rat:
            tirage = 'B'
            cb+=1
        else:
            tirage = 'W'
            cw+=1
        draw[j]=tirage
    #print draw
    return draw


## Aggregates neighbors' draws and updates urns (g=graph, draw=
    vector of draws)

def gather_neighbor(g,draw):
    nurn={}
    for i in g.nodes():
        N=dict()
        N[i]=nx.all_neighbors(g,i)
        own={}
        B=nx.get_node_attributes(g,'urn')[i]['B']
        W=nx.get_node_attributes(g,'urn')[i]['W']
        for j in N[i]:
            if 'W' in draw[j]:
                W+=1
            if 'B' in draw[j]:
                B+=1
        own['W']=W
        own['B']=B
        nurn[i]=own
    nx.set_node_attributes(g,nurn,'urn')


## Updates selecting on neighbor at random (g=graph, draw=
    vector of draws)

def rand_neighbor(g,draw):
    nurn={}
    for i in g.nodes():
        N=list(nx.all_neighbors(g,i))
        own={}
        B=nx.get_node_attributes(g,'urn')[i]['B']
        W=nx.get_node_attributes(g,'urn')[i]['W']
        s=np.random.randint(0,len(N))
        n=N[s]
```

```
        if 'W' in draw[n]:
            W+=1
        if 'B' in draw[n]:
            B+=1
        own['W']=W
        own['B']=B
        nurn[i]=own
    nx.set_node_attributes(g,nurn,'urn')


## Process functions combining draws and updating (g=graph, t=
    number of iterations, state=vector of proportions)

def com_sequence(g,t,state):
    l=1
    while l<t+1:
        draw=draw_step(g,state)
        gather_neighbor(g,draw)
        l+=1

def com_randneigh(g,t):
    l=1
    while l<t+1:
        draw=draw_step(g)
        rand_neighbor(g,draw)
        l+=1


###################################
### 2 Simulation Function
###################################

## 2.1 Star graph simulation
## (size=N, err=1-alpha, sims= number of simulations, lenproc=
    number of iterations)
###################################

def do_star(size,err,sims,lenproc):
    cons=[]
    for t in range(0,sims):
        g=star_gen(size)
        state={}
        for s in g.nodes():
            state[s]=[]
        init_urns(g,err,1)
        com_sequence(g,lenproc,state)
        cons.append(state[1][len(state[1])-1])
    print(cons)
    return(cons)


## 2.2 Regular graph simulation
```

```
##  (size=N, deg= degree of G, err=1-alpha, sims= number of
    iterations,
##  lenproc=number of iterations)#
  ##################################


def do_reg(size,deg,err,sims,lenproc):
    cons=[]
    for t in range(0,sims):
        g=nx.random_regular_graph(deg,size)
        while nx.is_connected(g) == "False":
            g=nx.random_regular_graph(deg,size)
        state={}
        for s in g.nodes():
            state[s]=[]
        init_urns(g,err,1)
        com_sequence(g,lenproc,state)
        cons.append(state[1][len(state[1])-1])
    print(cons)
    return(cons)


## 2.3 Complete graph simulation
##  (size=N, err=1-alpha, sims= number of simulations, lenproc=
    number of
##  iterations, res=collections of proportions)
###################################

def do_comp(size,err,sims,lenproc,res):
   cons=[]
   for t in range(0,sims):
       g=nx.complete_graph(size)
       state={}
       for s in g.nodes():
           state[s]=[]
       init_urns(g,err,1)
       com_sequence(g,lenproc,state)
       cons.append(state[1][len(state[1])-1])
    res.append(cons)
```

A.2 ADDITIONAL OUTPUTS

In this section, we provide additional outputs obtained from the simulations.

A.2.1 Beta distribution fit for varying values of α

The outputs below show the empirical and theoretical densities and cumulative distribution functions, quantile-quantile and P-P plots obtained in fitting a beta distribution on the outcomes of $N = 14000$ simulations run on a 5-regular graph of size 50 for values of α ranging from 0.1 to 0.9. Similar results were obtained in other graph structures. These outputs strongly support Conjecture 72 ad Conjecture 76.

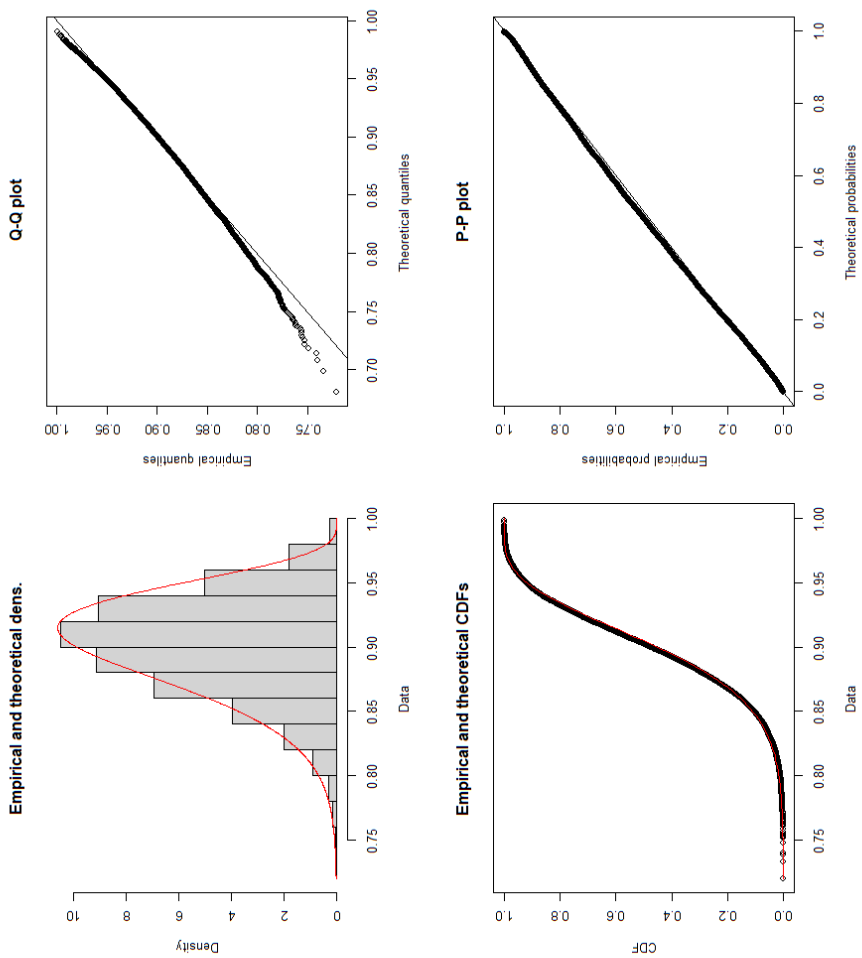


Figure a.1: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.90$, $n = 14000$ observations.

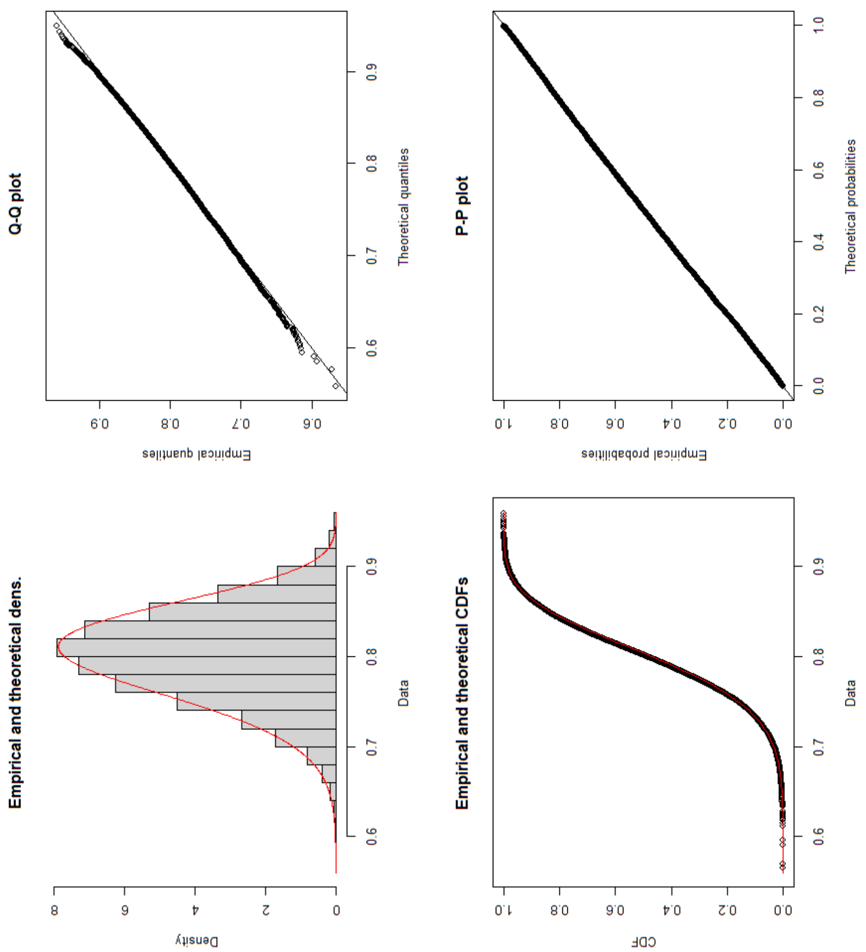


Figure a.2: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.80$, $n = 14000$ observations.

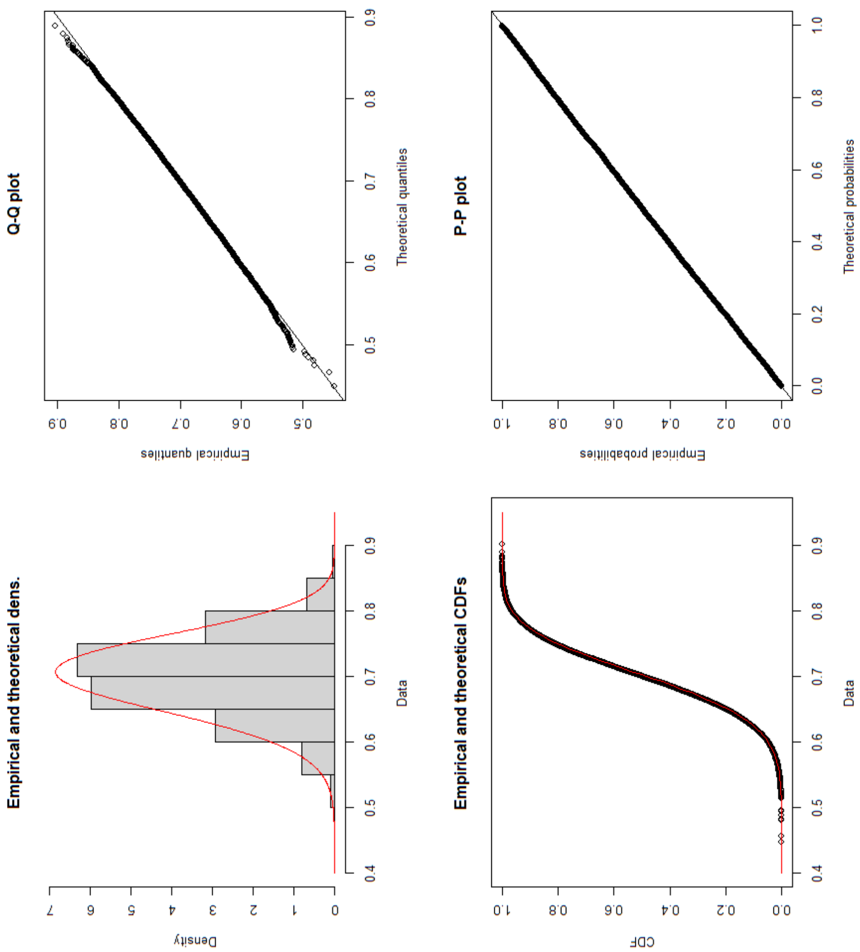


Figure a.3: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.70$, $n = 14000$ observations.

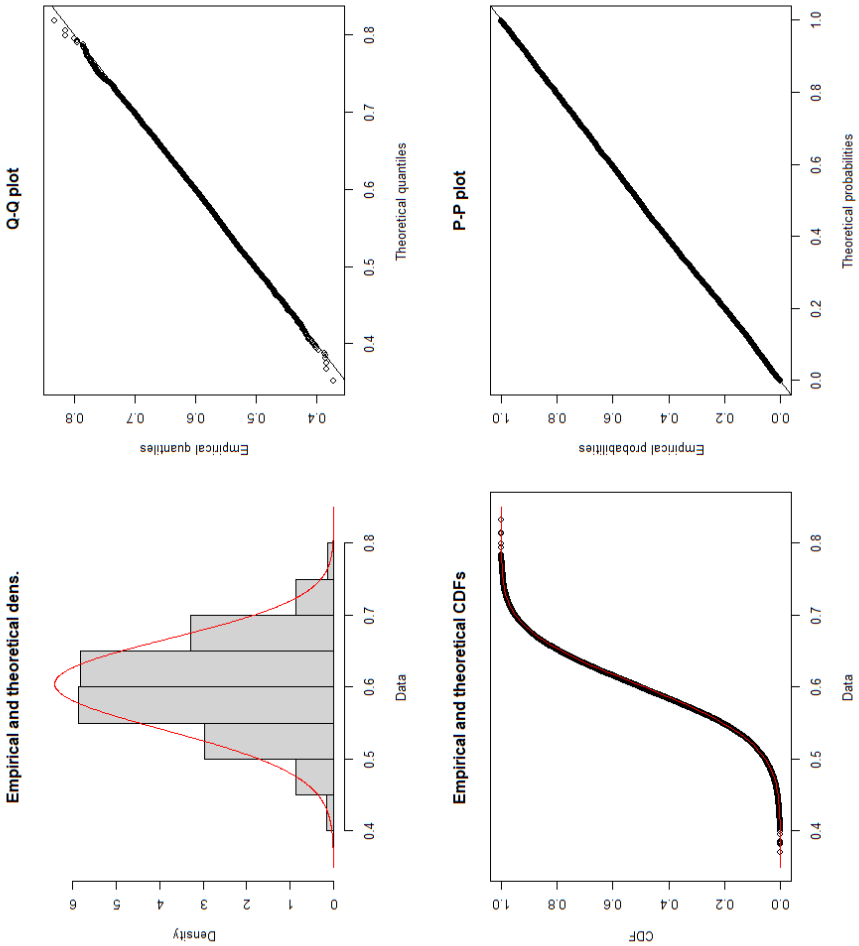


Figure a.4: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.60$, $n = 14000$ observations.

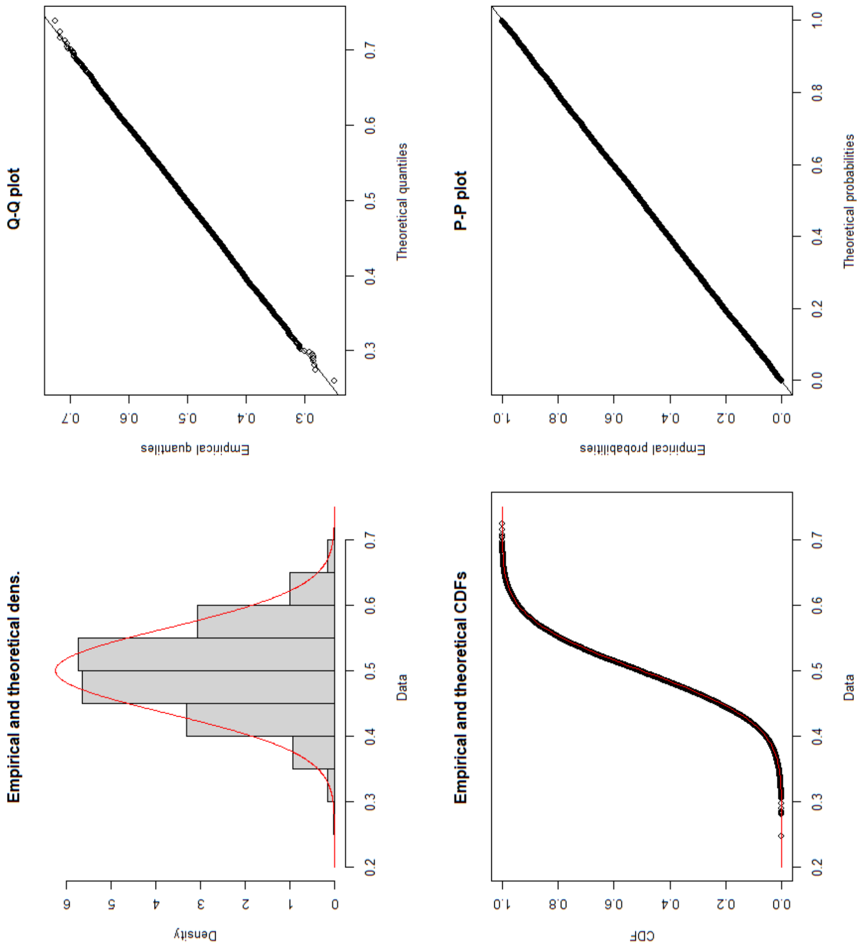


Figure a.5: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.50$, $n = 14000$ observations.

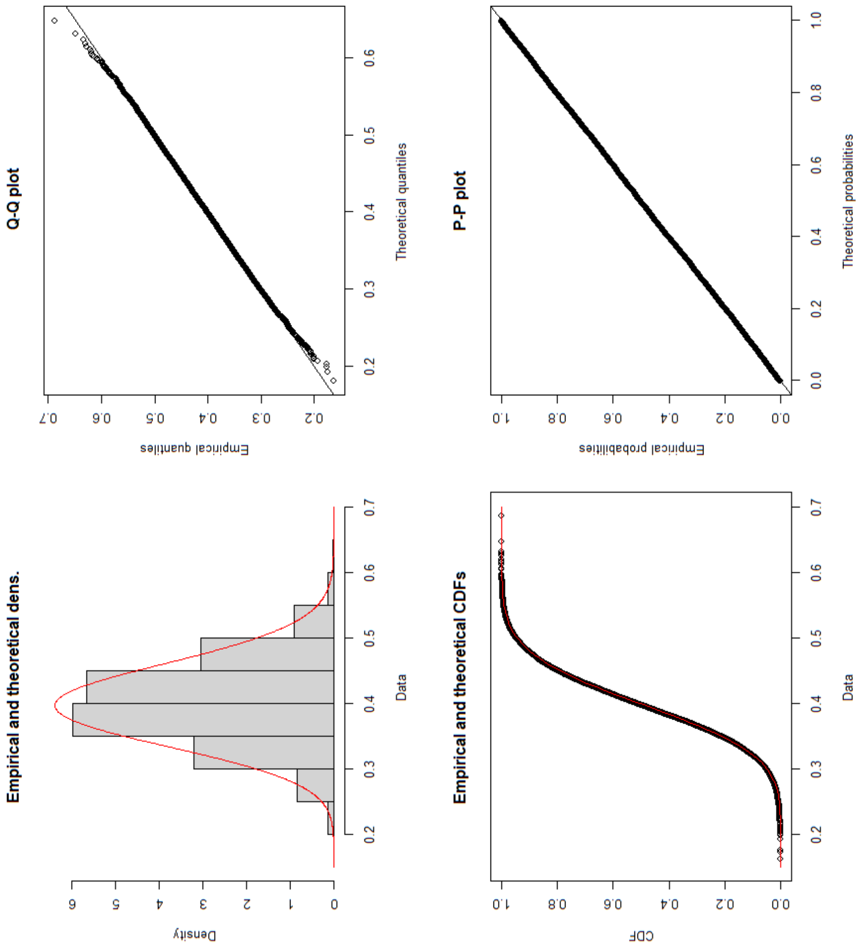


Figure a.6: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.40$, $n = 14000$ observations.

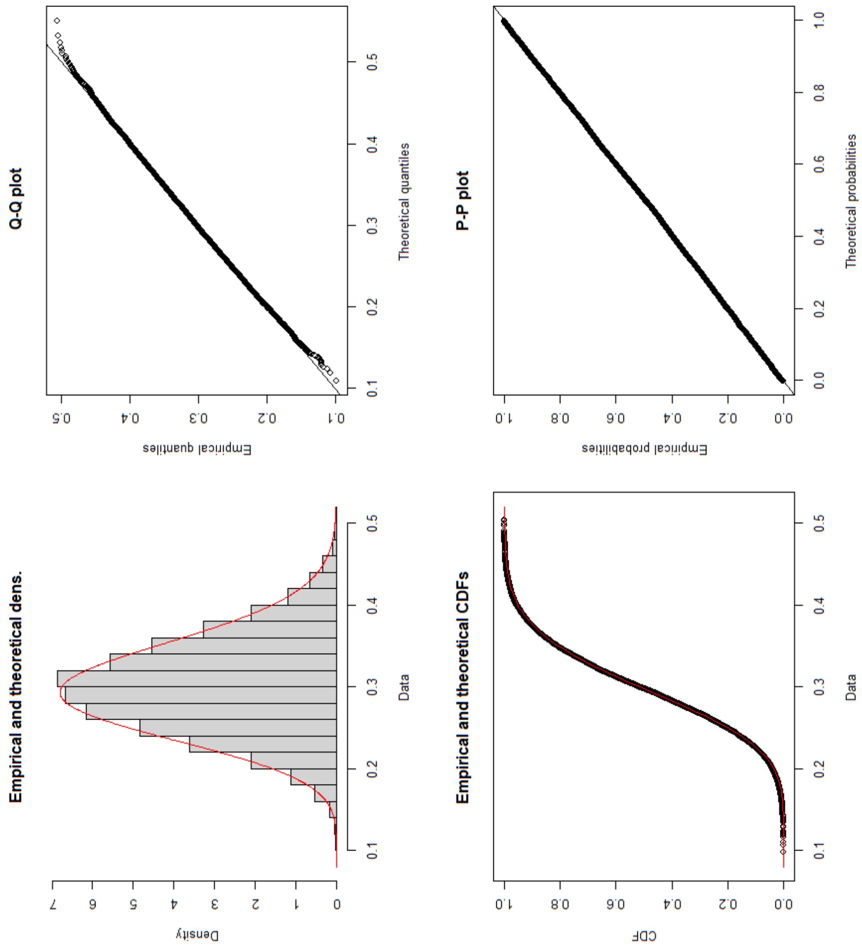


Figure a.7: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.30$, $n = 14000$ observations.

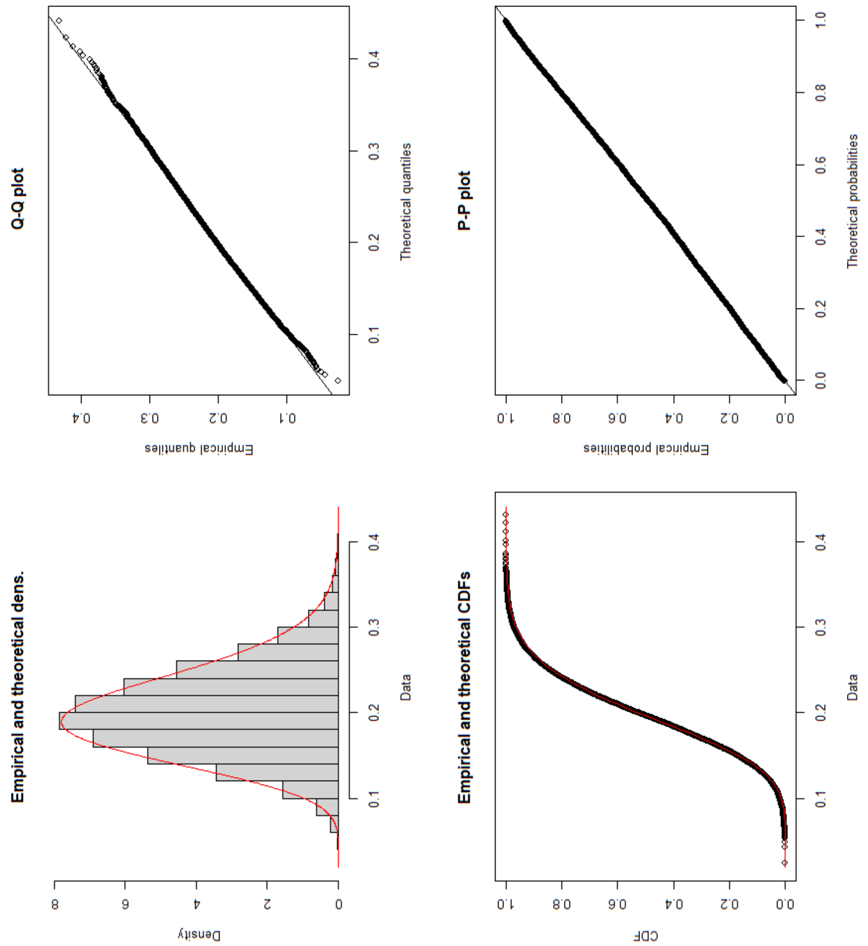


Figure a.8: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.20$, $n = 14000$ observations.

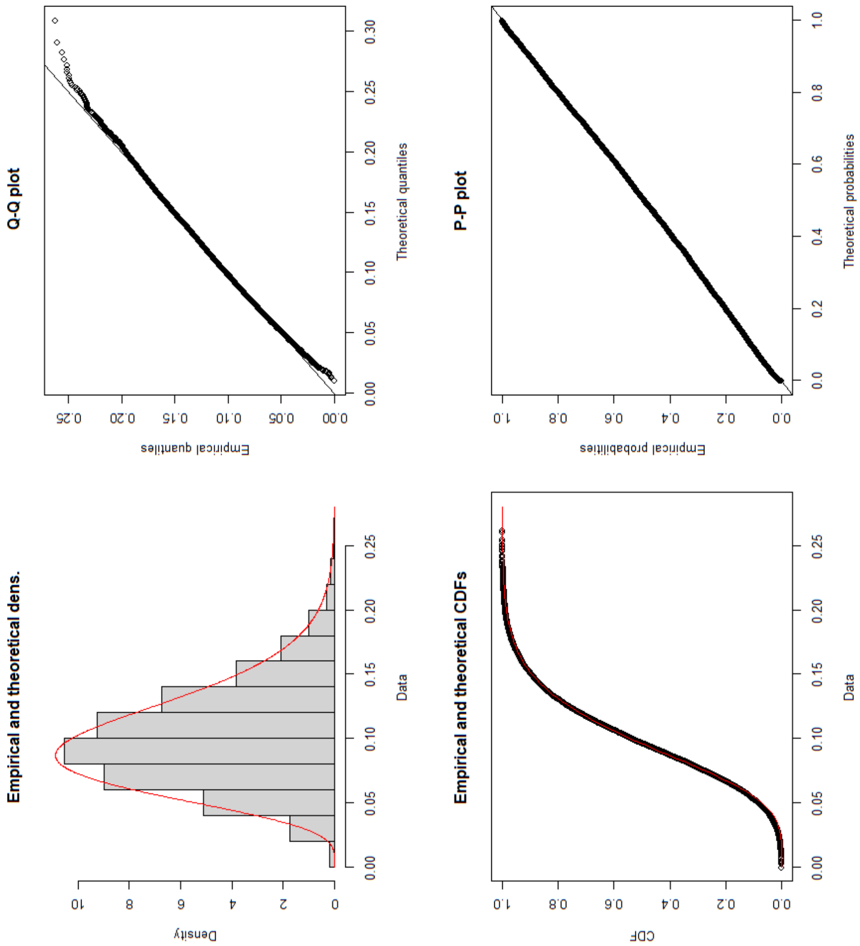


Figure a.9: Fitness measures for beta distribution on a 5-regular graph with $\alpha = 0.10$, $n = 14000$ observations.

BIBLIOGRAPHY

[1] Daron Acemoglu, Philippe Aghion, Leonardo Bursztyn, and David Hemous. "The environment and directed technical change." In: *American economic review* 102.1 (2012), pp. 131–66.

[2] Daron Acemoglu, Kostas Bimpikis, and Asuman Ozdaglar. "Dynamics of information exchange in endogenous social networks." In: *Theoretical Economics* 9.1 (2014), pp. 41–97.

[3] Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. "Bayesian learning in social networks." In: *The Review of Economic Studies* 78.4 (2011), pp. 1201–1236.

[4] Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. "Informational Braess' paradox: the effect of information on traffic congestion." In: *Oper. Res.* 66.4 (2018), pp. 893–917. ISSN: 0030-364X.

[5] Daron Acemoglu and Asuman Ozdaglar. "Opinion dynamics and learning in social networks." In: *Dynamic Games and Applications* 1.1 (2011), pp. 3–49.

[6] Nabil I Al-Najjar and Rann Smorodinsky. "Large nonanonymous repeated games." In: *Games and Economic Behavior* 37.1 (2001), pp. 26–39.

[7] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. "Trends in the diffusion of misinformation on social media." In: *Research & Politics* 6.2 (2019), p. 2053168019848554.

[8] Elliot Anshelevich, Anirban Dasgupta, Eva Tardos, and Tom Wexler. "Near-optimal network design with selfish agents." In: *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*. 2003, pp. 511–520.

[9] Itai Arieli and Manuel Mueller-Frank. "A general analysis of sequential social learning." In: *Math. Oper. Res.* forthcoming (2021).

[10] Itai Ashlagi, Dov Monderer, and Moshe Tennenholtz. "Two-terminal routing games with unknown active players." In: *Artificial Intelligence* 173.15 (2009), pp. 1441–1455. ISSN: 0004-3702.

[11] Robert J Aumann. "Mixed and Behavior Strategies in Infinite Extensive Games." In: *Advances in Game Theory.(AM-52), Volume 52*. Ed. by Melvin Dresher, Lloyd S Shapley, and Albert William Tucker. Princeton University Press, 1964. Chap. 28, pp. 193–216.

[12] Robert J Aumann. "Agreeing to disagree." In: *The annals of statistics* (1976), pp. 1236–1239.

[13] Robert J Aumann, Michael Maschler, and Richard E Stearns. *Repeated games with incomplete information*. MIT press, 1995.

[14] Robert J Aumann and Lloyd S Shapley. "Long-term competition—a game-theoretic analysis (1994 re-ed.)" In: *Essays in game theory*. Springer, 1976, pp. 1–15.

[15] Venkatesh Bala and Sanjeev Goyal. "Learning from neighbours." In: *The review of economic studies* 65.3 (1998), pp. 595–621.

[16] Abhijit V Banerjee. "A simple model of herd behavior." In: *The quarterly journal of economics* 107.3 (1992), pp. 797–817.

[17] Abhijit V. Banerjee. "A simple model of herd behavior." English. In: *Quart. J. Econom.* 107.3 (1992), pp. 797–817. ISSN: 00335533.

[18] Marco Battaglini and Bård Harstad. "Participation and duration of environmental agreements." In: *Journal of Political Economy* 124.1 (2016), pp. 160–204.

[19] Pierpaolo Battigalli and Danilo Guaitoli. "Conjectural equilibria and rationalizability in a game with incomplete information." In: *Decisions, Games and Markets*. Ed. by Pierpaolo Battigalli, Aldo Montesano, and Fausto Panunzi. Springer, 1997, pp. 97–124.

[20] Martin J. Beckmann, C.B McGuire, and Christopher B Winsten. *Studies in the Economics of Transportation*. New Haven, CT: Yale University Press, 1956.

[21] Michel Benaïm. "Dynamics of stochastic approximation algorithms." In: *Seminaire de probabilites XXXIII*. Springer, 1999, pp. 1–68.

[22] J-P Benoit and Vijay Krishna. "Finitely repeated games." In: *Econometrica* 53.4 (1985), pp. 905–922.

[23] J-P Benoit and Vijay Krishna. "Nash equilibria of finitely repeated games." In: *International Journal of Game Theory* 16.3 (1987), pp. 197–204.

[24] Petra Berenbrink and Oliver Schulte. "Evolutionary equilibrium in Bayesian routing games: specialization and niche formation." In: *Theoret. Comput. Sci.* 411.7-9 (2010), pp. 1054–1074. ISSN: 0304-3975.

[25] Edward L Bernays. *Propaganda*. Horace Liveright, New York, 1928.

[26] Umang Bhaskar, Katrina Ligett, Leonard J. Schulman, and Chaitanya Swamy. "Achieving target equilibria in network routing games without knowing the latency functions." In: *Games Econom. Behav.* 118 (2019), pp. 533–569. ISSN: 0899-8256.

[27] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. "A theory of fads, fashion, custom, and cultural change in informational cascades." In: *J. Polit. Econ.* 100.5 (1992), pp. 992–1026.

[28] David Blackwell and Lester Dubins. "Merging of opinions with increasing information." In: *Ann. Math. Statist.* 33 (1962), pp. 882–886. ISSN: 0003-4851.

[29] Vivek S. Borkar. *Stochastic Approximation: a Dynamical System Viewpoint*. Cambridge University Press, 2008.

[30] Dietrich Braess. "Über ein Paradoxon aus der Verkehrsplanung." In: *Unternehmensforschung* 12 (1968), pp. 258–268. ISSN: 0042-0573.

[31] Thomas Brenner and Ulrich Witt. "Melioration learning in games with constant and frequency-dependent payoffs." In: *Journal of Economic Behavior & Organization* 50.4 (2003), pp. 429–448.

[32] Riccardo Calcagno, Yuichiro Kamada, Stefano Lovo, and Takuo Sugaya. "Asynchronicity and coordination in common and opposing interest games." In: *Theoretical Economics* 9.2 (2014), pp. 409–434.

[33] Riccardo Calcagno and Stefano Lovo. "Preopening and equilibrium selection." In: (2010).

[34] Emilio Calvo, Santiago J Rubio, et al. "Dynamic models of international environmental agreements: a differential game approach." In: *International Review of Environmental and Resource Economics* 6.4 (2012), pp. 289–339.

[35] Guillermo Caruana and Liran Einav. "A theory of endogenous commitment." In: *The Review of Economic Studies* 75.1 (2008), pp. 99–116.

[36] Christophe P Chamley. *Rational herds: Economic models of social learning*. Cambridge University Press, 2004.

[37] Samprit Chatterjee and Eugene Seneta. "Towards consensus: Some convergence theorems on repeated averaging." In: *Journal of Applied Probability* 14.1 (1977), pp. 89–97.

[38] Bo Chen and Satoru Takahashi. "A folk theorem for repeated games with unequal discounting." In: *Games and Economic Behavior* 76.2 (2012), pp. 571–581.

[39] Xujin Chen, Zhuo Diao, and Xiaodong Hu. "Network characterizations for excluding Braess's paradox." In: *Theory Comput. Syst.* 59.4 (2016), pp. 747–780. ISSN: 1432-4350.

[40] Xujin Chen, Nikolai Gravin, Martin Hoefer, and Ruta Mehta. *Web and Internet Economics: 16th International Conference, WINE 2020, Beijing, China, December 7–11, 2020, Proceedings*. Vol. 12495. Springer Nature, 2020.

[41] Thiparat Chotibut, Fryderyk Falniowski, Michał Misiurewicz, and Georgios Piliouras. "The route to chaos in routing games: Population increase drives period-doubling instability, chaos & inefficiency with Price of Anarchy equal to one." In: *arXiv preprint arXiv:1906.02486* (2019).

[42] R. Cominetti, M. Scarsini, M. Schröder, and N. Stier-Moses. "Price of anarchy in stochastic atomic congestion games with affine costs." arXiv 1903.03309. 2019.

[43] R. Cominetti, M. Scarsini, M. Schröder, and N. Stier-Moses. "Convergence of large atomic congestion games." arXiv 2001.02797. 2020.

[44] Roberto Cominetti, Valerio Dose, and Marco Scarsini. "The price of anarchy in routing games as a function of the demand." arXiv:1907.10101. 2021.

[45] Roberto Cominetti, Valerio Dose, and Marco Scarsini. "The price of anarchy in routing games as a function of the demand." In: *Math. Program.* forthcoming (2021).

[46] Nicolas de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. de l'Imprimerie Royale, 1785.

[47] José Correa, Ruben Hoeksma, and Marc Schröder. "Network congestion games are robust to variable demand." In: *Transportation Res. Part B* 119 (2019), pp. 69–78. ISSN: 0191-2615.

[48] Pierre Coucheney, Bruno Gaujal, and Panayotis Mertikopoulos. "Penalty-regulated dynamics and robust learning procedures in games." In: *Mathematics of Operations Research* 40.3 (2015), pp. 611–633.

[49] Irene Crimaldi, Paolo Dai Pra, and Ida Germana Minelli. "Fluctuation theorems for synchronization of interacting Pólya's urns." In: *Stochastic processes and their applications* 126.3 (2016), pp. 930–947.

[50] Paolo Dai Pra, Pierre-Yves Louis, and Ida G Minelli. "Synchronization via interacting reinforcement." In: *Journal of Applied Probability* 51.2 (2014), pp. 556–568.

[51] Bruno De Finetti. "Funzione caratteristica di un fenomeno aleatorio." In: *Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928*. 1929, pp. 179–190.

[52] Morris H DeGroot. "Reaching a consensus." In: *Journal of the American Statistical Association* 69.345 (1974), pp. 118–121.

[53] Peter M DeMarzo, Dimitri Vayanos, and Jeffrey Zwiebel. "Persuasion bias, social influence, and unidimensional opinions." In: *The Quarterly journal of economics* 118.3 (2003), pp. 909–968.

[54] Reinhard Diestel. *Graph Theory (5th edition)*. Graduate Texts in Mathematics. Springer, 2017.

[55] Pradeep Dubey and Mamoru Kaneko. "Information patterns and Nash equilibria in extensive games. I." In: *Math. Social Sci.* 8.2 (1984), pp. 111–139. ISSN: 0165-4896.

[56] R. J. Duffin. "Topology of series-parallel networks." In: *J. Math. Anal. Appl.* 10 (1965), pp. 303–318. ISSN: 0022-247X.

[57] Florian Eggenberger and George Pólya. “Über die statistik verketteter vorgänge.” In: *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 3.4 (1923), pp. 279–289.

[58] L. R. Ford Jr. and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, N.J., 1962, pp. xii+194.

[59] Dylan J Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, and Eva Tardos. “Learning in games: Robustness of fast convergence.” In: *Advances in Neural Information Processing Systems*. 2016, pp. 4734–4742.

[60] Dimitris Fotakis, Vasilis Gkatzelis, Alexis C. Kaporis, and Paul G. Spirakis. “The impact of social ignorance on weighted congestion games.” In: *Theory Comput. Syst.* 50.3 (2012), pp. 559–578. ISSN: 1432-4350.

[61] Shane Frederick, George Loewenstein, and Ted O’donoghue. “Time discounting and time preference: A critical review.” In: *Journal of economic literature* 40.2 (2002), pp. 351–401.

[62] Noah E Friedkin and Eugene C Johnsen. “Social positions in influence networks.” In: *Social networks* 19.3 (1997), pp. 209–222.

[63] Drew Fudenberg and David K. Levine. “Self-confirming equilibrium.” In: *Econometrica* 61.3 (1993), pp. 523–545. ISSN: 0012-9682.

[64] Drew Fudenberg and Eric Maskin. “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information.” In: *Econometrica* 54.3 (1986), pp. 533–554.

[65] Martin Gairing. “Malicious Bayesian congestion games.” In: *Approximation and Online Algorithms*. Vol. 5426. Lecture Notes in Comput. Sci. Springer, Berlin, 2009, pp. 119–132.

[66] Martin Gairing, Burkhard Monien, and Karsten Tiemann. “Selfish routing with incomplete information.” In: *Theory Comput. Syst.* 42.1 (2008), pp. 91–130. ISSN: 1432-4350.

[67] Jason Gaitonde and Éva Tardos. “Stability and learning in strategic queuing systems.” In: *Proceedings of the 21st ACM Conference on Economics and Computation*. EC ’20. Virtual Event, Hungary: Association for Computing Machinery, 2020, pp. 319–347. ISBN: 9781450379755.

[68] Jason Gaitonde and Eva Tardos. “Virtues of patience in strategic queuing systems.” arXiv:2011.10205. 2020.

[69] Douglas Gale and Shachar Kariv. "Bayesian learning in social networks." In: *Games and economic behavior* 45.2 (2003), pp. 329–346.

[70] John D Geanakoplos and Heraklis M Polemarchakis. "We can't disagree forever." In: *Journal of Economic theory* 28.1 (1982), pp. 192–200.

[71] Fabien Gensbittel, Stefano Lovo, Jérôme Renault, and Tristan Tomala. "Zero-sum revision games." In: *Games and Economic Behavior* 108 (2018), pp. 504–522.

[72] Itzhak Gilboa. "Rationality and the Bayesian paradigm." In: *Journal of Economic Methodology* 22.3 (2015), pp. 312–334.

[73] Christian Gollier and Martin L Weitzman. "How should the distant future be discounted when discount rates are uncertain?" In: *Economics Letters* 107.3 (2010), pp. 350–353.

[74] Benjamin Golub and Matthew O Jackson. "Naive learning in social networks and the wisdom of crowds." In: *American Economic Journal: Microeconomics* 2.1 (2010), pp. 112–49.

[75] Benjamin Golub and Evan Sadler. "Learning in social networks." In: *Available at SSRN 2919146* (2017).

[76] Michel Grabisch and Agnieszka Rusinowska. "A survey on nonstrategic models of opinion dynamics." In: *Games* 11.4 (2020), p. 65.

[77] Edward J Green. "Noncooperative price taking in large dynamic markets." In: *Noncooperative Approaches to the Theory of Perfect Competition*. Elsevier, 1982, pp. 37–64.

[78] John C Harsanyi. "Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model." In: *Management science* 14.3 (1967).

[79] Bård Harstad. "Climate contracts: A game of emissions, investments, negotiations, and renegotiations." In: *Review of Economic Studies* 79.4 (2012), pp. 1527–1557.

[80] Alain Haurie and Patrice Marcotte. "On the relationship between Nash—Cournot and Wardrop equilibria." In: *Networks* 15.3 (1985), pp. 295–308.

[81] Edwin Hewitt and Leonard J Savage. "Symmetric measures on Cartesian products." In: *Transactions of the American Mathematical Society* 80.2 (1955), pp. 470–501.

[82] Bruce M Hill, David Lane, William Sudderth, et al. "Exchangeable urn processes." In: *The Annals of Probability* 15.4 (1987), pp. 1586–1592.

[83] Morris W Hirsch, Stephen Smale, and Robert L Devaney. *Differential equations, dynamical systems, and an introduction to chaos*. Academic press, 2012.

[84] Ron Holzman and Nissan Law-yone. "Network structure and strong equilibrium in route selection games." In: *Math. Social Sci.* 46.2 (2003), pp. 193–205. ISSN: 0165-4896.

[85] Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.

[86] Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.

[87] Matthew O Jackson and Leeat Yariv. "Present bias and collective dynamic choice in the lab." In: *American Economic Review* 104.12 (2014), pp. 4184–4204.

[88] Matthew O Jackson and Leeat Yariv. "Collective dynamic choice: the necessity of time inconsistency." In: *American Economic Journal: Microeconomics* 7.4 (2015), pp. 150–78.

[89] Reinoud Joosten, Thomas Brenner, and Ulrich Witt. "Games with frequency-dependent stage payoffs." In: *International journal of game theory* 31.4 (2003), pp. 609–620.

[90] Ehud Kalai and Ehud Lehrer. "Rational Learning Leads to Nash Equilibrium." In: *Econometrica* 61.5 (1993), pp. 1019–1045.

[91] Ehud Kalai and Ehud Lehrer. "Rational learning leads to Nash equilibrium." In: *Econometrica* 61.5 (1993), pp. 1019–1045. ISSN: 0012-9682.

[92] Ehud Kalai and Ehud Lehrer. "Rational learning leads to Nash equilibrium." In: *Econometrica* 61.5 (1993), pp. 1019–1045. ISSN: 0012-9682.

[93] Yuichiro Kamada and Michihiro Kandori. "Revision games." In: *Harvard University* (2009).

[94] Yuichiro Kamada and Michihiro Kandori. "Revision games." In: *Econometrica* 88.4 (2020), pp. 1599–1630.

[95] Mamoru Kaneko. "Some remarks on the folk theorem in game theory." In: *Mathematical Social Sciences* 3.3 (1982), pp. 281–290.

[96] Achim Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2013.

[97] Asen Kochov and Yangwei Song. "Repeated Games with Endogenous Discounting." In: *Available at SSRN 2714337* (2016).

[98] Ilan Kremer, Yishay Mansour, and Motty Perry. "Implementing the "wisdom of the crowd"." In: *Journal of Political Economy* 122.5 (2014), pp. 988–1012.

[99] Harold W Kuhn. "11. Extensive Games and the Problem of Information." In: *Contributions to the Theory of Games (AM-28), Volume II*. Princeton University Press, 1953, pp. 193–216.

[100] H. Kushner and G.G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Stochastic Modelling and Applied Probability. Springer New York, 2003. ISBN: 9780387008943.

[101] Rida Laraki, Jérôme Renault, and Sylvain Sorin. *Bases Mathématiques de la théorie des jeux*. Editions de l'Ecole Polytechnique, 2013.

[102] Rida Laraki, Jérôme Renault, and Tristan Tomala. *Théorie des jeux. Introduction à la théorie des jeux répétés (X-UPS 2006).* Editions de l'Ecole Polytechnique, 2008.

[103] Sophie Laruelle, Gilles Pages, et al. "Randomized urn models revisited using stochastic approximation." In: *Annals of Applied Probability* 23.4 (2013), pp. 1409–1436.

[104] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. "The science of fake news." In: *Science* 359.6380 (2018), pp. 1094–1096.

[105] Ehud Lehrer and Ady Pauzner. "Repeated games with differential time preferences." In: *Econometrica* 67.2 (1999), pp. 393–412.

[106] Sam Levin. "Facebookpromised to tackle fake nes. But the evidence shows it's not working." In: *The Guardian* (May 16th, 2017).

[107] Yehuda Levy. "Continuous-time stochastic games of fixed duration." In: *Dynamic Games and Applications* 3.2 (2013), pp. 279–312.

[108] Tien-Yien Li and James A Yorke. "Period three implies chaos." In: *The theory of chaotic attractors*. Springer, 2004, pp. 77–84.

[109] Stefano Lovo and Tristan Tomala. "Markov perfect equilibria in stochastic revision games." In: *HEC Paris Research Paper No. ECO/SCD-2015-1093* (2015).

[110] Thodoris Lykouris, Vasilis Syrgkanis, and Éva Tardos. "Learning and efficiency in games with dynamic population." In: *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics. 2016, pp. 120–129.

[111] Hosam Mahmoud. *Pólya urn models*. CRC press, 2008.

[112] Antoine Mandel and Xavier Venel. "Dynamic competition over social networks." In: *European Journal of Operational Research* 280.2 (2020), pp. 597–608.

[113] Jean-Francois Mertens and Shmuel Zamir. "Formulation of Bayesian analysis for games with incomplete information." In: *International Journal of Game Theory* 14.1 (1985), pp. 1–29.

[114] Igal Milchtaich. "Network topology and the efficiency of equilibrium." In: *Games Econom. Behav.* 57.2 (2006), pp. 321–346. ISSN: 0899-8256.

[115] Bojan Mohar. "Eigenvalues, diameter, and mean distance in graphs." In: *Graphs and combinatorics* 7.1 (1991), pp. 53–64.

[116] Dov Monderer and Lloyd S. Shapley. "Potential games." In: *Games Econom. Behav.* 14.1 (1996), pp. 124–143. ISSN: 0899-8256.

[117] Elchanan Mossel, Manuel Mueller-Frank, Allan Sly, and Omer Tamuz. "Social learning equilibria." In: *Econometrica* 88.3 (2020), pp. 1235–1267.

[118] Elchanan Mossel and Omer Tamuz. "Opinion exchange dynamics." In: *Probability Surveys* 14 (2017), pp. 155–204.

[119] John F Nash et al. "Equilibrium points in n-person games." In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49.

[120] Anna Maria Paganoni and Piercesare Secchi. "Interacting reinforced-urn systems." In: *Advances in applied probability* 36.3 (2004), pp. 791–804.

[121] Mallesh M Pai, Aaron Roth, and Jonathan Ullman. "An antifolk theorem for large repeated games." In: *ACM Transactions on Economics and Computation (TEAC)* 5.2 (2016), pp. 1–20.

[122] Christos Papadimitriou. "Algorithms, games, and the internet." In: *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. 2001, pp. 749–753.

[123] Rohit Parikh and Paul Krasucki. "Communication, consensus, and knowledge." In: *Journal of Economic Theory* 52.1 (1990), pp. 178–189.

[124] Robin Pemantle. "A survey of random processes with reinforcement." In: *Probability surveys* 4 (2007), pp. 1–79.

[125] Ron Peretz, Gideon Amir, Itai Arieli, and Galit Ashkeazi-Golan. "Robust naive learning in social networks." In: *Working Paper* (2021).

[126] A. C. Pigou. *The Economics of Welfare*. 1st. London: Macmillan and Co., 1920.

[127] John Riordan and C. E. Shannon. "The number of two-terminal series-parallel networks." In: *J. Math. Phys. Mass. Inst. Tech.* 21 (1942), pp. 83–93. ISSN: 0097-1421.

[128] Herbert Robbins and Sutton Monro. "A stochastic approximation method." In: *The annals of mathematical statistics* (1951), pp. 400–407.

[129] Julia Robinson. "An iterative method of solving a game." In: *Annals of mathematics* (1951), pp. 296–301.

[130] Dinah Rosenberg, Eilon Solan, and Nicolas Vieille. "Informational externalities and emergence of consensus." In: *Games and Economic Behavior* 66.2 (2009), pp. 979–994.

[131] Robert W. Rosenthal. "A class of games possessing pure-strategy Nash equilibria." In: *Internat. J. Game Theory* 2 (1973), pp. 65–67. ISSN: 0020-7276.

[132] Tim Roughgarden. "The price of anarchy is independent of the network topology." In: *Journal of Computer and System Sciences* 67.2 (2003), pp. 341–364.

[133] Tim Roughgarden. "The price of anarchy in games of incomplete information." In: *ACM Trans. Econ. Comput.* 3.1 (2015), Art. 6, 20. ISSN: 2167-8375.

[134] Tim Roughgarden. *Twenty lectures on algorithmic game theory*. Cambridge University Press, 2016.

[135] Tim Roughgarden and Éva Tardos. "How bad is selfish routing?" In: *Journal of the ACM (JACM)* 49.2 (2002), pp. 236–259.

[136] Ariel Rubinstein. "Equilibrium in supergames (1994 re-ed.)" In: *Essays in Game Theory*. Springer, 1977, pp. 17–27.

[137] Ariel Rubinstein and Asher Wolinsky. "Rationalizable conjectural equilibrium: between Nash and rationalizability." In: *Games Econom. Behav.* 6.2 (1994), pp. 299–311. ISSN: 0899-8256.

[138] Hamid Sabourian. "Anonymous repeated games with a large number of players and random outcomes." In: *Journal of Economic Theory* 51.1 (1990), pp. 92–110.

[139] Paul A Samuelson. "A note on measurement of utility." In: *The review of economic studies* 4.2 (1937), pp. 155–161.

[140] William H. Sandholm. "Potential games with continuous player sets." In: *J. Econom. Theory* 97.1 (2001), pp. 81–108. ISSN: 0022-0531.

[141] Leonard J Savage. *The foundations of statistics*. New York, Wiley, 1954.

[142] Marco Scarsini and Tristan Tomala. "Repeated congestion games with bounded rationality." In: *Internat. J. Game Theory* 41.3 (2012), pp. 651–669. ISSN: 0020-7276.

[143] Lloyd S Shapley. "Stochastic games." In: *Proceedings of the national academy of sciences* 39.10 (1953), pp. 1095–1100.

[144] Lones Smith and Peter Sørensen. "Pathological outcomes of observational learning." In: *Econometrica* 68.2 (2000), pp. 371–398. ISSN: 0012-9682.

[145] Lones Smith and Peter Sørensen. "Pathological outcomes of observational learning." In: *Econometrica* 68.2 (2000), pp. 371–398.

[146] Sylvain Sorin. "On repeated games with complete information." In: *Mathematics of Operations Research* 11.1 (1986), pp. 147–160.

[147] Nicholas Stern. *The economics of climate change: the Stern review*. cambridge University press, 2007.

[148] Robert Henry Strotz. "Myopia and inconsistency in dynamic utility maximization." In: *The review of economic studies* 23.3 (1955), pp. 165–180.

[149] Vasilis Syrgkanis. "Bayesian games and the smoothness framework." arXiv:1203.5155. 2012.

[150] Steve Tesich. "A Government of Lies." In: *The Nation* (1992).

[151] Xavier Vives. "The speed of information revelation in a financial market mechanism." In: *Journal of Economic Theory* 67.1 (1995), pp. 178–204.

[152] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton university press, 1944.

[153] Chenlan Wang, Xuan Vinh Doan, and Bo Chen. "Price of anarchy for non-atomic congestion games with stochastic demands." In: *Transportation Res. Part B* 70 (2014), pp. 90–111. ISSN: 0191-2615.

[154] John Glen Wardrop. "Some theoretical aspects of road traffic research." In: *Proceedings of the Institute of Civil Engineers, Part II*. Vol. 1. 1952, pp. 325–378.

[155] Martin L Weitzman. "Gamma discounting." In: *American Economic Review* 91.1 (2001), pp. 260–271.

[156] Robin J. Wilson. *Introduction to Graph Theory (5th edition)*. Pearson, 2010.

[157] Thomas Wiseman. "A partial folk theorem for games with unknown payoff distributions." In: *Econometrica* 73.2 (2005), pp. 629–645.

[158] Manxi Wu and Saurabh Amin. "Learning an unknown network state in routing games." In: *IFAC-PapersOnLine* 52.20 (2019), pp. 345–350.

[159] Manxi Wu, Saurabh Amin, and Asuman E. Ozdaglar. "Value of information in Bayesian routing games." In: *Oper. Res.* 69.1 (2021), pp. 148–163.

[160] Manxi Wu, Jeffrey Liu, and Saurabh Amin. "Informational aspects in a class of Bayesian congestion games." In: *2017 American Control Conference (ACC)*. IEEE. 2017, pp. 3650–3657.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Final Version as of January 11, 2022 (classicthesis v4.6).

Hermann Zapf's *Palatino* and *Euler* type faces (Type 1 PostScript fonts *URW Palladio L* and *FPL*) are used. The "typewriter" text is typeset in *Bera Mono*, originally developed by Bitstream, Inc. as "Bitstream Vera". (Type 1 PostScript fonts were made available by Malte Rosenau and Ulrich Dirr.)

Any unhappy deviation from these environment is due to the author's improving yet incomplete ability in mediating clashes between so many packages. Happy deviations are assumed to be entirely intentional.

ECOLE DOCTORALE

Titre : Information, Coordination et Coopération: Essais sur l'Apprentissage dans les Jeux Dynamiques

Mots clés : Jeux dynamiques, jeux de routage, jeux de révision, information incomplète, apprentissage social.

Résumé : Cette thèse étudie les dynamiques d'apprentissage dans les jeux dynamiques. Elle comporte trois parties indépendantes. Dans la première, j'étudie les modèles non-bayésiens de formation d'opinion dans un réseau. Ces modèles considèrent un ensemble d'individus structuré en réseau échangeant en temps discret leur croyance sur un état du monde. Je propose une approche stochastique fondée sur l'apprentissage par renforcement. J'analyse les propriétés de convergence et l'émergence d'un consensus au sein de la population. Dans la seconde partie, j'étudie les dynamiques d'apprentissage social dans les jeux de routage en information publique incomplète. En temps discret, un continuum d'agents traverse un réseau orienté d'un nœud origine à un nœud destination. Le temps de trajet, ou latence, de chaque arête est modélisé par une fonction dépendant de la masse d'agents utilisant l'arête ainsi que d'un paramètre global inconnu. Je caractérise les conditions nécessaires et suffisantes pour que soit les croyances des individus convergent vers la vérité, soit à la limite ils jouent à chaque étape un équilibre du jeu en information complète. Dans la dernière partie, je propose un modèle théorique d'économie environnementale. J'étudie comment deux états en concurrence économique arbitrent entre un flux de paiements de court-terme et un paiement dépendant de l'état du monde à long-terme lorsqu'ils décident de s'engager, ou non, dans une transition écologique. Ce modèle propose une approche transverse mêlant des éléments des jeux de révision et des jeux à paiements dépendant de la fréquence. Dans ce contexte, je tente de caractériser les stratégies optimales et d'identifier les conditions permettant l'émergence de comportements coopératifs.

Title : Information, Coordination and Cooperation: Essays on Learning in Dynamic Games.

Keywords : Dynamic games, routing games, revision games, incomplete information, social learning.

Abstract : This thesis studies leaning patterns in dynamic games. It consists in three independent parts. In the first one, I study non-Bayesian models of opinion formation in networks. These models consider a set of agents embedded in a network structure exchanging their opinion on some state of the world at discrete times. I propose a stochastic extension of non-Bayesian learning based on reinforcement learning models. I analyze convergence properties and characterize conditions that ensure the emergence of a consensus among a population. In the second part, I study the dynamics of social learning in routing games with incomplete public information. At discrete times, continua of agents route through a network from an origin node to a destination node. The travel time, or latency, of each edge is modeled as a function depending both on the mass of agents using that edge and a global unknown state parameter. I characterize necessary and sufficient conditions so that, in the limit, either the public belief converges to the truth or agents play an equilibrium of the full-information game at each stage. In the last part, I propose a theoretical model of environmental economics. I study how two competing states balance between a short-term payment flow and a long-run payoff depending on the state of the world when they can choose between engaging in environmental transition or keeping the status quo. This model proposes an approach combining elements from revision games and games with frequency-dependent payoffs. In this context, I aim at characterizing optimal strategies and identifying conditions that allow the emergence of cooperative behavior.

Institut Polytechnique de Paris
91120 Palaiseau, France