

Université de Paris

École doctorale de Sciences Mathématiques de Paris Centre (ED 386)

Institut de Recherche en Informatique Fondamentale

Algorithmic Game Theory Applied to Networks and Populations

par Simon Collet

Thèse de doctorat en Informatique

Dirigée par Pierre Fraigniaud
et Amos Korman

Présentée et soutenue publiquement le 9 décembre 2019

Devant un jury composé de :

Christoph	Dürr	Directeur de recherche	Sorbonne Université	rapporteur
Pierre	Fraigniaud	Directeur de recherche	Université de Paris	directeur
Amos	Korman	Directeur de recherche	Université de Paris	directeur
Sophie	Laplante	Professeure	Université de Paris	examinatrice
Giuseppe	Persiano	Professeur	Università di Salerno	rapporteur
Celine	Scornavacca	Chargée de recherche	Université de Montpellier	examinatrice
Yannick	Viossat	Maître de conférences	Université Paris-Dauphine	examineur



Contents

1	Introduction	1
1.1	Les différents courants de la théorie des jeux	1
1.1.1	L'ère classique	1
1.1.2	La théorie des jeux évolutionnaires	3
1.1.3	La théorie algorithmique des jeux	4
1.2	Contributions de la thèse	6
2	Equilibria of Games in Networks for Local Tasks	8
2.1	Introduction	8
2.1.1	Motivation and Objective	8
2.1.2	Our Results	13
2.1.3	Related Work	14
2.2	The Extensive Games Related to LCL Games	16
2.2.1	Basic Definitions	16
2.2.2	Well-Rounded Games	17
2.2.3	Strategies, Outcomes, and Expected Payoff	18
2.2.4	Equilibria, and Subgame Perfection	20
2.2.5	Metrics	20
2.3	Proof of Lemma 2.2	21
2.3.1	Perturbed Games and Perfect Equilibria	21
2.3.2	Truncated Games and Induced Equilibria	23
2.3.3	Truncation of Perturbed Games	25
2.3.4	The proof of Lemma 2.2	27
2.4	Proof of Lemma 2.3	28
2.4.1	Formal Definition of LCL Games	28
2.4.2	The proof of Lemma 2.3	29
2.5	Conclusion and Further Work	31
3	The impact of Competition on Coverage in a Dispersal Game	32
3.1	Introduction	33
3.1.1	The Dispersal Game	33
3.1.2	Evaluating Policies	35
3.1.3	Ideal Free Distribution (IFD)	36
3.1.4	Evolutionary Stable Strategy (ESS)	37

3.1.5	Our Results	38
3.1.6	Related Work	39
3.2	The Exclusive Reward Policy	40
3.2.1	Algorithm σ^*	40
3.2.2	σ^* has Optimal Coverage	42
3.3	The Ideal Free Distribution is an ESS	43
3.4	The Criticality of the Exclusive Collision Cost Function	46
3.5	Conclusion	47
3.5.1	Discussion and Future Directions	47
3.5.2	Informal Discussion on Animal Dispersal	48
4	Foraging of Leptos Bats : Models and Simulations	51
4.1	Introduction	51
4.1.1	Field Experiments	51
4.1.2	Computer Simulations	52
4.1.3	Theoretical model	52
4.2	Field Experiments	53
4.2.1	Description of the experiments	53
4.2.2	Qualitative results	54
4.3	Simulation Results	55
4.3.1	Experimental Framework	55
4.3.2	Simulations	57
4.3.3	Outcome of the Simulations	59
4.4	Analytical Results	67
4.4.1	Setting and notations	67
4.4.2	Model with complete information	68
4.4.3	Model with incomplete information	73
4.4.4	Uniform distribution	79
4.5	Conclusion and future research directions	87
5	Conclusion	89

Résumé

Cette thèse se propose d'étudier sous l'angle de la théorie algorithmique des jeux divers jeux inspirés de situations réelles rencontrées dans les domaines des réseaux informatiques et de la biologie. Ces jeux se caractérisent par un grand nombre de joueurs disposant chacun d'une information incomplète à propos des autres joueurs. Dans la théorie classique, ces jeux rentrent dans la catégorie des jeux extensifs à information imparfaite et sont modélisés à l'aide d'arbres. Ils restent toutefois très difficiles à analyser en détail, à cause de leur inhérente complexité, liée notamment à une profondeur d'arbre potentiellement infinie. Nous avons relevé le défi de cette tâche en diversifiant les méthodes de résolution et en mettant l'accent sur son aspect interdisciplinaire.

Cette thèse, outre l'introduction et la conclusion, se divise en trois parties. Dans la première, nous adoptons le point de vue de la théorie des jeux classique. Nous proposons un modèle de jeu qui correspond à une large classe de problèmes rencontrés dans la théorie du calcul distribué. Les principales contributions de cette partie sont, d'une part, de montrer comment passer d'un point de vue purement algorithmique du problème à un point de vue correspondant à la théorie des jeux, et, d'autre part, de prouver l'existence d'équilibres satisfaisants pour la classe de jeux obtenue. Ce deuxième point est essentiel et garantit que la théorie des jeux est adaptée à l'étude de tels jeux distribués, malgré leur complexité.

La seconde partie est consacrée à l'étude d'un jeu omniprésent au sein des systèmes biologiques que l'on nomme jeu de la dispersion. Il modélise la situation dans laquelle un groupe d'animaux doit se partager une certaine quantité de ressources répartie entre différents sites géographiques. La difficulté du jeu provient du fait que certains sites contiennent plus de ressources que d'autres, mais risquent également d'attirer plus de joueurs. Le partage est alors difficile. Nous proposons une règle de répartition des ressources qui permet de maximiser les ressources exploitées par l'ensemble du groupe. Cette partie est aussi l'occasion de revisiter les liens étroits entre le concept de distribution libre idéale, très présente dans la théorie du fourrageage, et le concept de stratégie évolutionnairement stable, concept clé de la théorie des jeux évolutionnaires.

La troisième partie se concentre sur l'étude du comportement d'une certaine espèce de petites chauve-souris vivant au Mexique, dans le désert de Sonora, et se nourrissant la nuit du nectar des cactus géant de l'espèce protégée Saguaro. Après la présentation des résultats expérimentaux obtenus sur le terrain, nous proposons une simulation informatique de leur comportement. Les résultats de ces simulations permettent de formuler des hypothèses intéressantes à propos du fonctionnement cérébral de ces petits mammifères. Nous étudions ensuite un modèle théorique de jeu inspiré de cette situation réelle. L'étude de ce modèle abstrait nous permet de distinguer les caractéristiques fondamentales du jeu, et de renforcer notre approche de théorisation du comportement de fourrageage. Cette étude ouvre également la voie à l'application de ce type de modèle à d'autres situations, impliquant un comportement animal ou humain.

Abstract

The aim of this thesis is to use algorithmic game theory tools to study various games inspired from real world problems in the fields of information networks and biology. These games are characterized by a large number of players, each with incomplete information about other players. In classical game theory, these games fall into the category of extensive games with imperfect information, which are modeled using trees. However, these games remain very difficult to analyze in details, because of their intrinsic complexity, which is linked with their possibly infinite tree depth. Nevertheless, we have taken up the challenge of this task, while diversifying the methods of resolution, and emphasizing its interdisciplinary aspect.

Besides the introduction and the conclusion, the thesis is divided into three parts. In the first part, we adopt the point of view of classical game theory. We propose a game that corresponds to a wide class of problems encountered in the theory of distributed computing. The main contributions of this part are, on the one hand, to show how to transform a purely algorithmic problem into a game, and, on the other hand, to prove the existence of satisfactory equilibria for the resulting class of games. This second point is essential, as it guarantees that game theory is adapted to the study of distributed games, despite their complexity.

The second part is dedicated to the study of a game omnipresent within biological systems, that we call dispersion game. This game models the situation in which a group of animals must share a certain amount of resources scattered among different geographical sites. The difficulty of the game comes from the fact that some sites contain more resources than others, but may also attract more players. We propose a rule for the distribution of resources which makes it possible to maximize the resources exploited by the whole group. This part of the thesis is also an opportunity to revisit the close links between the concept of ideal free distribution, very present in the theory of foraging, and the concept of evolutionarily stable strategy, a key concept of evolutionary game theory.

The third part focuses on the study of the behavior of a specific species of small bats living in Mexico, in the Sonoran Desert, and feeding at night from the nectar of the giant Saguaro cacti, a protected species. After the presentation of the experimental results obtained in the field, we propose a computer simulation of their behavior. The results of these simulations make it possible to formulate interesting hypotheses about the cerebral activities of these small mammals. We then study a theoretical model of game inspired by this real situation. The study of this abstract model allows us to distinguish the fundamental characteristics of the game, and to reinforce our approach of theorizing foraging behavior. This study also opens the way to applying this type of model to other situations, involving animal or human behavior.

Chapitre 1

Introduction

La théorie des jeux peut être brièvement décrite comme l'étude des interactions stratégiques entre plusieurs joueurs indépendants qui poursuivent chacun un objectif propre. Plus précisément, elle fournit un cadre mathématique permettant de modéliser de telles situations. Les termes de “jeu” et de “joueurs” font ici référence à l'aspect *stratégique* des situations modélisées, et non à leur aspect *ludique*. En effet, il peut s'agir de situations aussi variées que les interactions entre entreprises ayant le même secteur d'activités, entre automobilistes conduisant sur la même route, entre équipes de foot ou joueurs d'échecs participant à un tournoi, entre animaux se partageant un même territoire ou une même source de nourriture, voire les interactions entre des cellules cancéreuses et des cellules saines dans un organisme vivant. On comprend aisément comment une modélisation mathématique de situations aussi courantes peut avoir des applications dans des domaines aussi variés que l'économie, la sociologie, la psychologie, la biologie, la médecine, ainsi que bien sûr l'informatique. Il est aussi important de souligner que ces modélisations ne se limitent pas aux aspects *compétitifs* de ces interactions, mais englobent également leurs aspects *coopératifs*.

1.1 Les différents courants de la théorie des jeux

1.1.1 L'ère classique

Commençons par un petit tour d'horizon historique. Quelques problèmes relevant de la théorie des jeux ont été formulés, et plus rarement résolus, dès l'antiquité (on en trouve des traces dans certains dialogues de Platon), mais ce n'est qu'au début du 20^e siècle que l'on peut réellement parler d'une théorie mathématique des jeux. L'ouvrage *Théorie des jeux et comportements économiques* [94], publié en 1944 par le mathématicien John von Neumann et l'économiste Oskar Morgenstern, est souvent considéré comme l'ouvrage fondateur du domaine. Il propose une définition du concept de jeu suffisamment générale pour avoir des applications pratiques, ainsi qu'une façon de les résoudre assez intuitive et compatible avec l'expérience de joueurs professionnels d'échecs, par exemple. L'aspect le plus attrayant de cette théorie est le célèbre *théorème du minimax* qui nous assure que tous les jeux finis à deux joueurs et à somme nulle possèdent une solution. Toutefois la théorie se limitait alors aux jeux à deux joueurs et à somme nulle (que l'on devrait plutôt appeler jeux à somme constante puisqu'il s'agit précisément des jeux pour lesquels la

somme des gains des joueurs est constante et indépendante de leurs actions). Un tournoi dans lequel une victoire rapporte un point, une défaite zéro et un match nul un demi-point, est un exemple de jeu à somme nulle. De manière plus générale, tout jeu au cours duquel une quantité fixe de ressources (comme un gâteau par exemple) est partagée entre les joueurs est un jeu à somme nulle.

La théorie va s'étoffer considérablement en 1950 lorsque John Nash développe dans sa thèse de doctorat [68–70] la notion d'équilibre, qui sera dès ce moment appelée par la communauté l'*équilibre de Nash*. En quelques mots, il y a équilibre lorsque aucun joueur ne peut augmenter son gain en changeant de stratégie. Autrement dit, la stratégie de chaque joueur est optimale étant données les stratégies des autres joueurs. Cette notion, pourtant moins intuitive que les notions constructives déjà existantes, telles que le minimax ou le processus d'élimination des stratégies dominées, va s'imposer grâce à sa généralité et grâce au célèbre théorème de Nash : chaque jeu fini possède un équilibre, y compris les jeux à somme non-constante, et ce quelque soit le nombre de joueurs. De plus, lorsque le jeu est à deux joueurs et à somme nulle, l'équilibre de Nash coïncide avec la solution du minimax. Dès lors, grâce à ce théorème, la théorie des jeux va prendre son essor dès les années 50, et va rapidement trouver de nombreuses applications. Elle va notamment être utilisée par l'ONG *RAND corporation*¹ pour mettre au point les doctrines stratégiques de défense pendant la guerre froide comme la MAD : Destruction Mutuelle Assurée.

Du point de vue de la théorie, les principaux progrès réalisés depuis les années 50 jusqu'à la fin des années 70 consistent en des généralisations du concept de jeu, et en des raffinements du concept d'équilibre. En effet, le théorème de Nash garantit l'existence d'au moins un équilibre dans tout jeu fini, mais pas son unicité. Il est alors utile de définir des concepts d'équilibre plus restrictifs afin de résoudre les jeux possédant plusieurs équilibres de Nash, mais avec des espaces de solutions assez larges pour garantir un théorème d'existence. Ces recherches de généralisation et de raffinement vont de pair, car une extension du concept de jeu n'a d'intérêt que si l'on dispose d'un concept d'équilibre applicable à ce jeu et dont on peut prouver l'existence. Par ailleurs, plus les jeux se complexifient, plus il est probable qu'ils possèdent de nombreux équilibres. Ainsi, en 1965, Reinhard Selten définit la notion d'*équilibre parfait en sous-jeux* [85], raffinement de l'équilibre de Nash particulièrement adapté à l'étude des jeux sous forme extensive, c'est-à-dire les jeux au cours desquels les choix d'un joueur à un moment donné peuvent dépendre des choix précédents faits par chaque joueur ; comme les échecs, le go, et la majeure partie des jeux de société courants.

Par ailleurs, le concept d'*information* fait son apparition dans la théorie des jeux, et l'on distingue désormais les jeux à information parfaite (tels que les échecs ou le go) des jeux à information imparfaite (tels que les jeux de cartes où la main des adversaires nous est inconnue). Pour résoudre les jeux à information imparfaite, il est nécessaire d'inventer de nouveaux concepts d'équilibre. En 1975, Reinhard Selten propose l'*équilibre parfait en main tremblante* [86], qui tire son nom de l'idée selon laquelle un bon équilibre doit être stable même si les joueurs sont susceptibles de commettre des erreurs, comme si leurs

¹ONG américaine fondée en 1948 dans le but d'améliorer le processus décisionnel et stratégique de l'armée, toujours en activité.

maines tremblaient au moment de déplacer une pièce d'échecs. Sept ans plus tard, l'*équilibre séquentiel*, proposé par Kreps et Wilson en 1982 [54], introduit la notion de "croyance" des joueurs dans l'analyse des jeux, une notion qui permet d'éclairer de nombreux mécanismes subtils apparaissant dans les jeux à information imparfaite.

Malgré ces progrès dans la formalisation mathématique de la théorie, sa capacité à modéliser fidèlement des comportements humains sera progressivement mise en doute aux alentours des années 70. En effet, durant cette période, des psychologues tels que Daniel Kahneman s'emparent de la théorie et décident de la tester expérimentalement, ce qui donnera naissance aux champs de l'économie expérimentale et de l'économie comportementale. Les expériences réalisées donnent rapidement naissance à de nombreux paradoxes, en particulier certains jeux, pourtant assez simples, ont une solution mathématique à l'opposé de la solution intuitive adoptée par la plupart des humains testés. Le modèle de l'*homo economicus* est dépassé, il s'agit désormais de considérer les motivations humaines dans toute leur complexité, et d'identifier les différents biais qui faussent notre jugement.

1.1.2 La théorie des jeux évolutionnaires

C'est dans ce contexte de diversification des champs d'application de la théorie des jeux que les biologistes s'en emparent à leur tour pour créer ce qui deviendra la théorie des jeux *évolutionnaires*. À l'origine de cette théorie, on cite habituellement l'article de John Maynard Smith et George R. Price de 1973 intitulé *La logique des conflits entre animaux* [87], ainsi que le livre *L'évolution et la théorie des jeux* [88], publié en 1982 par John Maynard Smith.

Le pari de cette nouvelle branche de la théorie peut sembler loufoque au premier abord, puisqu'il s'agit d'appliquer à des animaux, des végétaux ou même des cellules, une théorie dont l'une des hypothèses est que les joueurs sont parfaitement rationnels. Néanmoins, il s'avère que la parfaite rationalité n'est pas la seule façon d'atteindre l'équilibre, et que le processus de l'évolution des espèces le permet lui aussi. Un nouveau modèle de jeu répété dans une population infinie, dans lequel les stratégies sont héritées de génération en génération, et altérées par des mutations, est inventé. Pour ce modèle, la notion de *stratégie évolutivement stable* (que l'on appelle parfois *état évolutivement stable*) est définie comme un état de la population dans lequel toute stratégie émergente, initialement en faible proportion (comme c'est le cas pour une stratégie issue d'une mutation), est vouée à disparaître. Le résultat surprenant de ce modèle est que toute stratégie évolutivement stable est un équilibre de Nash, c'est-à-dire que le processus de l'évolution permet en général d'atteindre les mêmes équilibres que le raisonnement déductif. Il est également intéressant de noter que ce nouveau modèle fait la part belle aux jeux symétriques : une stratégie évolutivement stable étant typiquement une stratégie symétrique d'un jeu symétrique.

Cependant, cette nouvelle notion d'équilibre n'est pas la seule contribution de la branche évolutionnaire de la théorie des jeux. Elle apporte également un éclairage sur les aspects dynamiques des jeux. En effet, les systèmes dynamiques modélisant des populations animales étaient déjà étudiés en biologie, et possédaient leurs propres notions d'équilibre : il s'agit des attracteurs, ensembles stables ou asymptotiquement stables. Une série de

résultats faisant partie du folklore de la théorie des jeux évolutionnaires établissent des liens, parfois subtils, entre les notions de stabilité d'un système dynamique et les notions d'équilibre du jeu correspondant [89], cette tâche étant rendue difficile par la diversité des modèles dynamiques que l'on peut associer à un même jeu. De l'étude des systèmes dynamiques, on importe également l'idée que certains attracteurs ou équilibres ne sont pas nécessairement ponctuels, ce qui donne naissance à la notion d'*ensemble évolutivement stable* [90] : un ensemble de stratégies telles que, si les joueurs choisissent chacun leur stratégie dans cet ensemble, toute stratégie émergente ne faisant pas partie de cet ensemble est vouée à disparaître.

Au fur et à mesure que les jeux évolutionnaires et leurs nombreuses variantes sont étudiés, on s'intéresse de plus en plus à des modèles de population possédant une dimension spatiale. Dans ces jeux, chaque joueur est en interaction avec un petit nombre de voisins, plutôt qu'avec l'ensemble de la population, ce qui amène rapidement la discipline à emprunter des idées à la théorie des graphes. Par ailleurs, dans le contexte des jeux à information imparfaite, les modèles étudiés étant de plus en plus complexes, il devient naturel de penser une stratégie comme un algorithme, plutôt que simplement comme une suite de décisions prédéterminées. Ces deux exemples illustrent l'intérêt croissant que les théoriciens des jeux vont porter au champ de l'informatique fondamentale, au point de créer, à la fin des années 90, une discipline à part entière.

1.1.3 La théorie algorithmique des jeux

Le mouvement de convergence entre théorie des jeux et informatique s'est effectué dans les deux sens. De nombreux informaticiens se sont en effet naturellement tournés vers la théorie des jeux pour résoudre certains problèmes, notamment en algorithmique parallèle ou distribuée, en théorie de la complexité de communication, en cryptographie et en intelligence artificielle. Théorie des jeux et informatique, deux domaines dont John von Neumann fut un pionnier, ont alors donné naissance à la *théorie algorithmique des jeux*. Comme pour de nombreux domaines à l'interface de plusieurs disciplines, il est difficile de donner une image complète de la théorie algorithmique des jeux, tant elle est foisonnante d'idées et de concepts divers. Néanmoins on peut distinguer trois contributions majeures à la théorie [72] : l'étude de la *complexité* des équilibres, la notion de *prix de l'anarchie*, et la théorie des *mécanismes d'incitation* (mechanism design).

La question de la complexité du calcul d'un équilibre est naturelle pour un informaticien. Dans le cadre de la théorie des jeux à deux joueurs et à somme nulle, le théorème du minimax permet de concevoir un algorithme calculant une solution du jeu. Il est appelé *algorithme du minimax*, mais on utilise plus souvent l'*algorithme alpha-bêta*, qui en est une version optimisée. C'est un algorithme polynômial vis-à-vis du nombre de stratégies de chaque joueur, mais exponentiel vis-à-vis du nombre de tours de jeux dans le cas des jeux sous forme extensive. Ceci est assez intuitif lorsqu'on considère l'arbre des positions atteignables du jeu, dans lequel chaque noeud a pour degré le nombre d'actions autorisées du joueur dans cette position. L'algorithme alpha-bêta est une composante principale des intelligences artificielles traditionnelles tels que *Deep Blue* en 1997, et est toujours présent (dans une version très optimisée et adaptée) dans les intelligences artificielles modernes

telles que *AlphaGo* en 2015.

Dans le cadre des jeux à somme non constante ou à au moins 3 joueurs, il s'agit d'évaluer la complexité du calcul de l'équilibre de Nash, ce qui s'avère plus complexe. La méthode la plus utilisée en pratique pour trouver des équilibres de Nash est l'algorithme de Lemke-Howson [57], qui est exponentiel dans le pire cas. Lorsque la question de la complexité du calcul d'un équilibre de Nash s'est posée à la fin des années 90, une des premières remarques qui ont été faites est que la formulation du problème sous forme de problème de décision : "est-ce que tel jeu fini possède un équilibre de Nash ?", n'avait pas d'intérêt, puisque le théorème de Nash garantit qu'il en existe toujours au moins un. En revanche, on peut montrer que la plupart des problèmes dérivés, tels que "existe-t-il au moins deux équilibres de Nash ?" ou "existe-t-il un équilibre de Nash pour lequel une fonction de bien-être social est supérieure à telle valeur ?" sont NP-complets. Le premier problème, quant à lui, est PPAD-complet [18, 23], la classe PPAD étant une sous-classe de NP [74].

Dans une approche comparable, une autre quantité que l'on a cherché à quantifier est l'inefficacité des équilibres. Dans ce but, Koutsoupias et Papadimitriou introduisent en 1999 le concept de *prix de l'anarchie* [52]. Il s'agit simplement de comparer les valeurs de différents équilibres d'un même jeu en utilisant une fonction de bien-être social qui quantifie globalement le caractère désirable d'une situation donnée. Le prix de l'anarchie est alors simplement la valeur optimale de cette fonction divisée par la valeur du plus mauvais équilibre. Ce concept simple, appliqué à de nombreux problèmes, permet de quantifier l'impact de l'égoïsme des agents sur l'efficacité d'un système. Il permet aussi de formuler simplement le problème suivant : "Comment faire baisser le prix de l'anarchie d'un jeu donné ?".

En effet, on sait depuis 1968 et la formulation par Braess de son célèbre paradoxe [10, 11], que la construction d'une nouvelle route dans un réseau routier peut conduire à l'augmentation du temps de trajet moyen pour chaque utilisateur. Mais l'inverse est également vrai : la suppression d'une route peut conduire à l'amélioration du temps de trajet moyen. L'étude des conditions dans lesquelles l'ajout ou la suppression d'une route peut provoquer l'augmentation ou la réduction du temps de trajet moyen peut être considéré comme un exemple d'application de la théorie des mécanismes d'incitation. D'une manière plus générale, il s'agit d'étudier, dans un cadre bien défini, l'impact de petites modifications d'un jeu sur l'ensemble de ses équilibres, en particulier sur son prix de l'anarchie. L'objectif étant bien entendu de faire baisser le prix de l'anarchie à faible coût. Cette idée a été appliquée notamment à des modèles d'enchères, d'élections, de marchés et de réseaux.

Parmi les résultats les plus importants de la théorie des mécanismes d'incitation, on peut citer les théorèmes d'Arrow, de Gibbard-Satterthwaite et de Gibbard, qui explicitent les limitations des systèmes de vote. Mais aussi des résultats plus positifs, tels que le mécanisme de Vickrey-Clarke-Groves qui permet, sous des hypothèses plus fortes mais raisonnables, de résoudre tout problème de choix social, de manière socialement optimale (au sens de l'utilitarisme) et sans coût additionnel. Son principe réside dans le fait de compenser les externalités de chaque joueur par des paiements. Un cas particulier bien

connu de ce mécanisme est l'enchère VCG : chaque joueur annonce (simultanément ou anonymement) le prix qu'il est prêt à payer pour obtenir l'objet vendu, et la personne qui annonce le prix le plus élevé se voit attribuer l'objet et doit payer le second prix le plus élevé annoncé. Si cette enchère est implémentée, aucun joueur n'a intérêt à annoncer un prix différent de la valeur réelle qu'il attribue à l'objet.

Outre le concept de prix de l'anarchie, l'approche algorithmique a apporté à la théorie des mécanismes d'incitation des considérations sur la combinatoire et la complexité des mécanismes ainsi que de nouveaux modèles de jeux inspirés de graphes ou d'algorithmes "online".

1.2 Contributions de la thèse

L'objectif principal de cette thèse est de combiner des aspects évolutionnaires et algorithmiques de la théorie des jeux afin d'étudier des environnements fortement compétitifs avec de nombreux joueurs, tels que des populations animales ou des réseaux informatiques à l'ère d'internet. De nombreux auteurs se sont déjà penchés sur des problèmes similaires, il est donc important de préciser en quoi notre approche se distingue des précédentes. La plupart des jeux en réseau (network games) analysés dans la littérature le sont grâce à une ou plusieurs hypothèses simplificatrices, c'est le cas par exemple des jeux à potentiel [64], pour lesquels il existe une fonction de potentiel permettant de résumer toutes les caractéristiques du jeu. D'autres hypothèses classiques incluent : supposer qu'il existe un joueur "central" qui possède une information parfaite sur l'état du réseau, supposer que chaque joueur possède une information parfaite sur le réseau, limiter l'ensemble des stratégies à un cadre bien défini mais assez restreint, etc. Dans cette thèse, nous avons essayé d'aborder le problème dans toute sa complexité, c'est-à-dire en proposant des modèles dans lesquels chacun des multiples joueurs possède une information imparfaite du jeu et des autres joueurs. La difficulté de la tâche nous a poussé à tenter des approches assez diverses, comme en témoignent les trois chapitres qui suivent cette introduction.

Dans le second chapitre, nous introduisons une nouvelle classe de jeux modélisant l'exécution d'algorithmes distribués dans un large réseau. L'idée de ce modèle provient de l'observation suivante : lorsqu'un algorithme distribué est exécuté par un réseau (on se place typiquement dans le modèle *LOCAL*), on suppose implicitement que tous les agents exécutent fidèlement l'algorithme. Mais il est courant d'exécuter de tels algorithmes sur des réseaux dont les machines sont possédées par des propriétaires différents, n'ayant pas nécessairement les mêmes objectifs, et confrontés à des problématiques liées aux coûts d'exploitation et de fonctionnement. Est-il alors raisonnable de penser que chaque entité va exécuter fidèlement l'algorithme ou doit-on s'attendre à des déviations allant dans le sens des intérêts particuliers ? Et si l'on suppose que les agents ont des intérêts particuliers et agissent de manière égoïste, dans quelle mesure est-il possible de concevoir des algorithmes robustes ?

Le troisième chapitre est dédiée à l'analyse d'un jeu d'exploration à multiples joueurs. Le modèle est fortement inspiré de la problématique du fourrageage, c'est-à-dire la recherche de nourriture dans le monde animal. La situation typique est la suivante : un large groupe d'animaux de la même espèce doit se disperser au cours de la journée (ou de la nuit pour

les animaux nocturnes) pour trouver de la nourriture. Les animaux ont une connaissance de leur environnement, et savent à quels endroits ils ont le plus de chances de trouver de la nourriture, en revanche ils ne savent pas si d'autres compétiteurs occupent déjà les lieux. La situation d'équilibre est alors telle que, en moyenne, chaque lieu de fourrageage est visité par un nombre d'animaux proportionnel à la quantité de nourriture s'y trouvant. Les animaux se trouvant à un même lieu devant partager la nourriture, il s'ensuit qu'à l'équilibre chaque animal obtient approximativement la même quantité de nourriture. Cette situation est bien connue et étudiée en théorie de l'approvisionnement (foraging theory), et porte le nom de "distribution libre idéale" (Ideal Free Distribution) [35, 48]. Nous avons reformulé cette problématique sous un angle différent, inspiré de la théorie des mécanismes d'incitation. Plus précisément, nous nous sommes posé la question suivante : quel est l'impact du niveau d'agressivité des animaux les uns envers les autres sur les performances du groupe dans son ensemble ? Nous montrons qu'un niveau d'agressivité plus élevé a pour conséquence que les animaux ont plus tendance à s'éviter, ce qui résulte en une meilleure dispersion du groupe. Nous caractérisons ainsi le niveau d'agressivité qui maximise la dispersion du groupe, et sa capacité à exploiter au mieux les ressources de nourriture.

Dans le quatrième chapitre, nous étudions un cas particulier du jeu introduit dans le second. Nous nous intéressons donc au comportement de chauves-souris de l'espèce *Leptonycteris yerbabuenae*, que l'on appelle plus couramment *Leptos*. Ces chauves-souris pèsent une vingtaine de grammes pour une longueur d'environ 8 centimètres. Au printemps, elles occupent une grotte dans la région du désert de Sonora, au Mexique. Pendant cette période, elles se nourrissent principalement de nectar de fleurs de cactus au début de la saison, puis de fruits de ces mêmes cactus vers la fin de la saison. À la tombée de la nuit, les 100 000 habitants de la grotte la quittent et se répartissent dans un rayon d'environ 70 km pour se nourrir, puis retournent à la grotte avant le lever du soleil. Elles participent ainsi à un jeu de dispersion d'une grande ampleur. Nous avons étudié leur comportement en mêlant une approche plus théorique et des observations de terrain. Nous proposons donc un modèle théorique simple permettant de comprendre conceptuellement le problème que chaque individu doit résoudre. Ensuite nous élaborons une série de simulations permettant de comprendre la dynamique du groupe.

Chapter 2

Equilibria of Games in Networks for Local Tasks

2.1 Introduction

Distributed tasks such as constructing a maximal independent set (MIS) in a network, or properly coloring the nodes or the edges of a network with reasonably few colors, are known to admit efficient distributed randomized algorithms. Those algorithms essentially proceed according to some simple generic rules, by letting each node choosing a tentative value at random, and checking whether this choice is consistent with the choices of the nodes in its vicinity. If this is the case, then the node outputs the chosen value, else it repeats the same process. Although such algorithms are, with high probability, running in a polylogarithmic number of rounds, they are not robust against actions performed by rational but selfish nodes. Indeed, such nodes may prefer specific individual outputs over others, e.g., because the formers suit better with some individual constraints. For instance, a node may prefer not being placed in a MIS as it is not willing to serve as a relay node. Similarly, a node may prefer not being assigned some radio frequencies (i.e., colors) as these frequencies would interfere with other devices running at that node. In this chapter, we show that the probability distribution governing the choices of the output values in the generic algorithm can be tuned such that no nodes will rationally deviate from this distribution. More formally, and more generally, we prove that the large class of so-called LCL tasks, including MIS and coloring, admit simple “Luby’s style” algorithms where the probability distribution governing the individual choices of the output values forms a Nash equilibrium. In fact, we establish the existence of a stronger form of equilibria, called symmetric trembling-hand perfect equilibria for those games.

2.1.1 Motivation and Objective

In networks, independent sets and dominating sets can be used as backbones to collect, transfer, and broadcast information, and/or as cluster heads in clustering protocols (see, e.g., [55, 66]). Hence, a node belonging to some selected independent or dominating set may be subject to future costs in term of energy consumption, computational efforts, and bandwidth usage. As a consequence, rational selfish nodes might be tempted to deviate from the instructions of an algorithm used to construct such sets, so that to avoid becoming member of the independent set, or dominating set, under construction. On the

other hand, the absence of a backbone, or of cluster heads, may penalize the nodes. Hence every node is subject to a tension between (1) facilitating the *obtention* of a solution, and (2) avoiding certain *forms* of solutions.

A large class of randomized algorithms [6, 59] for constructing maximal independent sets (MIS) proceed in synchronous *rounds*, where a round allows every node to exchange information with its neighbors in the network, and to perform some individual computation. Roughly, at each round of these algorithms, every node i which has not yet decided applies to enter the MIS with a certain probability p_i . If a node applies to enter the MIS, and none of its neighbors simultaneously apply, then the former node enters the MIS, and, subsequently, all its neighbors renounce to enter the MIS. If two adjacent nodes simultaneously apply to enter the MIS, then there is a conflict, and both nodes remain undecided, and go to the next round. The round complexity of the algorithm heavily depends on the choice of the probability p_i that node i applies to enter the MIS, which may typically depend on the degree of node i , and may vary along with the execution of the algorithm, as node i accumulates more and more information about its neighborhood. Hence, a node i aiming at avoiding entering the MIS might be tempted to deviate from the algorithm by setting p_i small. However, if all nodes deviate in this way, then the algorithm may take a very long time before converging to a solution. The same holds whenever all nodes are aiming at entering the MIS.

Similar phenomenons may appear for other problems, like, e.g., coloring [8], that is, an abstraction of frequency assignment in radio networks. For solving this task, typical algorithms provides every node i with a probability distribution \mathcal{D}_i over the colors, and node i chooses color c at random with probability $\mathcal{D}_i(c)$. If this color does not conflict with the chosen colors of its neighbors, then node i adopts this color, else it performs another random choice, and repeats until no conflicts with the neighbors occur. However, some frequencies might be preferred to others because, e.g., some frequencies might be conflicting with local devices present at the node. As a consequence, not all colors are equal for the nodes, and while each node is aiming at constructing a coloring quickly (in order to take benefit from the resulting radio network), it is also aiming at avoiding being assigned a color that it does not want. Therefore, in a random assignments of colors, every node might be tempted to give more weight to its preferred colors than to its non desired colors, and if all nodes deviate in this way, then the algorithm may take a long time before converging to a solution, if converging at all.

In fact, such phenomena as those listed above are susceptible to occur for many network problems, typically for solving so-called *locally checkable labeling* tasks [67], or LCL tasks for short.

Locally Checkable Labelings

LCL tasks [67] form a large class of classical problems, including maximal independent set, coloring, maximal matching, minimal dominating set, etc., studied for more than 20 years in the framework of distributed computing in networks. An LCL task is characterized by

a finite set of *labels*, and a set of *good* labeled balls¹ of radius at most t , for some fixed $t \geq 0$. For instance, in the MIS task, balls are of radius 1, a label is either \bullet (interpreted as being member of the independent set), or \circ (interpreted as not being member of the independent set), and a labeled ball is good if either (1) its center is labeled \bullet , and all its neighbors are labeled \circ , or (2) its center is labeled \circ , and at least one of its neighbors is labeled \bullet . Similarly, in k -coloring, the labels are in $\{1, \dots, k\}$, balls are of radius 1, and a ball is good if the label of the center is different from the label of each of its neighbors.

Solving an LCL task consists in designing a distributed algorithm resulting in all nodes collectively assigning a label to each of them, such that all resulting balls are good.

In the following, we restrict ourselves to the large class of LCL tasks that are sequentially solvable by a greedy algorithm that (1) picks the nodes one by one in an arbitrary order, and (2) sets the label of each node when picked, after solely consulting the vicinity of the node. For instance, MIS is greedily constructible, as well as $(\Delta + 1)$ -coloring in networks of maximum degree Δ . Instead, Δ -coloring is not greedily constructible, as witnessed by 2-coloring even cycles. We restrict ourselves to greedily constructible LCL tasks because non greedily constructible tasks are hard to handle in the distributed network computing setting. Indeed, for solving such tasks, far away nodes might be forced to coordinate, yielding poor locality, as witnessed by, again, 2-coloring even cycles (which cannot be solved in less than $\Omega(n)$ rounds [58]).

A Generic “Luby’s Style” Randomized Algorithm for LCL Tasks

A generic randomized algorithm for LCL tasks, directly inspired from [6, 59], and therefore often referred to as “Luby’s style” algorithm, performs as follows. Every node v aims at computing its label, $label(v)$, for a given LCL task. The labels should be such that all resulting labeled balls are good with respect to the considered task. Node v starts with initial value $label(v) = \perp$. Let t be the radius of the task, i.e., the maximum radius of the labeled balls defining the tasks (both MIS and coloring have radius 1).

¹A *ball* of radius t is a graph with one identified node, called *center*, and with all the other nodes at distance at most t from the center. In a graph G , a ball of radius t centered at some node v is the subgraph induced by all nodes at distance at most t from v in G . A *labeled* ball is a ball whose every node is provided with a label (i.e., a bit-string).

Distributed Construction Algorithm: At each round, every node v which has not yet terminated observes the ball of radius t around it in the network (including its structure, and the already fixed nodes' labels). Then v chooses a random *temporary* label, $tmp\text{-}label(v)$, compatible with the current fixed labels of the nodes in the observed ball. Next, v observes the ball of radius t around it in the network again, and recovers the temporary labels randomly chosen by the nodes in the ball. If $tmp\text{-}label(v)$ is compatible with all fixed labels, and with all temporary labels in the ball of radius t centered at v (i.e., if the observed labeled ball is good w.r.t. the considered LCL task), then v sets $label(v)$ as equal to $tmp\text{-}label(v)$. The value of $label(v)$ is then fixed, and it is not subject to any modification in the future. Otherwise, node v goes to the next round.

Note that assuming that the LCL task is greedily constructive prevents nodes from being blocked by nodes that terminated at previous rounds: every node has always at least one label at its disposal for building a good ball around it. (It can be easily checked that, for all greedily constructible LCL tasks, the generic distributed construction algorithm terminates, and outputs correctly, as long as every label has non zero probability to be chosen). Note that each phase of the generic algorithm takes only $2t$ rounds for LCL task of radius t , as every node performs two snapshots of the labels (fixed or temporary) in its t -ball.

The random choice of $tmp\text{-}label(v)$ is governed by some probability distribution \mathcal{D} , which is actually characterizing the algorithm, and must be tuned according to the LCL task at hand. Importantly, \mathcal{D} may depend on the round, and may also depend on the structure of v 's neighborhood as observed during the previous rounds. It is known that, for many tasks such as Maximal Independent Set (MIS), and $(\Delta + 1)$ -coloring, there exist distributions \mathcal{D} enabling the generic distributed construction algorithm to terminate in $O(\log n)$ rounds in n -node networks (see, e.g., [6, 8, 59]). In this chapter, we consider the following issue:

What if selfish nodes are not playing according to the desired distribution \mathcal{D} ?

To address this issue, we define *LCL games*.

LCL Games

To every LCL task can be associated a game, that we call *LCL game*, and that we define as follows. Let G be a connected simple graph. Every node v of G is a rational and selfish *player*, playing the game with the ultimate goal of maximizing its payoff while performing the generic distributed construction algorithm described in Section 2.1.1.

Strategy. A *strategy* for a node v is a probability distribution \mathcal{D} over the labels compatible with the ball of radius t centered at v , which may depend on the history of v during the execution of the generic algorithm. For instance, in the MIS game, a strategy is a probability p to propose itself for entering the MIS. Similarly, in the $(\Delta + 1)$ -coloring game, a strategy is a distribution of probabilities over the set of remaining colors

compatible with the colors already assigned to the neighbors. (This set may even include a “fake” color 0 if nodes do not need to be systematically participating to a choice of color at every round [8]).

Remark. The distribution \mathcal{D} over the labels compatible with the ball of radius t centered at v is the unique item subject to non-orthodox behaviors. In particular, in LCL games, every node executes the prescribed algorithm, forwards messages correctly, and does not lie about its internal state, apart from what is concerning its private strategy for choosing its temporary label at random.

The *strategy* of a node at a round, i.e., the distribution \mathcal{D} of probability over the labels, may depend on the history of that node at any point in time during the execution of the generic distributed construction algorithm. On the other hand, the individual strategies depend only on the knowledge accumulated by the nodes along with the execution of the algorithm. In fact, at the beginning of the algorithm, player v does not even know which node she will play in the network, and just knows that the network may belong to some given graph family (like, e.g., cycles, planar graphs, etc.).

Payoff. The *payoff* of the nodes is aiming at capturing the tension between the objective of every node to compute a global solution rapidly (as this global solution brings benefits to every node), versus avoiding certain forms of solutions (which may not be desirable from an individual perspective). We denote by pref_v a preference function, which is an abstraction of how much node v will “suffer” in the future according to a computed solution. For instance, in the MIS game where nodes do not want to belong to the constructed MIS, one could set

$$\text{pref}_v(I) = \begin{cases} 0 & \text{if } v \in I \\ 1 & \text{otherwise} \end{cases}$$

for every MIS I . More specifically, we define, for each node v ,

$$\text{pref}_v : \{\text{good balls}\} \rightarrow [0, 1]$$

by associating to each good ball B centered at v the preference $\text{pref}_v(B)$ of v for that ball. The *payoff function* π_v of node v at the completion of the algorithm is decaying with the number k of rounds before the algorithm terminates at v . More precisely, we set

$$\pi_v = \delta^k \text{pref}_v(B_v)$$

where $0 < \delta < 1$ is a discount factor, B_v is the good ball centered at v as returned by the algorithm, and k is the number of rounds performed before all nodes in B_v fix their labels.

The choice of $\delta \in (0, 1)$ reflects the tradeoff between the quality of the solution from the nodes’ perspective, and their desire to construct a global solution quickly. Note that k is at least the time it takes for v to fix its label, and at most the time it takes for the algorithm to terminate at all nodes. If the algorithm does not terminate around v , that is, if a label remains perpetually undecided in at least one node of B_v , then we set $\pi_v = 0$.

- The payoff of a node v will thus be large if all nodes in B_v decide quickly (i.e., k is small), and if the labels computed in B_v suits node v (i.e., $\text{pref}_v(B_v)$ is close to 1).
- Conversely, if v or another node in its ball B_v takes many rounds before deciding a label (i.e., k is large), or if node v is not satisfied by the computed solution in B_v (i.e., $\text{pref}_v(B_v)$ is close to 0), then the payoff of v will be small.

In particular, if the preference for every good ball is the same, then maximizing the payoff is equivalent to completing the task as quickly as possible. Instead, if the preference is very small for some balls, then nodes might be willing to slow down the completion of the task, with the objective of avoiding being the center of such a ball, in order to maximize their payoff. That is, such nodes may bias their distribution \mathcal{D} towards preferred good balls, even if this is at the price of increasing the probability of conflicting with the choices of close nodes, resulting in more iterations before reaching convergence.

2.1.2 Our Results

We show that LCL games have *trembling-hand perfect equilibria*, that is, a stronger form of sequential equilibria due to Reinhard Selten [86], which are themselves a stronger form of Nash equilibria. Trembling-hand perfect equilibria include the possibility of off-the-equilibrium play, i.e., players may, with small probabilities, choose unintended strategies. In contrast, in Nash equilibria, players are assumed to play precisely as specified by the equilibrium. We show the following:

Theorem 2.1. *For any greedily constructible locally checkable labeling, the LCL game associated to that labeling has a symmetric trembling-hand perfect equilibrium.*

Therefore, in particular, for many tasks occurring in the context of distributed network computing such as MIS, and $(\Delta + 1)$ -coloring, there exist strategies played by the nodes of the network for solving these tasks such that no nodes have incentive to deviate from these strategies. Moreover, the related equilibria are strong forms of Nash equilibria which ensure that the players behave rationally even off the equilibrium path.

To establish Theorem 2.1, we first notice that LCL games belongs to the class of *extensive* games with *imperfect information*, because a node plays arbitrarily many times, and is not necessarily aware of the actions taken by far away nodes in the network. Also, LCL games belongs to the class of games with *infinite horizon* and *finite action set*: the horizon is infinite because neighboring nodes may perpetually prevent each other from terminating, and the action set is supposed to be finite (as long as the set of labels is finite). However, the classical game theoretical result [36] does not explicitly apply to LCL games. Indeed, first, in LCL games the actions of far-away nodes are not observable. Second, the imperfect information in [36] is solely related to the fact that players play simultaneously, while, again, in LCL games, imperfect information also refers to the fact that each node is not aware of the states of far away nodes in the network. It follows that the first step in our proof consists of revisiting the results in [36] for extending them, as specified in the following result.

Lemma 2.2. *Every infinite, continuous, measurable, well-rounded, extensive game with perfect recall and finite action set has a trembling-hand perfect equilibrium. Moreover, if the game is, in addition, symmetric, then it has a symmetric trembling-hand perfect equilibrium.*

The hypotheses regarding the nature of the strategy, and the nature of the payoff function (continuity, measurability, etc.) are standard in the framework of extensive games. The notion of *well-rounded* game is new, and is used to capture the fact that the nodes play in synchronous rounds in LCL games. The fact that the equilibrium is symmetric is crucial as far as games in networks are concerned since, in LCL games, as in randomized distributed computing in general, the instructions given to all nodes are identical, and the behavior of the nodes only vary along with the execution of the algorithm when they progressively discover their environment. Extending the results in [36] is quite technical, but follows the standard methods for establishing such results in game theory.

The more interesting part of the proof, as far as local distributed computing in networks is concerned, is to show that LCL games satisfy all requirements stated in Lemma 2.2. This is the role of the following result:

Lemma 2.3. *LCL games are symmetric, infinite, continuous, measurable, well-rounded, extensive games with perfect recall and finite action set.*

Lemmas 2.2 and 2.3 together prove Theorem 2.1. The rest of the chapter is therefore focussing on formalizing LCL games, and on proving Lemma 2.3.

2.1.3 Related Work

Let us first position our result into the various settings of game theory. Indeed, games take various forms, and the types of equilibria that can be satisfied by these games vary according to their forms. Table 2.1 surveys the results regarding equilibria for various game settings, from the finite strategic games to the extensive games with imperfect information (we restrict our attention to games with a finite number of players). Recall that *trembling-hand perfect* equilibria are refinements of *sequential* equilibria, which are themselves refinements of *subgame-perfect* equilibria, all of them being Nash equilibria. In Table 2.1, we distinguish strategic games (i.e., 1-step games like, e.g., prisoner's dilemma) from extensive games (i.e., game trees with payoffs, like, e.g., monetary policy in economy). For the latter class, we also distinguish games with perfect information (i.e., every player knows exactly what has taken place earlier in the game), from the games with imperfect information. We also distinguish finite games (i.e., games with a finite number of pure strategies, and finite number of repetitions) from games with infinite horizon (i.e., games which can be repeated infinitely often). The latter class of games is also split into games with finite numbers of actions, and games with infinite set of actions (like, e.g., when fixing the price of a product). In particular, Fudenberg and Levine [36] have proved that, under specific assumptions, every extensive game with imperfect information and finite action set has a sequential equilibrium (for behavior strategies). The specific class of games for which this result holds can be described as extensive games with observable actions, simultaneous moves, perfect recall, and finite action set, plus some continuity

	Strategic Games	Extensive games with perfect information	Extensive games with imperfect information
Finite games	[70] Nash equilibrium	[85] Subgame-perfect equilibrium Pure strategies	[86] Trembling-hand perfect eq. Behavior strategies
Games with a finite action set	Mixed strategies	[36] Subgame-perfect equilibrium Pure strategies	[36] Sequential equilibrium Behavior strategies
Games with an infinite action set	[28] [39] Nash equilibrium Mixed strategies	[44] Subgame-perfect equilibrium Pure strategies	[17] Nash equilibrium Behavior strategies

Table 2.1 – A summary of results about the existence of equilibria

requirements. Although this class of games captures repeated games, and contains natural games in economy, LCL games are not explicitly included into this class. Indeed, as we already mentioned, the actions of far-away nodes are not observable in LCL games, and, in these latter games, imperfect information also refers to the fact that each node is not aware of the states of far away nodes in the network.

We now list some previous works related to games in networks (for network formation games, see, e.g., [7, 46]). Many games in networks have complete information, and, among games with incomplete information, a large part of the literature is dedicated to single-stage games where players are not initially aware of the network topology (see the survey [47]). Repeated games in networks have also been considered a lot in the literature (again, see [47]). These games differ from LCL games since, in repeated games, the utility of a player depends on each round, and it is computed pairwise with each neighbor, while, in LCL games, the utility is computed solely when the player terminates, and may depend on the whole neighborhood. Regarding games with incomplete information involving communications in networks, it is worth mentioning [2, 15, 16, 37, 42]. However, all these work mostly refer to games in which the players' actions consist in choosing which information to reveal, and to whom it should be revealed. Instead, in LCL games, players actions are always fully observable by their neighbors at distance $\leq t$, where t is the maximum radius of the good balls for the considered LCL task.

Probably the first contribution to distributed computing by rational agents is [1], which studies leader election in various networks, including complete networks and rings. Different forms of Nash equilibria are shown to exist, for both synchronous and asynchronous computing. The contribution in [3] extended and generalized the results in [1] by considering other problems (consensus, renaming, etc.), and by identifying different utility functions that encompass different preferences of players in a distributed system: communication preference, solution preference, and output preference. The article [4] carried on this line of research, by enlarging the considered set of problems to coloring, partition, orientation, etc., and by addressing the question of how much global information agents should know a priori about the network in order for equilibria to exist. All these previous work differ from our approach in many ways. First, in [1, 3, 4], the

agents strategies define the algorithm itself, including which messages to send, which information to reveal, etc. Instead, in this chapter, the agents strategies solely consist in choosing a probability distribution on the possible outputs (at each round, depending on the history of the player). Second, the algorithms in the three aforementioned articles are “global” in the sense that they can take $\Omega(n)$ rounds in n -node networks. Instead, our (generic) algorithm is in essence “local”, i.e., it is expected to converge in a polylogarithmic number of rounds, even in networks with large diameter. Last but not least, we consider a whole family of tasks at once (all “reasonable” LCL tasks) while the three aforementioned articles address each task separately, each one with its own algorithm.

2.2 The Extensive Games Related to LCL Games

In this section, we specify the type of games we are interested in, aiming at capturing the characteristics of LCL games. We focus on extensive games with imperfect information, and we include infinite horizon in the analysis of such games. We formally define all the concepts appearing in the statement of Lemma 2.2, and/or useful for formally defining LCL games, and proving Lemma 2.3. In particular, we define the novel notion of *well-rounded* games, which fits with distributed network computing in the LOCAL model [78].

2.2.1 Basic Definitions

Recall that an extensive game is a tuple $\Gamma = (N, A, X, P, U, p, \pi)$, where:

- $N = \{1, \dots, n\}$ is the set representing the *players* of the game. An additional player, denoted by c , and called *chance*, is modeling all external random effects that might occur in the course of the game.
- A is the (finite) *action set*, i.e., a finite set representing the actions that can be made by each player when she has to play.
- X is the *game tree*, that is, a subset of $A^* \cup A^\omega$ where A^* (resp., A^ω) denotes the set of finite (resp., infinite) strings with elements in A , satisfying the following properties:
 - the empty sequence $\emptyset \in X$;
 - X is stable by prefix;
 - if $(a_i)_{i=1, \dots, k} \in X$ for every $k \geq 1$, then $(a_i)_{i \geq 1} \in X$.

The set X is partially ordered by the prefix relation, denoted by \preceq , where $x \preceq y$ means that x is a prefix of y , and $x \prec y$ means that x is a prefix of y distinct from y . The elements of X are called *histories*. A history x is *terminal* if it is a prefix of no other histories in X . In particular, every infinite history in X is terminal. The set of terminal histories is denoted by Z . If the longest history is finite then the game has *finite horizon*, otherwise it has *infinite horizon*. For every non-terminal history x , we denote by $A(x) = \{a \in A : (x, a) \in X\}$ the set of available actions after the history x .

- P is the *player partition*, i.e., a function $P : X \setminus Z \rightarrow N \cup \{c\}$ that assigns a player to each non-terminal history. $P(x)$ is the player who has to play after the history x . The sets $P_i = \{x \in X \setminus Z : P(x) = i\}$, for $i \in N \cup \{c\}$, called *player sets*, form a partition of $X \setminus Z$.
- U is the *information partition*, that is, a refinement of the player partition, whose elements are called *information sets*, such that for every $u \in U$, and for every two histories x, y in this *information set* u , we have $A(x) = A(y)$, i.e., the sets of available actions after x and after y are identical. We can therefore define $A(u)$ as the set of actions available after the information set u . Formally, $A(u) = \{a \in A : (x, a) \in X \text{ for every } x \in u\}$. For every history x , the information set containing x is denoted by $u(x)$. We also define $P(u)$ as the player who has to play after the information set u has been reached, and for every player i , the set $U_i = \{u \in U : P(u) = i\}$. The collection $\{U_i, i \in N \cup \{c\}\}$ forms a partition of U . Information sets regroup histories that are indistinguishable to players. Since the chance player c is not expected to behave rationally, we will simply put $U_c = \{\{x\}, x \in P_c\}$.
- p is a function that assigns to every history x in P_c (the player set of the chance c) a probability distribution over the set $A(x)$ of available actions after the history x . This *chance function* p is supposed to be common knowledge among the players.
- π is the *payoff function*, that is, $\pi : Z \rightarrow \mathbb{R}^n$ assigns the payoff (a real value) to every player in N for every terminal history of the game. We assume that every payoff is in $[-M, M]$ for some $M \geq 0$.

2.2.2 Well-Rounded Games

We introduce the concept of *rounds* in extensive games, and of *well-rounded* games.

Definition 2.1. The round function r of an extensive game assigns a positive integer to every non terminal history x , defined by $r(x) = |Rec(x)|$ where

$$Rec(x) = \{x' \in X \mid x' \prec x \text{ and } P(x') = P(x)\}.$$

We call $r(x)$ the round of x . The round of a finite terminal history is the round of its predecessor, and the round of an infinite history z is $r(z) = \infty$. An extensive game Γ for which the round function is non decreasing with respect to the prefix relation, i.e.,

$$y \preceq x \implies r(y) \leq r(x),$$

is said to be well-rounded.

Note that not every game is well-rounded, because two histories x and y such that $x \preceq y$ do not necessarily satisfy $P(x) = P(y)$. In a well-rounded game, since r is non decreasing, we have that, for any non terminal history x , every player has played at most $r(x) + 1$ times before x . Moreover, every player which has played less than $r(x)$ times before x will never play again after x .

Let $u \in U_i$ and $u' \in U_i$ be two (non necessarily distinct) information sets of the same player i , for which there exist $x \in u$, $x' \in u'$, and $a \in A(u')$, such that $(x', a) \preceq x$. Recall that an extensive game is said to have *perfect recall* if, for every such i , u , u' and a , we have:

$$\forall y \in u, \exists y' \in u' \mid (y', a) \preceq y.$$

The following lemma will allow us to safely talk about the round of an information set.

Lemma 2.4. *Let Γ be an extensive game with perfect recall, and let $x \in X$ and $x' \in X$ be two non terminal histories in the same information set $u \in U$. Then x and x' have the same round.*

Proof. We first observe the following. Let Γ be an extensive game with perfect recall, and let $y \in X$ be a finite history. Let $y' \in X$ and $y'' \in X$ for which there exists $u \in U$ such that

$$y' \in \text{Rec}(y) \cap u \text{ and } y'' \in \text{Rec}(y) \cap u.$$

Then $y' = y''$. Indeed, since both y' and y'' are in $\text{Rec}(y)$, we have that both are prefixes of y , and thus one of the two is a prefix of the other. Assume, w.l.o.g., that $y'' \prec y' \prec y$ (as, if $y' = y''$ then we are done). Let a be the action such that $(y'', a) \preceq y'$. Since the game has perfect recall, there must exist a history $y''' \in u$ such that $(y''', a) \preceq y''$. Thus $y''' \prec y'' \prec y' \prec y$. We can repeat the same reasoning for y''' and y'' as we did for y'' and y' . In this way, we construct an infinite strictly decreasing sequence of histories, which contradicts the fact that y is finite.

If both $\text{Rec}(x)$ and $\text{Rec}(x')$ are empty, then x and x' have the same round. Assume, w.l.o.g., that $\text{Rec}(x) \neq \emptyset$, and let $y \in \text{Rec}(x)$. Let a be the action such that $(y, a) \preceq x$. Since the game has perfect recall, there exists $y' \in u(x')$ such that $(y', a) \preceq x'$. Therefore $y' \prec x'$ and $P(y') = P(y) = P(x) = P(x')$. It follows that $y' \in \text{Rec}(x')$. Thus, for any $y \in \text{Rec}(x)$, we have identified a corresponding $y' \in \text{Rec}(x')$. This mapping from $\text{Rec}(x)$ to $\text{Rec}(x')$ is one-to-one. Indeed, let y_1 and y_2 in $\text{Rec}(x)$, and let y'_1 and y'_2 in $\text{Rec}(x')$ be the corresponding histories. If $y'_1 = y'_2$, then, since $u(y_1) = u(y'_1)$ and $u(y_2) = u(y'_2) = u(y'_1)$, we get that

$$y_1 \in \text{Rec}(x) \cap u(y'_1) \text{ and } y_2 \in \text{Rec}(x) \cap u(y'_1).$$

It follows from the above observation that $y_1 = y_2$. Thus the mapping is one-to-one, and hence $r(x) \leq r(x')$. It follows that we also have $\text{Rec}(x') \neq \emptyset$. Therefore, we can apply the same reasoning by switching the roles of x and x' , which yields $r(x') \leq r(x)$. Thus $r(x) = r(x')$. \square

2.2.3 Strategies, Outcomes, and Expected Payoff

In this section, we first recall several basic concepts about extensive games with perfect recall. Without loss of generality, we restrict our attention to *behavioral strategies* since such strategies are outcome-equivalent to mixed strategies thanks to [56]. The main objective of this section is to define the *expected payoff function*, which is novel as it is adapted to infinite games.

Recall that, for an information set u , the *local strategy* $b_{i,u}$ of a player i is a probability distribution over the set $A(u)$ of actions available given u . The set of local strategies of player i for u is denoted by $B_{i,u}$. The *behavioral strategy* b_i of a player i is a function which assigns a local strategy $b_{i,u}$ to every information set u of this player. The set of all behavioral strategies of player i is denoted by B_i . A *strategy profile* is a n -tuple of behavioral strategies, one for each player. The set of all strategy profiles is $B = \times_{i \in N} B_i$. For each player i , we denote by B_{-i} the set $\times_{j \neq i} B_j$. Since

$$B = B_i \times B_{-i} = \times_{i \in N} B_i,$$

a strategy profile b can be identified different ways, as $b = (b_i, b_{-i}) = (b_1, b_2, \dots, b_n)$. If every player plays according to a strategy profile b , then the outcome of the game is entirely determined, in the sense that every history x has a probability $\rho_b(x)$ of being reached. For every strategy profile b , and every history $x = (a_i)_{i=1, \dots, k}$ where $k \in \mathbb{N} \cup \infty$, the *realization probability* of x is defined by $\rho_b(\emptyset) = 1$, and

$$\rho_b(x) = \prod_{i=0}^{k-1} b_{P(x_i), u(x_i)}(a_{i+1})$$

where $x_0 = \emptyset$, and, for every positive $i \leq k$, $x_i = (a_1, a_2, \dots, a_i)$. For the chance player c , we simply identify its strategy with the chance function p :

$$P(x_i) = c \Rightarrow b_{P(x_i), u(x_i)}(a_{i+1}) = p(x_i, a_{i+1}).$$

The function $\rho_b : X \rightarrow [0, 1]$ is called the *outcome* of the game under the strategy profile b . An outcome $\rho : X \rightarrow [0, 1]$ is *feasible* if and only if there exists a strategy profile b such that $\rho = \rho_b$. The set of feasible outcomes of Γ is denoted by O .

We are now ready to define the *expected payoff*. Note that, in a game with infinite horizon, there can be uncountably many terminal histories. Therefore the definition of the probability measure on Z requires some care. For any finite history x , let

$$Z_x = \{z \in Z \mid x \preceq z\}.$$

Note that Z_x might be uncountable. Let Σ be the σ -algebra on Z generated by all sets of the form Z_x for some finite history x . For each strategy profile b , the measure μ_b on Σ is defined by: for every set Z_x , $\mu_b(Z_x) = \rho_b(x)$. This definition ensures that μ_b is a probability measure because $\mu_b(Z) = \mu_b(Z_\emptyset) = \rho_b(\emptyset) = 1$.

Definition 2.2. Let π be a payoff function that is measurable on Σ . The expected payoff function Π assigns a real value $\Pi(b)$ to every strategy profile $b \in B$, defined by

$$\Pi(b) = \int_{\Sigma} \pi \, d\mu_b.$$

Note that each component of the expected payoff function is bounded by M , where M is the upper bound on every payoff. A game Γ whose payoff function π is measurable on Σ is said to be a *measurable game*. In the following, we always assume that the considered games are measurable.

2.2.4 Equilibria, and Subgame Perfection

We now show how to adapt the standard notion of ϵ -equilibria (cf., e.g., [83]) to infinite games (Nash equilibria are special cases of ϵ -equilibria, with $\epsilon = 0$). Recall that a strategy profile b is a ϵ -equilibrium if and only if, for every player i , and every behavior strategy $b'_i \in B_i$ of this player, we have $\Pi_i(b'_i, b_{-i}) - \Pi_i(b) \leq \epsilon$. Similarly, we recall the notions of *subgames* and *subgame perfect equilibria* (see, e.g., [85]). A subtree X' of X is said to be *regular* if no information sets contain both a history in X' and a history not in X' . To each regular subtree X' is associated a game $\Gamma' = (N, A, X', P', U', p', \pi')$, where P' , U' , p' and π' are the restrictions of P , U , p and π to X' , called a *subgame*. The notions of outcomes and expected payoff functions for subgames follow naturally.

Definition 2.3. *A strategy profile b is a subgame perfect ϵ -equilibrium of an infinite game Γ if and only if, for every subgame Γ' , the restriction of b to Γ' is an ϵ -equilibrium.*

Note that a subgame perfect ϵ -equilibrium of Γ is an ϵ -equilibrium.

2.2.5 Metrics

In this section, we now define specific metrics on the set O of feasible outcomes, and on the set of behavior strategy profiles. These definitions are inspired from [36], with adaptations to fit our infinite setting.

Definition 2.4. *Let $\rho^1, \rho^2 \in O$ be two feasible outcomes of the same extensive game Γ . We define the following metric d on O : $d(\rho^1, \rho^2) = \sup_{x \in X, x \text{ finite}} 2^{-r(x)} \cdot |\rho^1(x) - \rho^2(x)|$ where $r(x)$ is the round of the finite history x .*

Lemma 2.5. *The function $d : O \times O \rightarrow \mathbb{R}$ specified in Definition 2.4 is a metric.*

Proof. We first show that d satisfies the triangle inequality. Let ρ^1, ρ^2 and ρ^3 be three feasible outcomes of Γ . For any finite history x we have the following:

$$\begin{aligned}
& |\rho^1(x) - \rho^3(x)| \leq |\rho^1(x) - \rho^2(x)| + |\rho^2(x) - \rho^3(x)| \\
\Rightarrow & 2^{-r(x)} |\rho^1(x) - \rho^3(x)| \leq 2^{-r(x)} |\rho^1(x) - \rho^2(x)| + 2^{-r(x)} |\rho^2(x) - \rho^3(x)| \\
\Rightarrow & \sup_{\substack{x \in X \\ x \text{ finite}}} 2^{-r(x)} |\rho^1(x) - \rho^3(x)| \leq \sup_{\substack{x \in X \\ x \text{ finite}}} \left(2^{-r(x)} |\rho^1(x) - \rho^2(x)| + 2^{-r(x)} |\rho^2(x) - \rho^3(x)| \right) \\
\Rightarrow & \sup_{\substack{x \in X \\ x \text{ finite}}} 2^{-r(x)} |\rho^1(x) - \rho^3(x)| \leq \sup_{\substack{x \in X \\ x \text{ finite}}} 2^{-r(x)} |\rho^1(x) - \rho^2(x)| + \sup_{\substack{x \in X \\ x \text{ finite}}} 2^{-r(x)} |\rho^2(x) - \rho^3(x)| \\
\Rightarrow & d(\rho^1, \rho^3) \leq d(\rho^1, \rho^2) + d(\rho^2, \rho^3).
\end{aligned}$$

Next, we prove that d separates different outcomes. Let ρ^1 and ρ^2 be two feasible outcomes such that $d(\rho^1, \rho^2) = 0$. By definition, this implies that, for every finite history x , we have $\rho^1(x) = \rho^2(x)$. Let b^1 and b^2 be two strategy profiles such that $\rho^1 = \rho_{b^1}$ and $\rho^2 = \rho_{b^2}$. Let $x = (a_k)_{k \geq 1}$ be an infinite history, and, for $k \geq 1$, let $x_k = (a_1, a_2, \dots, a_k)$ be the corresponding increasing sequence of its prefixes. By definition, $\rho_{b^1}(x)$ is the limit, for $k \rightarrow \infty$, of the sequence $\rho_{b^1}(x_k)$, and $\rho_{b^2}(x_k) \rightarrow \rho_{b^2}(x)$ when $k \rightarrow \infty$ as well. Since the two sequences are equal, they have the same limit, and therefore $\rho^1(x) = \rho^2(x)$. Since this equality holds for every infinite history x , it follows that $\rho^1 = \rho^2$. \square

We can use d to define a metric on behavioral strategy profiles as follows. Let $b^1, b^2 \in B$ be two behavioral strategy profiles of the same game Γ . We define the metric d on B by:

$$d(b^1, b^2) = \max \left\{ d(\rho_{b^1}, \rho_{b^2}), \sup_{\substack{i \in N \\ b_i \in B_i}} d(\rho_{(b_i, b_{-i}^1)}, \rho_{(b_i, b_{-i}^2)}) \right\}.$$

Finally, we define the continuity of the expected payoff function using the sup norm over \mathbb{R}^n . Specifically, the expected payoff function Π is *continuous* if, for every sequence of strategy profiles $(b^k)_{k \geq 1}$, and every strategy profile b , we have:

$$d(b^k, b) \xrightarrow{k \rightarrow \infty} 0 \implies \sup_{i \in N} |\Pi_i(b^k) - \Pi_i(b)| \xrightarrow{k \rightarrow \infty} 0.$$

An extensive game Γ is continuous if its expected payoff function is continuous.

2.3 Proof of Lemma 2.2

In this section, we prove that, as stated in Lemma 2.2, every (symmetric) infinite, continuous, measurable, well-rounded, extensive game with perfect recall and finite action set has a (symmetric) trembling-hand perfect equilibrium. To this purpose, we begin by establishing results about perturbed games, and about truncated games.

2.3.1 Perturbed Games and Perfect Equilibria

In this section, we extend the classical notion of trembling-hand perfect ϵ -equilibria to infinite games, generalizing [86]. We start by revisiting the notion of a perturbed game in the context of infinite games. A *perturbed game* $\hat{\Gamma}$ is a pair (Γ, η) , where Γ is an extensive game and η is a function that assigns a positive probability $\eta_u(a)$ to *every* action a available after the information set u , such that

$$\forall u \in U, \quad \sum_{a \in A(u)} \eta_u(a) < 1. \quad (2.1)$$

The perturbed game $\hat{\Gamma}$ is to be interpreted as a restriction of the game Γ in the following sense. The probability $\eta_u(a)$ is the *minimal probability* that player $P(u)$ can assign to the action a in his local strategy at information set u . Therefore, we define the set of strategies of the perturbed game as follows. In the perturbed game $\hat{\Gamma}$, a *local strategy* $\hat{b}_{i,u}$ of a player i at one of its information sets u is a probability distribution over the set $A(u)$ of actions available at this information set, satisfying:

$$\forall (i, u, a) \in N \times U_i \times A(u), \quad \hat{b}_{i,u}(a) \geq \eta_u(a). \quad (2.2)$$

The set of local strategies of player i at the information set u in the perturbed game $\hat{\Gamma}$ is denoted by $\hat{B}_{i,u}$. Note that $\hat{B}_{i,u}$ is not reduced to a single strategy thanks to Eq. (2.1). From the above, we can define the set \hat{B}_i of behavioral strategies of player i , the set \hat{B} of strategy profiles, and the notion of ϵ -equilibria of a perturbed game in the same way as for regular extensive games.

Definition 2.5. A strategy profile b^* is a trembling-hand perfect ϵ -equilibrium of Γ if the following two conditions hold:

1. There exists a sequence $(\hat{\Gamma}^k) = (\Gamma, \eta^k)$, $k \geq 1$, of perturbations of Γ such that $\eta^k \rightarrow 0$ when $k \rightarrow \infty$ (where $\eta^k \rightarrow 0$ means that, for every player i , every information set $u \in U_i$, and every action $a \in A(u)$, $\eta_u^k(a) \rightarrow 0$);
2. There exists a sequence of ϵ -equilibria \hat{b}^k of these perturbed games such that $\hat{b}^k \rightarrow b^*$ when $k \rightarrow \infty$ (where $\hat{b}^k \rightarrow b^*$ in the sense of the metric on behavioral strategy profiles).

Such a sequence $(\hat{\Gamma}^k)_{k \geq 1}$ of perturbed games is called a test sequence for Γ .

The next lemma shows that the notion of trembling-hand perfect ϵ -equilibrium is a refinement of the notion of subgame perfect ϵ -equilibrium.

Lemma 2.6. Every trembling-hand perfect ϵ -equilibrium of a continuous game is a subgame perfect ϵ -equilibrium.

Proof. We first show that every trembling-hand perfect ϵ -equilibrium of a continuous game is an ϵ -equilibrium. Let b^* , Γ , $\hat{\Gamma}^k$ and \hat{b}^k be as in Definition 2.5. Since \hat{b}^k are ϵ -equilibria of the games $\hat{\Gamma}^k$, the following holds:

$$\forall i \in N, \forall k, \forall b_i \in \hat{B}_i^k, \quad \Pi_i(b_i, \hat{b}_{-i}^k) - \Pi_i(\hat{b}^k) \leq \epsilon.$$

Let B_i^m be the intersection of all \hat{B}_i^k with $k \geq m$. Then we have:

$$\forall i \in N, \forall k \geq m, \forall b_i \in B_i^m, \quad \Pi_i(b_i, \hat{b}_{-i}^k) - \Pi_i(\hat{b}^k) \leq \epsilon.$$

Since Π is assumed to be continuous, these inequalities remain valid if, on both sides, we take the limits for $k \rightarrow \infty$. Therefore,

$$\forall i \in N, \forall b_i \in B_i^m, \quad \Pi_i(b_i, b_{-i}^*) - \Pi_i(b^*) \leq \epsilon.$$

This latter inequality holds for every m . The closure of the union of all B_i^m is B_i . Since Π is continuous, this yields:

$$\forall i \in N, \forall b_i \in B_i, \quad \Pi_i(b_i, b_{-i}^*) - \Pi_i(b^*) \leq \epsilon,$$

as desired. Moreover, every ϵ -equilibrium of a perturbed game is subgame perfect. Indeed, this follows from the fact that in a perturbed game, every subgame is reached with a strictly positive probability.

We now have all the ingredients to establish the lemma. Let b^* be a trembling-hand perfect ϵ -equilibrium of the game Γ , and let \hat{b}^k be a sequence of equilibria of the test sequence $\hat{\Gamma}$ such that $\hat{b}^k \rightarrow b^*$. Every subgame Γ' of Γ induces a sequence $\hat{\Gamma}'^k$ of subgames of the test sequence. Since every ϵ -equilibrium of a perturbed game is subgame perfect, it follows that the sequence \hat{b}^k induces a sequence of equilibria \hat{b}'^k of these perturbed subgames, which converges to b'^* , the strategy profile induced by b^* on Γ' . Therefore b'^* is a trembling-hand perfect ϵ -equilibrium of Γ' , which proves that b^* is subgame perfect since every trembling-hand perfect ϵ -equilibrium of a continuous game is an ϵ -equilibrium. \square

2.3.2 Truncated Games and Induced Equilibria

The goal of this section is to introduce the concept of *truncation* of an infinite game, and to draw links between equilibria of the truncated (finite) game, and the equilibria of the corresponding infinite game. The purpose of truncation is to deal with finite games, which are easier to handle than infinite ones. The material in this section is a generalization of [36] to extensive games (the original article [36] only applies to repeated games).

For every well-rounded extensive game $\Gamma = (N, A, X, P, U, p, \pi)$ with perfect recall, and for every positive integer t , we define the *truncated game* $\Gamma_t = (N, A, X_t, \tilde{P}, \tilde{U}, \tilde{p}, \tilde{\Pi})$ as follows.

- X_t is the set of all histories $x \in X$ of round at most t . The set of terminal histories \tilde{Z} of Γ_t is defined as the set of all histories with a predecessor in X_t and no successor in X_t .
- \tilde{P} , \tilde{U} and \tilde{p} are the restrictions of P , U , and p to X_t . Note that these restrictions do not break information sets thanks to Lemma 2.4.
- Strategies, and strategy profiles of Γ are mapped to strategies, and to strategy profiles of Γ_t , respectively, as follows.
 - Every strategy profile b of Γ is mapped to a strategy profile \tilde{b} of Γ_t obtained by the restriction of b to the information sets in X_t . The set of strategy profiles of Γ_t is denoted by \tilde{B} . The strategy profile \tilde{b} is said to be *induced* by b .
 - Conversely, every strategy profile \tilde{b} of Γ_t is mapped to the strategy profile b (said to be induced by \tilde{b}) of Γ defined by:
 - * $b_{i,u} = \tilde{b}_{i,u}$ for every information set $u \in \tilde{U}$, and every player i ,
 - * $b_{i,u}(a) = \frac{1}{|A(u)|}$ for every information set $u \notin \tilde{U}$, every player i , and every action a (i.e., $b_{i,u}$ is a uniform distribution).

Note that the first induction mapping is not necessarily one-to-one, while the second is not necessarily onto.

- In order to define the new payoff function $\tilde{\pi}$ over the set \tilde{Z} , we rather define it implicitly, via the expected payoff function $\tilde{\Pi}$ which is set such that, for any strategy profile $\tilde{b} \in \tilde{B}$ that induces $b \in B$, we have $\tilde{\Pi}(\tilde{b}) = \Pi(b)$.

The following lemma characterizes ϵ -equilibria, using truncated games.

Lemma 2.7. *In a continuous well-rounded game with perfect recall Γ , a strategy profile b^* is an ϵ -equilibrium if and only if there exist sequences ϵ^k , t^k , and b^k such that, for every k , b^k is an ϵ^k -equilibrium in Γ_{t^k} , and, as k goes to infinity, $\epsilon^k \rightarrow \epsilon$, $t^k \rightarrow \infty$, and $b'^k \rightarrow b^*$, where b'^k is induced by b^k on Γ .*

Proof. Given a game Γ , for every $t \geq 1$, let

$$w^t = \max_{i \in N} \sup_{b^1, b^2 \in B : \tilde{b}^1 = \tilde{b}^2} |\Pi_i(b^1) - \Pi_i(b^2)|$$

where $\tilde{b}^1 = \tilde{b}^2$ stands for the fact that b^1 and b^2 induce the same strategy profiles over the truncated game Γ_t . Note that $w^t \leq 2M$ for every $t \geq 1$.

Claim 2.8. *A continuous game Γ satisfies $w^t \xrightarrow{t \rightarrow \infty} 0$.*

To establish the claim, let b^1 and b^2 be two strategy profiles of Γ that induce the same strategy profiles over Γ_t . This implies that, for any history x of round at most t , and for any deviation b_i of some player i , the relations $\rho_{b^1}(x) = \rho_{b^2}(x)$ and $\rho_{(b_i, b_{-i}^1)}(x) = \rho_{(b_i, b_{-i}^2)}(x)$ hold. Therefore by the definition of the metric d on feasible outcomes (cf. Definition 2.4); and of its extension to behavioral strategy profiles, we have $d(b^1, b^2) \leq 2^{-(t+1)}$. Therefore:

$$w^t \leq \sup_{\substack{i \in n \\ b^1, b^2 \in B \\ d(b^1, b^2) \leq 2^{-(t+1)}}} |\Pi_i(b^1) - \Pi_i(b^2)|.$$

Since Γ is continuous, we get

$$\sup_{\substack{i \in n \\ b^1, b^2 \in B \\ d(b^1, b^2) \leq 2^{-(t+1)}}} |\Pi_i(b^1) - \Pi_i(b^2)| \xrightarrow{t \rightarrow \infty} 0$$

which completes the proof of Claim 2.8.

Claim 2.9. *Let Γ be a well-rounded game with perfect recall, and let Γ_t be its truncation. We have that: (1) Any ϵ -equilibrium \tilde{b}^* in Γ_t induces a $(\epsilon + w^t)$ -equilibrium in Γ ; (2) Any ϵ -equilibrium b^* in Γ induces a $(\epsilon + 2w^t)$ -equilibrium in Γ_t .*

To establish the first item of the claim, let \tilde{b}^* be an ϵ -equilibrium in Γ_t which induces b^* in Γ . Let i be a player, $b'_i \in B_i$ a deviation of this player in Γ , \tilde{b}_i the deviation induced in Γ_t , and b_i the deviation induced by \tilde{b}_i in Γ . Since \tilde{b}^* is an ϵ -equilibrium in Γ_t , we get that $\tilde{\Pi}_i(\tilde{b}_i, \tilde{b}_{-i}^*) - \tilde{\Pi}_i(\tilde{b}^*) \leq \epsilon$. By definition of $\tilde{\Pi}$, the following two equalities hold: $\tilde{\Pi}_i(\tilde{b}_i, \tilde{b}_{-i}^*) = \Pi_i(b_i, b_{-i}^*)$, and $\tilde{\Pi}_i(\tilde{b}^*) = \Pi_i(b^*)$. Furthermore, by definition of w^t , we have $\Pi_i(b'_i, b_{-i}^*) - \Pi_i(b_i, b_{-i}^*) \leq w^t$. We eventually derive that $\Pi_i(b'_i, b_{-i}^*) - \Pi_i(b^*) \leq \epsilon + w^t$. The latter equality proves that b^* is a $(\epsilon + w^t)$ -equilibrium, since i is any player, and since b'_i is any deviation in B_i .

To establish the second item of Claim 2.9, let b^* be an ϵ -equilibrium in Γ , \tilde{b}^* be the strategy profile it induces on Γ_t , and b'^* be the strategy profile induced by \tilde{b}^* on Γ . Let i be a player, $\tilde{b}_i \in \tilde{B}_i$ be a deviation in Γ_t , and b_i be the deviation it induces in Γ . Since b^* is an ϵ -equilibrium, we get: $\Pi_i(b_i, b_{-i}^*) - \Pi_i(b^*) \leq \epsilon$. By definition of $\tilde{\Pi}$, we have: $\tilde{\Pi}_i(\tilde{b}^*) = \Pi_i(b^*)$, and $\tilde{\Pi}_i(\tilde{b}_i, \tilde{b}_{-i}^*) = \Pi_i(b_i, b_{-i}^*)$. Moreover, by definition of w^t , we have: $\Pi_i(b'_i, b_{-i}^*) - \Pi_i(b_i, b_{-i}^*) \leq w^t$, and $\Pi_i(b_i, b_{-i}^*) - \Pi_i(b_i, b_{-i}^*) \leq w^t$. The above relations put together

eventually yield: $\tilde{\Pi}_i(\tilde{b}_i, \tilde{b}_{-i}^*) - \tilde{\Pi}_i(\tilde{b}^*) \leq \epsilon + 2w^t$. This latter inequality shows that \tilde{b}^* is a $(\epsilon + 2w^t)$ -equilibrium in Γ_t . This completes the proof of Claim 2.9.

Note that Claim 2.9 applies to perturbed games as well, as long as the induced strategy profiles satisfy the constraints of the minimal probabilities, as specified in Eq. (2.2).

Claim 2.10. *Let b^k be a sequence of ϵ -equilibria of a continuous well-rounded game Γ with perfect recall such that $b^k \rightarrow b^*$. Then b^* is also an ϵ -equilibrium of Γ . In other words, for every ϵ , the set of ϵ -equilibria of Γ is closed.*

To establish the claim, suppose that b^* is not an ϵ -equilibrium of Γ . Therefore for some player i , and for some $\delta > 0$, there exists a deviation $b_i \in B_i$ such that

$$\Pi_i(b_i, b_{-i}^*) - \Pi_i(b^*) \geq \epsilon + 3\delta.$$

By continuity of Π , we have that, for k large enough:

$$\begin{aligned} \Pi_i(b^k) - \Pi_i(b^*) &\leq \delta; \\ \Pi_i(b_i, b_{-i}^*) - \Pi_i(b_i, b_{-i}^k) &\leq \delta. \end{aligned}$$

Subtracting the two above inequalities to the previous one yields:

$$\Pi_i(b_i, b_{-i}^k) - \Pi_i(b^k) \geq \epsilon + \delta$$

which contradicts the premise that b^k is an ϵ -equilibrium. This completes the proof of Claim 2.10.

We now have all the ingredients to prove Lemma 2.7. Assume first that b^* is an ϵ -equilibrium in Γ , and let us define b^k as the strategy profile induced by b^* on Γ_k , and b'^k as the strategy profile induced by b^k on Γ . By Claim 2.9, it follows that, for every k , b^k is a $(\epsilon + 2w^k)$ -equilibrium in Γ_k . By Claim 2.8, we get $w^k \rightarrow 0$. Furthermore, $d(b'^k, b^*) \leq 2^{-k}$ by definition of d . Therefore, we have $b'^k \rightarrow b^*$.

Conversely, assume that there exists sequences ϵ^k , t^k , and b^k as in the statement of the lemma. By Claim 2.9, for every k , the strategy profile b'^k induced by b^k on Γ is an $(\epsilon^k + w^{t^k})$ -equilibrium in Γ . Since Γ is continuous, it follows from Claim 2.8 that $(\epsilon^k + w^{t^k}) \rightarrow \epsilon$. For every $\delta > 0$, we have that, for k large enough, $\epsilon^k + w^{t^k} \leq \epsilon + \delta$. By Claim 2.10, this implies that b^* is a $(\epsilon + \delta)$ -equilibrium. Since this holds for every $\delta > 0$, we have shown that b^* is an ϵ -equilibrium in Γ . \square

2.3.3 Truncation of Perturbed Games

The goal of this section is to extend the results of the previous section to trembling-hand perfect ϵ -equilibria (instead of just for ϵ -equilibria). We define the truncation of a perturbed game in a straightforward way: for every perturbed game $\hat{\Gamma} = (\Gamma, \eta)$, and every positive integer t , the *truncated perturbed game* $\hat{\Gamma}_t$ is $(\Gamma_t, \tilde{\eta})$, where $\tilde{\eta}$ is simply the restriction of η to X_t .

Lemma 2.11. *In a continuous game Γ , a strategy profile b^* is a trembling-hand perfect ϵ -equilibrium if and only if there exist sequences ϵ^k , t^k and b^k such that, for every k , b^k is a trembling-hand perfect ϵ^k -equilibrium in Γ_{t^k} , and, as k goes to infinity, $\epsilon^k \rightarrow \epsilon$, $t^k \rightarrow \infty$, and $b'^k \rightarrow b^*$, where b'^k is induced by b^k on Γ .*

Proof. We first show the following.

Claim 2.12. *Let Γ be a continuous game, and let Γ_t be its truncation.*

1. *Any trembling-hand perfect ϵ -equilibrium \tilde{b}^* in Γ_t induces a trembling-hand perfect $(\epsilon + w^t)$ -equilibrium in Γ .*
2. *Any trembling-hand perfect ϵ -equilibrium b^* in Γ induces a trembling-hand perfect $(\epsilon + 2w^t)$ -equilibrium in Γ_t .*

For proving the first item of Claim 2.12, let \tilde{b}^* be a trembling-hand perfect ϵ -equilibrium in Γ_t . By definition of a trembling-hand perfect equilibrium, there exists a sequence \tilde{b}^k of ϵ -equilibrium of a test sequence $\hat{\Gamma}_t^k = (\Gamma_t, \tilde{\eta}^k)$ of the game Γ_t , such that $\tilde{b}^k \rightarrow \tilde{b}^*$. This test sequence can be extended into a sequence $\hat{\Gamma}^k = (\Gamma, \eta^k)$ of infinite games by (1) choosing the same minimal probabilities for every information set of round at most t , and for every information set u of round larger than t , and (2) defining $\eta_u^k(a) = \frac{2^{-k}}{|A(u)|}$ for every action $a \in A(u)$. Therefore the sequence $\hat{\Gamma}^k$ is, by construction, a test sequence of Γ . Furthermore, η^k is specifically set such that Claim 2.9 applies. That is, the strategy profiles b^k induced by \tilde{b}^k are therefore $(\epsilon + w^t)$ -equilibria of $\hat{\Gamma}_t^k$. We also have that b^k is induced by \tilde{b}^k , and b^* is induced by \tilde{b}^* and $\tilde{b}^k \rightarrow \tilde{b}^*$. It follows that $b^k \rightarrow b^*$. This proves that b^* is a trembling-hand perfect $(\epsilon + w^t)$ -equilibrium of Γ .

For proving the second item of Claim 2.12, let b^* be a trembling-hand perfect ϵ -equilibrium in Γ . By definition, there exists a sequence b^k of ϵ -equilibria of a test sequence $\hat{\Gamma}^k$ of Γ such that $b^k \rightarrow b^*$. For k large enough, the following holds:

$$\forall u \in U, \forall a \in A(u), \quad \eta_u^k(a) \leq \frac{1}{|A(u)|}.$$

Therefore we can apply Claim 2.9. We get that the strategy profiles \tilde{b}^k induced by b^k in $\hat{\Gamma}_t^k$ are $(\epsilon + 2w^t)$ -equilibria. Note that $\hat{\Gamma}_t^k$ is a test sequence for Γ_t . Also, since \tilde{b}^k is induced by b^k , and \tilde{b}^* is induced by b^* and $b^k \rightarrow b^*$, it follows that \tilde{b}^k converges to \tilde{b}^* . This proves that \tilde{b}^* is a $(\epsilon + 2w^t)$ -equilibrium in Γ_t , which completes the proof of Claim 2.12.

Claim 2.13. *Let b^k be a sequence of trembling-hand perfect ϵ -equilibria of a game Γ such that $b^k \rightarrow b^*$. Then b^* is a trembling-hand perfect ϵ -equilibrium of Γ .*

To see why the claim holds, observe that, for every k , since b^k is a trembling-hand perfect ϵ -equilibrium, there exists a sequence $b^{k,n}$ of ϵ -equilibria of a test sequence $\hat{\Gamma}^n$ of Γ , such that $b^{k,n} \xrightarrow{n \rightarrow \infty} b^k$. The sequence $b^{k,k}$ of ϵ -equilibria of the test sequence $\hat{\Gamma}^k$ converges to b^* , thus it is a trembling-hand perfect ϵ -equilibrium of Γ .

The proof of the lemma is then identical to the proof of Lemma 2.7, by replacing Claims 2.9 and 2.10 by Claims 2.12 and 2.13, respectively. \square

2.3.4 The proof of Lemma 2.2

We now have all the ingredients to prove that, as stated in Lemma 2.2 every infinite, continuous, measurable, well-rounded, (symmetric) extensive game with perfect recall and finite action set has a (symmetric) trembling-hand perfect equilibrium.

Regarding symmetric games, recall that Γ is a *symmetric game* if it satisfies the following requirements.

- For every player i , there exist a bijection $type_i : U_i \rightarrow [1, \dots, I]$ such that for every two players i and i' , and every two information sets $u \in U_i$ and $u' \in U_{i'}$ with $type_i(u) = type_{i'}(u')$, we also have $A(u) = A(u')$ and $actions_i(u) = actions_{i'}(u')$.
- For every two players i and i' , and every two strategy profiles b and b' with $b R_{i,i'} b'$, we have

$$\Pi_i(b) = \Pi_{i'}(b') \text{ and } \Pi_i(b') = \Pi_{i'}(b),$$

where, for every two players i and i' , we define the relation $R_{i,i'}$ on strategy profiles as follows. For every two strategy profiles b and b' , we have $b R_{i,i'} b'$ if and only if, for every positive integer k ,

$$b_{i,type_i^{-1}(k)} = b'_{i',type_{i'}^{-1}(k)} \text{ and } b'_{i,type_i^{-1}(k)} = b_{i',type_{i'}^{-1}(k)},$$

and, for every player j different from i and i' , $b_j = b'_j$.

In a symmetric game Γ , we say that a profile of strategies $b \in B$ is *symmetric* if and only if for every two players i and i' , we have $b R_{i,i'} b$.

First observe that the set of strategy profiles B of an extensive game Γ with a finite action set is sequentially compact. Indeed, let Γ be a game with a finite action set A . For every player i , and for every information set u , the set of local strategies B_{iu} is a simplex in the space $\mathbb{R}^{A(u)}$, and therefore it is sequentially compact. Since a countable product of sequentially compact spaces is sequentially compact, we get that

$$B = \prod_{i \in N} \prod_{u \in U_i} B_{iu}$$

is sequentially compact. Indeed, there are countably many finite histories in any game with a finite action set, and therefore there are also countably many information sets.

Consider now a (possibly infinite) continuous well-rounded extensive game Γ with perfect recall and a finite action set, and consider the corresponding sequence of truncated games Γ_k . [86] has shown that every finite extensive game with perfect recall has a trembling-hand perfect equilibrium. By this result, every game Γ_k has a trembling-hand perfect equilibrium. We denote by b^k this equilibrium, and we denote by b'^k the sequence of strategy profiles induced by b^k in the game Γ . Since B is sequentially compact, we can extract a convergent subsequence b'^{t_k} . let us denote by b^* its limit. Lemma 2.11 shows that b^* is a trembling-hand perfect equilibrium of Γ .

Since a limit of a sequence of symmetric strategy profiles is symmetric, we can derive that b^* is a symmetric trembling-hand perfect equilibrium. \square

2.4 Proof of Lemma 2.3

In this section, we show that LCL games satisfy all hypotheses of Lemma 2.2, from which we derive Theorem 2.1. We start by formally defining LCL games.

2.4.1 Formal Definition of LCL Games

Let A be a finite alphabet, \mathcal{F} a family of graphs with at most n vertices, \mathcal{D} a probability distribution over \mathcal{F} , and \mathcal{L} a greedily constructible LCL task over graphs in \mathcal{F} . Let $\text{good}(\mathcal{L})$ be the set of good balls in \mathcal{L} , and let t be the radius of \mathcal{L} , that is, the largest radius of the good balls. Let $\text{pref}_i : \text{good}(\mathcal{L}) \mapsto [0, 1]$, be a preference function of player i over good balls, and let $\delta \in (0, 1)$, called *discounting factor*. We define the game

$$\Gamma(\mathcal{L}, \mathcal{D}, \text{pref}, \delta) = (N, A, X, P, U, p, \pi)$$

associated to the LCL task \mathcal{L} , the distribution \mathcal{D} , the preference function $\text{pref} = (\text{pref}_i)_{i \in N}$, and the discounting factor δ , as follows.

- The player set is $N = \{1, \dots, n\}$.
- The action set is $A \cup \mathcal{F}$ where the actions in \mathcal{F} are only used by the chance player c in the initial move, and the actions in A are used by the actual players in N .
- The first move of the game is made by the chance player, that is, $P(\emptyset) = c$. As a result, a graph $G \in \mathcal{F}$ is selected at random according to the probability distribution \mathcal{D} , and a one-to-one mapping of the players to the nodes of G is chosen uniformly at random. From now on, the players are identified with the vertices of the graph G , labeled from 1 to n . Note that \mathcal{F} might be reduced to a single graph, e.g., $\mathcal{F} = \{C_n\}$, and the chance player just selects, for each vertex v , which player $i \in N$ is playing at v (in a one-to-one manner).
- The game is then divided into rounds (corresponding to the intuitive meaning in synchronous distributed algorithms). At each round, the *active* players play in increasing order, from 1 to n . At round 0 every player is active and plays, and every action in A is available.
- At the end of each round (i.e., after every active player has played the same number of times), some players might become *inactive*, depending on the actions chosen during the previous rounds. For every $i \in N$, let $s(i)$ denote the last action played by player i , which we call the *state* of i , and let $\text{ball}(i)$ denote the ball of radius t centered at node i . Every player i such that $\text{ball}(i) \in \text{good}(\mathcal{L})$ at the end of a round becomes *inactive*.
- In subsequent rounds, the set of available actions might be restricted. For every round $r > 0$, and for every active player i , an action $a \in A$ is available to player i if and only if there exists a ball $b \in \text{good}(\mathcal{L})$ compatible with the states of inactive players in which $s(i) = a$.

- A history is terminal if and only if either it is infinite, or it comes after the end of a round with every player being inactive after that round.
- Let x be a history. We denote by $actions_i(x)$ the sequence of actions extracted from x by selecting all actions taken by player i during rounds before $r(x)$. (The action possibly made by player i at round $r(x)$, and actions made by a player $j \neq i$ are not included in $actions_i(x)$).
- Let x and y be two non terminal histories such that $P(x) = P(y) = i$. Then x and y are in the same information set if and only if, for every $j \in ball(i)$, we have

$$actions_j(x) = actions_j(y).$$

This can be interpreted by the fact that a player i “knows” every action previously taken by any player at distance at most t from i in the graph.

- Let i be a player, and let z be a terminal history. We define the *terminating time* of player i in history z by $time_i(z) = \max\{|actions_j(z)|, j \in ball(i)\} - 1$. The payoff function π of the game is then defined as follows. For every player i , and every terminal history z , we have $\pi_i(z) = \delta^{time_i(z)} \cdot pref_i(ball(i))$. And $\pi_i(z) = 0$ if $time_i(z) = \infty$.

2.4.2 The proof of Lemma 2.3

We survey the properties of LCL games, with emphasis on those listed as pre-conditions in the statement of Lemma 2.2.

Lemma 2.14. *LCL games are well-rounded.*

Proof. This follows directly from the fact that, in a LCL game, (1) every active player plays at every round until it becomes inactive, and (2) once inactive, a player cannot become active again. \square

Lemma 2.15. *LCL games are symmetric.*

Proof. This follows directly from the fact that, in a LCL game, the position of every player in the actual graph (which might be fixed, or chosen at random in some given family of graphs according to some given distribution) is chosen uniformly at random. \square

Lemma 2.16. *LCL games have perfect recall.*

Proof. Let $\Gamma = (\mathcal{L}, \mathcal{D}, pref, \delta) = (N, A, X, P, U, p, \pi)$ be an LCL game. Let u and u' be two information sets of the same player i , for which there exists $x \in u$, $x' \in u'$, and $a \in A(u')$ such that $(x', a) \preceq x$. Let y be a history in u . Since x and y are in the same information set u , it follows that, for every player $j \in ball(i)$, we have $actions_j(x) = actions_j(y)$. In particular, this implies that x and y are in the same round. Let y' be the unique history which is a prefix of y with $P(y') = i$, and with $r(y') = r(x')$. (Such a history exists because $r(x') < r(x)$, and $r(y) = r(x)$). Since the players play in the same order at every round,

we get that, for every player $j \in \text{ball}(i)$, $\text{actions}_j(x') = \text{actions}_j(y')$. As a consequence, we have $y' \in u'$. Furthermore, since $\text{actions}_i(x) = \text{actions}_i(y)$, the action played by i after y' must be a , which implies $(y', a) \preceq y$, and concludes the proof. \square

Lemma 2.17. *The payoff function π of a LCL game is measurable on the σ -algebra Σ corresponding to the game.*

Proof. We prove that, for every player i , and for every $a \in \mathbb{R}$, $\pi_i^{-1}(]a, +\infty[) \in \Sigma$, which implies that π is measurable on Σ . In LCL game, we have $\pi_i : Z \mapsto [0, 1]$. For every $a < 0$, we have $\pi_i^{-1}(]a, +\infty[) = Z \in \Sigma$. Similarly, for every $a > 1$, we have $\pi_i^{-1}(]a, +\infty[) = \emptyset \in \Sigma$. So, let us assume that $a \in]0, 1]$, and let z be a terminal history such that $\pi_i(z) > a$. We have $\text{time}_i(z) < \ln a / \ln \delta$, i.e., every player in $\text{ball}(i)$ has played only a finite number of times in the history z . Let x be the longest history such that $x \preceq z$, and $r(x) = \text{time}_i(z)$. By this setting, the history x' that comes right after x in z is the shortest prefix of z satisfying that every player in $\text{ball}(i)$ is inactive. Let z' be a terminal history such that $x' \preceq z'$. Since every player $j \in \text{ball}(i)$ is inactive after x' , it follows that the state of any such player in z' is the same as its state in z , and thus $\pi_i(z') = \pi_i(z)$. It follows from the above that, for any terminal history z such that $\pi_i(z) > a$, there exists a finite history x' in round $\text{time}_i(z) + 1$ such that $z \in Z_{x'} \subseteq \pi_i^{-1}(]a, +\infty[)$. Since there are finitely many histories in round $\text{time}_i(z)$, we get that $\pi_i^{-1}(]a, +\infty[)$ is the union of a finite number of sets of the form $Z_{x'}$. As a consequence, it is measurable in Σ . It remains to prove that $\pi_i^{-1}(]0, +\infty[) \in \Sigma$. This simply follows from the fact that

$$\pi_i^{-1}(]0, +\infty[) = \bigcup_{k \geq 1} \pi_i^{-1}(] \frac{1}{k}, +\infty[),$$

and from the fact that Σ is stable by countable unions. \square

Lemma 2.18. *LCL games are continuous.*

Proof. Let b be a strategy profile, and let $(b^k)_{k \geq 0}$ be a sequence of strategy profiles such that $d(b^k, b) \rightarrow 0$ when $k \rightarrow \infty$. By definition of the metric d on B (cf. subsection 2.2.5), we have that $d(\rho_{b^k}, \rho_b) \rightarrow 0$ when $k \rightarrow \infty$. By definition of the metric on O , we have that, for any finite history x , $|\rho_{b^k}(x) - \rho_b(x)| \xrightarrow[k \rightarrow \infty]{} 0$. It follows that, for any set of the form Z_x as defined in subsection 2.2.3, $|\mu_{b^k}(Z_x) - \mu_b(Z_x)| \xrightarrow[k \rightarrow \infty]{} 0$. In other words the sequence of measures μ_{b^k} strongly converges to μ_b . Since, for every player i , the function π_i is measurable and bounded, it follows that

$$\int_{\Sigma} \pi_i d\mu_{b^k} \xrightarrow[k \rightarrow \infty]{} \int_{\Sigma} \pi_i d\mu_b.$$

Therefore, $\Pi_i(b^k) \xrightarrow[k \rightarrow \infty]{} \Pi_i(b)$, and thus the expected payoff function Π is continuous. \square

Lemmas 2.15-2.18 show that every LCL game satisfies the requirements of Lemma 2.2, that is, every LCL game satisfies Lemma 2.3. \square

2.5 Conclusion and Further Work

In this chapter, we have proved that natural games occurring in the framework of local distributed network computing have trembling-hand perfect equilibria, a strong form of Nash equilibria. Further study includes the analysis of the performances of the robust algorithms resulting from these equilibria. This study is challenging as determining the performances of iterative distributed construction algorithms such as the generic algorithm in Section 2.1.1 is non trivial, even if nodes are altruistic, and follow the prescribed actions imposed by the algorithm. On the other hand, this line of study is of the utmost importance as, in the framework of large scale distributed computing, it is unreasonable to assume that no nodes will be tempted to deviate from the prescribed actions, for optimizing its own benefit, at the expense of the performances of the algorithms, and of the quality of the solutions.

Chapter 3

The impact of Competition on Coverage in a Dispersal Game

In this chapter, we consider a game-theoretic setting in which selfish individuals compete over resources of varying quality. The motivating example is a group of animals that disperse over patches of food of different abundances. In such scenarios, individuals are biased towards selecting the higher quality patches, while, at the same time, aiming to avoid costly collisions or overlaps. Our goal is to investigate the impact of collision costs on the parallel coverage of resources by the whole group.

Consider M sites, where a site x has value $f(x)$. We think of $f(x)$ as the reward associated with site x , and assume that if a single individual visits x exclusively, it receives this exact reward. Typically, we assume that if $\ell > 1$ individuals visit x then each receives at most $f(x)/\ell$. In particular, when competition costs are high, each individual might receive an amount strictly less than $f(x)/\ell$, which could even be negative. Conversely, modeling cooperation at a site, we also consider cases where each one gets more than $f(x)/\ell$. There are k identical players that compete over the rewards. They independently act in parallel, in a one-shot scenario, each specifying a single site to visit, without knowing which sites are explored by others. The group performance is evaluated by the expected coverage, defined as the sum of $f(x)$ over all sites that are explored by at least one player. Since we assume that players cannot coordinate before choosing their site we focus on symmetric strategies.

The main takeaway message of this chapter is that the optimal symmetric coverage is expected to emerge when collision costs are relatively high, so that the following “Judgment of Solomon” type of rule holds: If a single player explores a site x then it gains its full reward $f(x)$, but if several players explore it, then neither one receives any reward. Under this policy, it turns out that there exists a unique symmetric Nash Equilibrium strategy, which is, in fact, evolutionary stable. Moreover, this strategy yields the best possible coverage among all symmetric strategies. Viewing the coverage measure as the social welfare, this policy thus enjoys a (*Symmetric*) *Price of Anarchy* of precisely 1, whereas, in fact, any other congestion policy has a price strictly greater than 1.

Our model falls within the scope of mechanism design, and more precisely in the area of incentivizing exploration. It finds relevance in evolutionary ecology, and further connects to studies on Bayesian parallel search algorithms.

3.1 Introduction

Studying the way humans and other animals distribute themselves in their environment is a cornerstone of ecology and the social sciences [13, 38, 41, 45]. In many of these contexts, dispersal is governed by two contradicting forces. On the one hand, individuals are biased towards selecting the higher quality patches, and, on the other hand, they aim to avoid costly collisions or overlaps, which can significantly deteriorate the value of a patch [34, 49, 91]. In nature, collision costs can be caused by various factors, including aggressive behavior, or merely due to sharing the patch equally between colliding individuals (a.k.a., scramble competition [71]). The purpose of this chapter is to investigate the impact of collision costs on the overall coverage of resources by the whole group.

Understanding the impact of collisions on the overall parallel coverage can also have implications to the way centralized entities incentive individuals in human organizations. For example, research foundations are often interested in promoting specific research topics, and offer grants to attract researchers to such topics. The question of how to properly define a grant policy that is expected to yield a desirable distribution of researchers can have significant implication on the scientific community [50].

This chapter proposes a framework to study the impact of collision on the group performances through an *algorithmic mechanism design* approach. We focus on a relatively simple scenario, which already provides interesting, and perhaps surprising, insights. See Section 3.5.1 for suggested generalizations left for future work.

3.1.1 The Dispersal Game

Think about the following imaginary scenario. A group of individuals are searching for resources in some space. Eventually, they gather all resources found for the benefit of the public, and hence the group performances is defined as the sum of the resources values. In order to avoid the free-rider problem, the group incentivizes individuals by granting those who find more resource with higher social status. When an individual exclusively finds a resource of value f , a simple mechanism is to let its social status be proportional to f . However, how should the group define the social status of each of ℓ individuals in case they all found a resource simultaneously?

Formally, we have M sites indexed by $x \in [1, M]$. Each site x has some *importance value* $f(x)$, and we assume without loss of generality that sites are ordered such that lower index sites have higher values, i.e., $f(x) \geq f(x + 1)$ for each $x < M$.

We have k players that act a one-shot scenario, each specifying a cite x to visit. Crucially, they cannot coordinate, and each player must commit to a site x , without knowing which sites are selected by other players [32]. Moreover, we assume that players cannot change their mind after committing. Formally, each player specifies an index x according to some law, called *strategy*. This is simply a probability distribution, indicating, for each x , the probability $p(x)$ to explore x . A *strategy profile* is a vector of k strategies, one for each player. It is called *symmetric* when all players play the same strategy. For short, we shall use the term *symmetric strategy* to refer to the strategy being played in the case of a symmetric strategy profile.

The group performance is evaluated by the *weighted coverage*, defined as the sum of

the values $f(x)$ over visited sites x . With the lack of coordination between users, the most the group can hope for is the best achievable coverage among all symmetric strategies, that is, when assuming that all players play the same strategy. Formally, the *coverage* of a strategy p is defined as:

$$\text{Cover}(p) = \sum_{x=1}^M f(x)(1 - (1 - p(x))^k) \quad (3.1)$$

Let p^* be a strategy that maximizes $\text{Cover}(p)$. Using a compactness argument, it is not difficult to show that such a strategy exists. The following observation implies that the optimal symmetric strategy p^* has coverage which is, up to a small constant factor, as good as the best achievable coverage in the full coordination case, when the k users can be assigned to precisely cover the k highest value sites.

Observation 3.1. $\text{Cover}(p^*) > (1 - \frac{1}{e}) \sum_{x=1}^k f(x)$.

Proof. Let \hat{p} be the uniform distribution in $[k]$, that is, we have $\hat{p}(x) = 1/k$ for each $x \leq k$. Then, $\text{Cover}(p^*) \geq \text{Cover}(\hat{p}) = \sum_{x=1}^k f(x) \left(1 - \left(1 - \frac{1}{k}\right)^k\right) > (1 - \frac{1}{e}) \sum_{x=1}^k f(x)$. \square

A *reward policy* is a function $I(x, \ell)$ specifying the *payoff* (or reward) that a player exploring x receives, given that $\ell - 1$ other players are also exploring x . We shall be particularly interested in *congestion* reward policies, which can be written as:

$$I(x, \ell) = f(x) \cdot C(\ell)$$

where $C(\cdot)$ is some *congestion* function, for which $C(1) = 1$, and C is non-increasing. Note that the specification of the policy does not depend on the total number of players k , and is influenced only by the number of players that visit x .

The assumption that $C(1) = 1$ implies that a player visiting x exclusively will gain exactly¹ $f(x)$. A natural example of a congestion policy is the *sharing policy* in which $C_{\text{share}}(\ell) = 1/\ell$, and hence if ℓ players select site x in parallel then each receives a payoff of $f(x)/\ell$. This policy has received a lot of attention in both the ecology literature and algorithmic game theory community due to its simplicity [49, 50, 91]. Cases where $C(\ell) > 1/\ell$ can model cooperation, in the sense that the presence of several players at a site would result in each having more than their relative share. On the other extreme, note that we do not restrict C to be a positive function. Taking $C(\ell)$ to be negative implies that the presence of ℓ players at x damages each of them quite significantly. In the context of animals, this could represent aggressive behavior that increases with the amount of potential benefit $f(x)$ and could result in severe injuries.

¹Alternatively, we could have defined $C(1)$ to be any constant that doesn't depend on x . This would mean that the reward a single player gets when visiting x exclusively is proportional to the value $f(x)$. (This assumption contrasts with the setting of [50], see more details in Section 3.1.6.) Note that in this sense, $f(x)$ plays a double role - it has a meaning from the perspective of the group and a meaning from the perspective of an individual.

Let X_i be the site specified by player i . Given a policy I , the goal of each player i is to maximize the *expected payoff* it receives, defined as

$$\mathbb{E}\left[I(X_i, \sum_j \mathbb{1}_{X_j=X_i})\right]$$

3.1.2 Evaluating Policies

The game we define is symmetric and possesses both mixed symmetric equilibria and pure non-symmetric equilibria. Most previous works dealing with such situations prefer to focus on pure equilibria (rather than on the mixed equilibria), because of their conceptual simplicity. The classical example of this situation is the well-know game “battle of the sexes” [73]. However we argue otherwise: the number of pure equilibria grows exponentially with the number of players, and choosing an equilibrium among those requires coordination between the players. In some contexts, because coordination is not possible or communication is too costly, it does not seem reasonable to expect that a pure equilibrium arises in practice, and instead, it is more natural to expect the emergence of a symmetric equilibrium. Since we focus on large distributed systems, where players do not always possess individual identities and often have identical behaviors, we will restrict our attention to symmetric equilibria, and disregard pure equilibria².

Viewing p^* as the symmetric strategy that maximizes the social welfare, we adopt an interpretation of the concept of Price of Anarchy to evaluate a policy $I(x, \ell)$, by measuring the coverage of the worst symmetric Nash Equilibrium under $I(x, \ell)$ in comparison with the coverage of p^* . Specifically, let $C(\ell)$ be a congestion function, and $f(x)$ be an importance value function. Define

$$\text{SPoA}(C, f) := \sup_{p \in \mathcal{P}(C, f)} \left\{ \frac{\text{Cover}(p^*)}{\text{Cover}(p)} \right\},$$

where $\mathcal{P}(C, f)$ is the set of symmetric Nash Equilibrium under $I(x, \ell) = f(x) \cdot C(\ell)$. The *Symmetric Price of Anarchy* (SPoA) of the congestion function $C(\ell)$ is defined as the sup of $\text{SPoA}(C, f)$ over all functions $f : [M] \rightarrow \mathbb{R}^+$, where we take the freedom to choose their domain M as well, i.e.,

$$\text{SPoA}(C) := \sup_{\substack{M \in \mathbb{N} \\ f: [M] \rightarrow \mathbb{R}^+}} \{\text{SPoA}(C, f)\}$$

By definition, $\text{SPoA}(C) \geq 1$ for any congestion function C . On the other hand, there exist congestion functions with much higher SPoA. For example, taking $C(\ell) \equiv 1$ yields SPoA

²At first glance, our focus on symmetric strategies may seem non-compatible with the fact that conspecific animals often appear not to have identical strategies. However, it is important to note that a symmetric strategy profile in the context of ESS (see Section 3.1.4) does not imply that every player follows the same strategy! Indeed, a symmetric strategy profile can arise also if the population contains pure strategies in proportions as specified by the mixed strategy. In this case, since the framework of ESS assumes that at each time k players are selected at random from the population, each of the selected random players ends up having the same mixed strategy. For this reason, restricting attention to symmetric equilibria is in fact very common in the ecology literature [34, 49, 91].

of roughly k . Indeed, under this function, and assuming a strictly decreasing f , a strategy at equilibrium would explore the first site with probability 1, and for distributions f that decrease very slowly (and thus are close to uniform), this would yield a gap $\Omega(k)$, with respect to, e.g., the uniform strategy. Note, however, that this policy is unlikely to occur in nature, as it implies that a value at a site is fully obtained by all visitors.

The PoA of the sharing policy C_{share} was studied by Kleinberg and Oren in a somewhat similar model [50] (see also Section 3.1.6). Thanks to a more general result from Vetta [92], they showed that the SPoA of the sharing policy, defined as the ratio between the best achievable coverage and the worst coverage among Nash equilibria, is at most 2. Since the worst symmetric equilibrium coverage is at least as good as the worst equilibrium coverage, and the best achievable coverage is at least as good as the best achievable coverage with a symmetric strategy profile, this directly implies that our notion of SPoA also satisfies $\text{SPoA}(C_{share}) \leq 2$.

3.1.3 Ideal Free Distribution (IFD)

The setting of animals competing over patches of resources, often referred to as *playing the field*, has been extensively studied in the ecology discipline through the theory of *Ideal Free Distribution (IFD)*, see reviews in [49, 91]. Introduced by Fretwell and Lucas [34], the standard model assumes that each individual follows the same distribution $p(\cdot)$ (i.e., a symmetric strategy) and hence the fraction of the population that is expected to occupy a site x is $p(x)$.

An IFD is a probability distribution $p(\cdot)$ in which every site which is explored with positive probability yields the same payoff, and every other site yields a lower payoff if explored. More precisely, as before, let $I(x, \ell)$ denote the payoff received by an individual selecting a site x together with $\ell - 1$ other individuals. Fix a player and let $P(x, \ell) = \binom{k-1}{\ell} p(x)^\ell (1 - p(x))^{k-\ell-1}$ be the probability that among $k - 1$ players, ℓ of them selected site x . The *value_p* of a site x corresponds to the expected gain for exploring site x and is defined as:

$$\nu_p(x) := \sum_{\ell=1}^k I(x, \ell) \cdot P(x, \ell - 1) \quad (3.2)$$

By definition, the expected payoff that a player receives is $\sum_x p(x) \cdot \nu_p(x)$. The IFD is a distribution p , that for some integer W , satisfies the following *IFD conditions*.

- $p(x) > 0 \iff x \leq W$
- $\nu_p(1) = \nu_p(2) = \dots = \nu_p(W) = \nu_p$, and for all $W + 1 \leq x$, we have $\nu_p > \nu_p(x)$

The following observation is well-known, see e.g., [21, 34]. For the sake of completeness, a sketch of the proof is provided below.

Observation 3.2. *Assume that $I(x, \cdot)$ is non-increasing for any x (note that this is indeed the case for congestion policies). There exists a unique strategy satisfying the IFD conditions. Moreover, this IFD is the only symmetric Nash Equilibrium of the game.*

Proof Sketch The existence of the IFD follows by the fact that $\nu_p(x)$ is non-increasing with $p(x)$. By definition, the IFD is a symmetric Nash Equilibrium of the game. We now prove its uniqueness. Imagine there are two different symmetric Nash Equilibria, π and π' , respectively associated with their values ν_π and $\nu_{\pi'}$. Note that for any fixed x , the value $\nu_p(x)$ is a strictly decreasing function of $p(x)$. Therefore

$$\pi = \pi' \iff \nu_\pi = \nu_{\pi'}$$

Assume by contradiction that $\nu_\pi > \nu_{\pi'}$. This implies that for any $x \in [\pi]$, $\pi(x) < \pi'(x)$, where $[\pi]$ denotes the support of π , i.e the set of sites explored with positive probability. Summing over all sites in $[\pi]$, we get

$$\sum_{x \in [\pi]} \pi(x) < \sum_{x \in [\pi]} \pi'(x)$$

Since the left-hand side is equal to 1, we get a contradiction. This proves the uniqueness of the symmetric Nash Equilibrium, and of the IFD. \square

3.1.4 Evolutionary Stable Strategy (ESS)

In addition to the notion of Nash Equilibrium, we shall adopt the stronger concept of an *Evolutionary Stable Strategy (ESS)* [12–14, 40, 62, 88]. This concept has become a cornerstone of evolutionary game theory, and has been extensively used to study the evolution of animal and human behavior [13, 62]. Informally, an ESS is a strategy σ such that, in a population composed of a majority of σ -strategists and a minority of “mutants” playing strategy π , the σ -strategists have a higher payoff, provided that the proportion of mutants is not too high.

The standard view of ESS is to consider an infinite population whose individuals play repeatedly against each other in pairs, chosen uniformly at random. In our setting, individuals play in groups of k , and hence a generalization is required. One possible generalization is to consider a fixed population of k individuals that play the game between themselves. Unfortunately, generalizing ESS to finite populations is quite challenging, and the known generalizations are often too restrictive to be applicable. Instead, we chose the generalization in which the population remains infinite and players are randomly matched in k -tuples, for any fixed $k \geq 2$, see e.g., the book [13]. Modeling biological settings, this generalization may correspond, for example, to scenarios in which a large colony of bats breaks on a daily basis into smaller groups, each of which forages in a different field of patched resources, see [24, 63].

Formally, let $\mathcal{E}(\rho; \sigma^\ell, \pi^{k-\ell-1})$ denote the payoff received by an individual playing strategy ρ against ℓ players playing strategy σ and $(k - \ell - 1)$ players playing strategy π . When $\sigma = \pi$, we use the abridged notation $\mathcal{E}(\rho; \sigma^{k-1})$. Consider a population in which a proportion $(1 - \epsilon)$ of it are playing strategy σ and a proportion of ϵ are playing π . The average payoff of a player playing ρ matched against $k - 1$ opponents drawn uniformly at random from this population is denoted by $\mathcal{U}[\rho; (1 - \epsilon)\sigma + \epsilon\pi]$, and is equal to

$$\mathcal{U}[\rho; (1 - \epsilon)\sigma + \epsilon\pi] = \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} (1 - \epsilon)^\ell \epsilon^{k-\ell-1} \mathcal{E}(\rho; \sigma^\ell, \pi^{k-\ell-1}) \quad (3.3)$$

A strategy σ is an ESS if for every other strategy π , there exist $\epsilon_\pi > 0$ such that for every $\epsilon \leq \epsilon_\pi$, $\mathcal{U}[\sigma, (1 - \epsilon)\sigma + \epsilon\pi] > \mathcal{U}[\pi, (1 - \epsilon)\sigma + \epsilon\pi]$, i.e., σ cannot be “invaded” by π , as long as the proportion of π -strategists remains below ϵ_π .

We also recall the following *characterization of ESS* [13]: a strategy σ is an ESS if for every other strategy π , there is an integer $0 \leq m_\pi \leq k - 1$, such that both conditions below hold:

- $\mathcal{E}(\sigma; \sigma^{k-m_\pi-1}, \pi^{m_\pi}) > \mathcal{E}(\pi; \sigma^{k-m_\pi-1}, \pi^{m_\pi})$
- $\forall \ell, 0 \leq \ell < m_\pi$, we have $\mathcal{E}(\sigma; \sigma^{k-\ell-1}, \pi^\ell) = \mathcal{E}(\pi; \sigma^{k-\ell-1}, \pi^\ell)$

It should be clear from this definition that every ESS is a symmetric Nash equilibrium of the k -player game. Indeed, these conditions ensure that $\mathcal{E}(\sigma; \sigma^{k-1}) \geq \mathcal{E}(\pi; \sigma^{k-1})$ for every $\pi \neq \sigma$. Further details about the ESS notion appear in Section 3.5.1.

3.1.5 Our Results

We study the interplay between the congestions and the expected coverage of strategies at equilibrium. Our main takeaway message is that optimal coverage is expected to emerge by a congestion policy which is defined by a “Judgment of Solomon” type of rule: If a single player attempts to exploit a site then it gains the full reward $f(x)$, but when several players attempt to exploit it, they all receive nothing. Formally, define the *exclusive congestion function* as follows:

$$C_{exc}(\ell) = \begin{cases} 1 & \text{if } \ell = 1 \\ 0 & \text{otherwise} \end{cases}$$

and define the *exclusive reward policy* as

$$I_{exc}(\ell, x) = f(x) \cdot C_{exc}(\ell)$$

We identify I_{exc} as a critical reward policy, in the sense that its IFD yields optimal coverage, while any other congestion policy yields strictly worse coverage.

More formally, let σ^* be the IFD for the exclusive reward policy. We first claim that under this policy, σ^* is not only a Nash Equilibrium but also an ESS.

Theorem 3.3. σ^* is an ESS under I_{exc} .

Next, we claim that σ^* yields the best possible coverage, among all symmetric strategies.

Theorem 3.4. For any strategy σ , $\text{Cover}(\sigma^*) \geq \text{Cover}(\sigma)$, with equality if and only if $\sigma = \sigma^*$.

Since the IFD is the only symmetric Nash Equilibrium under I_{exc} , the price of anarchy equals 1.

Corollary 3.5. $\text{SPoA}(C_{exc}) = 1$.

The criticality of I_{exc} follows by the fact that among all congestion policies, it is the only one whose IFD yields optimal coverage for all value functions f .

Theorem 3.6. *For any congestion function $C \neq C_{exc}$, we have $\text{SPoA}(C) > 1$.*

We find the fact that $\text{SPoA}(C_{exc}) = 1$ rather surprising. Indeed, although it appears intuitive that increasing the competition will result in better coverage, the exclusive policy C_{exc} is, in fact, not the one with highest levels of competition. Indeed, one could define a congestion policy where in case of conflict, each of the colliding players receives a negative payoff (i.e., $C(\ell)$ being negative, see Section 3.1.1). This means that the competition level could significantly exceed the level of C_{exc} . The fact that the best coverage occurs exactly at that level is thus quite surprising. A second important factor to note is that even if one identifies the best policy (i.e., the one for which the SPoA is smallest), it is a priori unclear that the resulted coverage would actually be optimal among all symmetric strategies (including the non-competing ones), i.e., that $\text{SPoA} = 1$.

An informal discussion regarding the implications of our results in the context of animal behavior is presented in the conclusion (section 3.5.2).

See Figure 3.1 for an illustration of the coverage as a function of the competition extent, in specific instances of 2 players and 2 sites.

3.1.6 Related Work

A research theme that is currently quite popular in the Operations Research, and Economic Theory research communities goes by the name “incentivizing exploration”. Similarly to us, such articles tend to focus on the problem of designing mechanisms to coordinate the activities of independent selfish agents exploring a space of alternatives [33, 53, 60, 75]. These articles however typically focus on sequential rather than simultaneous exploration, leading them to model the problem as a multi-armed bandit problem rather than a congestion game. For this reason, there is no real technical connection between the two lines of work, but there is nevertheless a thematic connection.

Closest to our work is the article by Kleinberg and Oren [50] which considers a general model that is motivated by research foundations offering grants to incentivize researchers to work on particular topics. That article proposes two mechanisms for selecting grant policies that are expected to incentivize researchers to adopt an optimal distribution, while assuming that researchers are restricted to the sharing policy. The first mechanism relies on the ability to freely choose the rewards (grants) associated with sites (topics), despite the fact that the significance of a site from the perspective of the group is fixed. In contexts of animals, however, the overall rewards (termed here as $f(x)$) correspond to the amount of food in patch x and therefore cannot be modified. Moreover, implementing their mechanism of choosing rewards requires the knowledge of the number of players k , whereas the specification of the congestion policies studied here does not require this knowledge (see Section 3.1.1). The second mechanism proposed in [50] requires that the players receive different rewards, set in a particular way by the central entity, despite doing the same actions. This might be seen as unfair and unsatisfying. Moreover, implementing this would require high levels of coordination which again seem unlikely in the context of animals.

The settings of selfish routing, job scheduling, and congested games [64, 84] all bear similarities to the dispersal game, however, the measurement of the global performances considered here, namely, the coverage, is very different from the measures studied in the former frameworks, such as makespan or latency [5, 9, 22, 72].

Finally, many articles have informally mentioned the IFD as an example of an ESS [49, 65, 76, 77, 80, 81]. However, on a formal level, it is only recently that a rigorous proof was given [20], and even this proof was with respect a limited definition of ESS (essentially, in our notation, they proved the first item of the characterization with $m_\pi = k - 1$, without considering the second item). Since that article, several other works have also addressed this question rigorously, but, to the best of our knowledge, they all considered either weaker versions than our ESS definition, or other contexts, such as continuous populations dynamics [19, 82].

3.2 The Exclusive Reward Policy

This section considers the exclusive reward policy I_{exc} . In Section 3.2.1 we explicitly calculate the IFD for this policy, called σ^* . Then, in Section 3.2.2, we prove that σ^* yields optimal coverage.

3.2.1 Algorithm σ^*

Consider a strategy p that satisfies the IFD conditions for the exclusive reward policy. We know that it exists by Observation 3.2 and we calculate it explicitly. In the context of C_{exc} , the *value* _{p} of a site x , as given by Eq. (3.2), is:

$$\nu_p(x) = f(x)(1 - p(x))^{k-1}$$

The second IFD condition implies that for any two $x, y \in [1, W]$, we have $f(x)(1 - p(x))^{k-1} = f(y)(1 - p(y))^{k-1}$, or in other words,

$$\frac{1 - p(x)}{1 - p(y)} = \left(\frac{f(y)}{f(x)} \right)^{1/(k-1)} \quad (3.4)$$

A natural guess for a distribution that satisfies this is the following Pareto distribution:

$$p(x) = \begin{cases} 1 - \frac{\alpha}{f(x)^{\frac{1}{k-1}}} & \text{if } x \leq W \\ 0 & \text{otherwise} \end{cases}$$

Since p is a distribution, we must have $\sum_{x=1}^W p(x) = 1$, from which we can extract

$$\alpha = \frac{W - 1}{\sum_{x \leq W} f(x)^{-\frac{1}{k-1}}}.$$

To complete the description of our candidate IFD strategy, it remains to calculate W . By the second IFD condition, we get that if $W < M$, then for every $x \leq W$:

$$f(x)(1 - p(x))^{k-1} = \nu_p(x) > \nu_p(W + 1) = f(W + 1) \quad (3.5)$$

Rearranging, we get

$$p(x) < 1 - \left(\frac{f(W+1)}{f(x)} \right)^{\frac{1}{k-1}}.$$

Since $p(x)$ is a distribution whose support is $[W] := \{1, \dots, W\}$ we get

$$1 = \sum_{x \leq W} p(x) < \sum_{x \leq W} \left(1 - \left(\frac{f(W+1)}{f(x)} \right)^{\frac{1}{k-1}} \right).$$

This means that we can define W as the first index that satisfies the equation above, if one exists, or M otherwise. Equivalently, we can define W as the largest index that satisfies

$$\sum_{x \leq W} \left(1 - \left(\frac{f(W)}{f(x)} \right)^{\frac{1}{k-1}} \right) \leq 1$$

Altogether, this leads to strategy σ^* , as defined below.

$$\sigma^*(x) = \begin{cases} 1 - \frac{\alpha}{f(x)^{\frac{1}{k-1}}} & \text{if } x \leq W \\ 0 & \text{otherwise} \end{cases}$$

where W and the normalization factor α are calculated as follows.

$$W = \arg \max_{y \leq M} \left\{ \sum_{x \leq y} \left(1 - \left(\frac{f(y)}{f(x)} \right)^{\frac{1}{k-1}} \right) \leq 1 \right\}$$

$$\alpha = \frac{W-1}{\sum_{x \leq W} f(x)^{-\frac{1}{k-1}}}$$

Pseudocode of Algorithm $\sigma^*(x)$

Interestingly, it turns out that algorithm σ^* is actually identical to the first round in the algorithm A^* used in [51] in the context of Bayesian search.

The following claim verifies that the constructed strategy is indeed the IFD.

Claim 3.7. σ^* satisfies the IFD conditions under the exclusive reward policy I_{exc} .

Proof. The first IFD condition is satisfied by σ^* by construction. It remains to show that $\nu_p(1) = \nu_p(2) = \dots = \nu_p(W)$, and for all $x > W$, we have $\nu_p(W) > \nu_p(x)$. Under the exclusive reward policy I_{exc} , the *value_p* of a site $x \leq W$, as given by Eq. (3.2), is:

$$\nu_p(x) = f(x)(1 - \sigma^*(x))^{k-1} = \alpha^{k-1}$$

which proves the first part. For the second part, we first derive $f(W + 1) < \alpha^{k-1}$ as follows:

$$\begin{aligned}
\sum_{x \leq W+1} \left(1 - \left(\frac{f(W+1)}{f(x)} \right)^{\frac{1}{k-1}} \right) &> 1 \\
\sum_{x \leq W} \left(1 - \left(\frac{f(W+1)}{f(x)} \right)^{\frac{1}{k-1}} \right) &> 1 \\
\sum_{x \leq W} \left(\frac{f(W+1)}{f(x)} \right)^{\frac{1}{k-1}} &< W - 1 \\
f(W+1)^{\frac{1}{k-1}} &< \frac{W-1}{\sum_{x \leq W} f(x)^{-\frac{1}{k-1}}} \\
f(W+1) &< \alpha^{k-1}
\end{aligned}$$

Now it suffices to observe that the *value_p* of a site $x > W$ is $\nu_p(x) = f(x) \leq f(W+1) < \alpha^{k-1} = \nu_p(W)$. \square

As claimed in Theorem 3.3, it turns out that under the exclusive policy I_{exc} , strategy σ^* is not only a Nash Equilibrium, but in fact an ESS.

3.2.2 σ^* has Optimal Coverage

In this subsection we prove Theorem 3.4, that is, we prove that for any strategy σ , $\text{Cover}(\sigma^*) \geq \text{Cover}(\sigma)$, with equality if and only if $\sigma = \sigma^*$. Before dwelling into the proof we note that algorithm A^* was shown in [51] to be optimal for the setting in which k independent searchers are searching for a treasure placed in one of M boxes, according to a Bayesian distribution. The proof therein is long and relies on deep techniques that were developed in [32]. Since σ^* agrees with A^* on the first round of A^* and since there are some similarities between the search and the coverage objectives, it might be possible to translate the proof in [32, 51] to show the optimality of σ^* in terms of coverage. However, we managed to find a simple and concise proof of this fact, that can also be used to show the uniqueness of σ^* . We provide it below.

Note that maximizing $\text{Cover}(p)$ is equivalent to minimizing $\mathcal{T}(p) = \sum_x f(x)(1-p(x))^k$. We next show that σ^* minimizes the latter expression.

First observe that since the number of resources is finite, it follows by a compactness argument that the infimum of $\{\mathcal{T}(\sigma) \mid \sigma \text{ is a distribution}\}$ is attained by some distribution σ_{\min} . Assume by contradiction that $\sigma_{\min} \neq \sigma^*$. Since both σ_{\min} and σ^* are distributions, we must have some $x_0 \leq W$ such that $\sigma^*(x_0) > \sigma_{\min}(x_0)$ and some x_1 such that $\sigma^*(x_1) < \sigma_{\min}(x_1)$. For sufficiently small $\epsilon > 0$, let us define $\sigma'(x)$ which equals σ_{\min} everywhere except that $\sigma'(x_1) = \sigma_{\min}(x_1) - \epsilon$ and $\sigma'(x_0) = \sigma_{\min}(x_0) + \epsilon$. In other words, we create σ' from σ_{\min} by shifting a small mass of probability from x_1 to x_0 .

Aiming to contradict the minimality of σ_{\min} , our goal is to show that $\mathcal{T}(\sigma') < \mathcal{T}(\sigma_{\min})$. In other words, we aim to show that $\mathcal{T}(\sigma_{\min}) - \mathcal{T}(\sigma')$ is positive, and we rewrite this

expression as follows:

$$\begin{aligned}\mathcal{T}(\sigma_{\min}) - \mathcal{T}(\sigma') &= f(x_0) \left((1 - \sigma_{\min}(x_0))^k - (1 - \sigma_{\min}(x_0) - \epsilon)^k \right) \\ &\quad + f(x_1) \left((1 - \sigma_{\min}(x_1))^k - (1 - \sigma_{\min}(x_1) + \epsilon)^k \right)\end{aligned}$$

By a Taylor expansion, for a small $\epsilon > 0$, we obtain:

$$\mathcal{T}(\sigma_{\min}) - \mathcal{T}(\sigma') \approx \epsilon k \left(f(x_0)(1 - \sigma_{\min}(x_0))^{k-1} - f(x_1)(1 - \sigma_{\min}(x_1))^{k-1} \right) \quad (3.6)$$

By equation (3.5) and the definitions of x_0 and x_1 , we know that

$$\begin{aligned}f(x_0)(1 - \sigma_{\min}(x_0))^{k-1} &> f(x_0)(1 - \sigma^*(x_0))^{k-1}, \\ f(x_0)(1 - \sigma^*(x_0))^{k-1} &\geq f(x_1)(1 - \sigma^*(x_1))^{k-1}, \\ \text{and } f(x_1)(1 - \sigma^*(x_1))^{k-1} &> f(x_1)(1 - \sigma_{\min}(x_1))^{k-1}.\end{aligned}$$

Hence, for sufficiently small $\epsilon > 0$, we get that $\mathcal{T}(\sigma_{\min}) > \mathcal{T}(\sigma')$, contradicting the minimality of σ_{\min} . This establishes the fact that $\sigma^* = \sigma_{\min}$, implying that σ^* maximizes the coverage $\mathbf{Cover}(p)$.

Finally, the fact that $\sigma^* = \sigma_{\min}$ also implies that the distribution that minimizes $\mathcal{T}(p)$ (and hence maximizes $\mathbf{Cover}(p)$) is unique. Hence, $\mathbf{Cover}(\sigma^*) = \mathbf{Cover}(\sigma)$ can occur only if $\sigma = \sigma^*$. This concludes the proof of Theorem 3.4. \square

3.3 The Ideal Free Distribution is an ESS

The goal of this section is to prove Theorem 3.3, i.e, we aim to show that under I_{exc} , strategy σ^* is an ESS.

Let σ be a strategy, distinct from σ^* . First note that, when playing against strategy σ^* , because it is the IFD, the expression of $\mathcal{E}_k(\sigma; \sigma^{*(k-1)})$ becomes :

$$\mathcal{E}_k(\sigma; \sigma^{*(k-1)}) = \sum_{x=1}^M \sigma(x) \nu_{\sigma^*}(x) = \sum_{x=1}^W \sigma(x) \nu_{\sigma^*}(W) + \sum_{x=W+1}^M \sigma(x) f(x) \quad (3.7)$$

Since for every site $x > W$, we have $f(x) < \nu_{\sigma^*}(W)$, it is clear that every strategy σ whose support is not a subset of $[1, W]$ has

$$\mathcal{E}_k(\sigma; \sigma^{*(k-1)}) < \mathcal{E}_k(\sigma^*; \sigma^{*(k-1)})$$

and therefore satisfies the ESS characterization conditions with $m_{\sigma} = 0$.

Next, let us assume that the support of σ is included in $[1, W]$, and show that the ESS characterization conditions are satisfied for $m_{\sigma} = 1$, i.e., we aim to show that

$$\begin{aligned}\forall \ell, 1 \leq \ell \leq k-2, \text{ we have} \\ \mathcal{E}_k(\sigma^*; \sigma^{\ell}, \sigma^{*(k-\ell-1)}) &> \mathcal{E}_k(\sigma; \sigma^{\ell}, \sigma^{*(k-\ell-1)}),\end{aligned} \quad (3.8)$$

$$\text{and } \mathcal{E}_k(\sigma^*; \sigma^{*(k-1)}) = \mathcal{E}_k(\sigma; \sigma^{*(k-1)}). \quad (3.9)$$

In fact, to show that the ESS characterization conditions are satisfied for $m_\sigma = 1$, we need only the case $\ell = 1$ in Inequality (3.8), but we prove here a stronger stability criteria, showing that this inequality holds for every $1 \leq \ell \leq k - 2$.

Eq. (3.9) follows quite simply from the observation above, coupled with the assumption that the support of σ is in $[1, W]$, which gives both $\mathcal{E}_k(\sigma^*; \sigma^{*(k-1)}) = \nu_{\sigma^*}$ and $\mathcal{E}_k(\sigma; \sigma^{*(k-1)}) = \nu_{\sigma^*}$.

Eq. (3.8) requires more elaboration. We first show that for every $1 \leq \ell \leq k - 2$:

$$\mathcal{E}_k(\sigma^*; \sigma^\ell, \sigma^{*(k-\ell-1)}) \geq \mathcal{E}_k(\sigma; \sigma^\ell, \sigma^{*(k-\ell-1)})$$

and then show that equality holds iff $\sigma = \sigma^*$. We start with the following claim.

Claim 3.8. *Expanding each side in Eq. (3.8), we obtain:*

$$\mathcal{E}_k(\sigma^*; \sigma^\ell, \sigma^{*(k-\ell-1)}) = \alpha^{k-\ell-1} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^\ell - \alpha \sum_{x \leq W} f(x)^{\frac{\ell-1}{k-1}} (1 - \sigma(x))^\ell \right), \quad (3.10)$$

$$\mathcal{E}_k(\sigma; \sigma^\ell, \sigma^{*(k-\ell-1)}) = \alpha^{k-\ell-1} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^\ell - \sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} \right) \quad (3.11)$$

Proof of Claim 3.8. We start by expanding the left hand side of Eq. (3.8).

$$\begin{aligned} \mathcal{E}_k(\sigma^*; \sigma^\ell, \sigma^{*(k-\ell-1)}) &= \sum_x f(x) \sigma^*(x) (1 - \sigma^*(x))^{k-\ell-1} (1 - \sigma(x))^\ell \\ &= \sum_{x \leq W} (f(x) - \alpha f(x)^{\frac{k-2}{k-1}}) \alpha^{k-\ell-1} f(x)^{\frac{\ell+1-k}{k-1}} (1 - \sigma(x))^\ell \\ &= \alpha^{k-\ell-1} \sum_{x \leq W} (f(x)^{\frac{\ell}{k-1}} - \alpha f(x)^{\frac{\ell-1}{k-1}}) (1 - \sigma(x))^\ell \\ &= \alpha^{k-\ell-1} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^\ell - \alpha \sum_{x \leq W} f(x)^{\frac{\ell-1}{k-1}} (1 - \sigma(x))^\ell \right) \end{aligned}$$

This establishes the first part of the claim.

Next, we expand the right hand side of Eq. (3.8).

$$\begin{aligned}
\mathcal{E}_k(\sigma; \sigma^\ell, \sigma^{*(k-\ell-1)}) &= \sum_x f(x) \sigma(x) (1 - \sigma^*(x))^{k-\ell-1} (1 - \sigma(x))^\ell \\
&= \sum_{x \leq W} f(x) \sigma(x) (1 - \sigma^*(x))^{k-\ell-1} (1 - \sigma(x))^\ell \\
&= \sum_{x \leq W} f(x) \sigma(x) \alpha^{k-\ell-1} f(x)^{\frac{\ell+1-k}{k-1}} (1 - \sigma(x))^\ell \\
&= \alpha^{k-\ell-1} \sum_{x \leq W} \sigma(x) f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^\ell \\
&= \alpha^{k-\ell-1} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^\ell - \sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} \right)
\end{aligned}$$

Concluding the proof of Claim 3.8.

By Claim 3.8, it follows that the desired inequality is equivalent to:

$$\alpha \sum_{x \leq W} f(x)^{\frac{\ell-1}{k-1}} (1 - \sigma(x))^\ell \leq \sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1}$$

By definition of α , we rearrange this inequality into

$$(W-1) \sum_{x \leq W} f(x)^{\frac{\ell-1}{k-1}} (1 - \sigma(x))^\ell \leq \sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1}$$

Decomposing $f(x)^{\frac{\ell-1}{k-1}} = \frac{f(x)^{\frac{\ell}{\ell+1} \cdot \frac{\ell-1}{k-1}}}{f(x)^{\frac{1}{(\ell+1)(k-1)}}}$, and applying Hölder's Inequality, we obtain:

$$\sum_{x \leq W} f(x)^{\frac{\ell-1}{k-1}} (1 - \sigma(x))^\ell \leq \left(\sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \right)^{\frac{1}{\ell+1}} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} \right)^{\frac{\ell}{\ell+1}}$$

This implies that to prove that Eq. (3.8) holds, it is sufficient to establish the following inequality:

$$(W-1) \left(\sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \right)^{\frac{1}{\ell+1}} \leq \sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} \right)^{\frac{1}{\ell+1}}$$

Using the fact that σ is a distribution, and hence $\sum_{x \leq W} (1 - \sigma(x)) = W - 1$, it suffices to prove that

$$\sum_{x \leq W} (1 - \sigma(x)) \left(\sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \right)^{\frac{1}{\ell+1}} \leq \sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} \right)^{\frac{1}{\ell+1}}$$

Rearranging, this is equivalent to proving that

$$\sum_{x \leq W} (1 - \sigma(x)) \leq \left(\sum_{x \leq W} \frac{1}{f(x)^{\frac{1}{k-1}}} \right)^{\frac{\ell}{\ell+1}} \left(\sum_{x \leq W} f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} \right)^{\frac{1}{\ell+1}} \quad (3.12)$$

Which finally follows by applying Hölder's Inequality.

The fact that the inequality is strict for $\ell \geq 1$ follows from the following.

Claim 3.9. For $\ell \geq 1$, $\mathcal{E}_k(\sigma^*; \sigma^\ell, \sigma^{*(k-\ell-1)}) = \mathcal{E}_k(\sigma; \sigma^\ell, \sigma^{*(k-\ell-1)}) \implies \sigma = \sigma^*$.

Proof of Claim 3.9. The proof of the claim is based on the observation that the former equality implies, in particular, an equality in Hölder's inequality in Eq. (3.12). For $\ell \geq 1$, this can be the case only if the two sides are linearly dependent, i.e., for each $x \leq W$, we have

$$f(x)^{\frac{\ell}{k-1}} (1 - \sigma(x))^{\ell+1} = \frac{\beta^{\ell+1}}{f(x)^{\frac{1}{k-1}}}$$

where β is some fixed constant. Together with the fact that the support of σ is included in $[W]$, the strategy σ must therefore be of the form

$$\sigma(x) = \begin{cases} 1 - \frac{\beta^{\frac{1}{\ell+1}}}{f(x)^{\frac{1}{k-1}}} & \text{if } x \leq W \\ 0 & \text{otherwise} \end{cases}$$

Finally, we know that

$$\sum_{x \leq W} \sigma(x) = \sum_{x \leq W} \left(1 - \frac{\beta^{\frac{1}{\ell+1}}}{f(x)^{\frac{1}{k-1}}} \right) = 1$$

Solving this for β gives

$$\beta = \frac{W - 1}{\sum_{x \leq W} f(x)^{-\frac{1}{k-1}}} = \alpha$$

and therefore $\sigma = \sigma^*$. This completes the proof of Claim 3.9, and thus established the proof for Theorem 3.3. \square

3.4 The Criticality of the Exclusive Collision Cost Function

The goal in this section is to prove Theorem 3.6. That is, fixing a congestion function $C(\cdot) \neq C_{exc}(\cdot)$, we aim to show that $\mathbf{SPoA}(C) > 1$.

First recall that by definition of congestion functions, $C(\cdot)$ is non-increasing. Hence Observation 3.2 applies, and the existence of the IFD is guaranteed for every value function f . Our plan is to show that there exists a value function f , for which this IFD is different than σ^* . Once this is established, the fact that $\mathbf{SPoA}(C) > 1$ follows from Theorem 3.4.

Assume, towards contradiction, that for every f , the IFD under $C(\cdot)$ is precisely σ^* . Let $M \gg k$, and let f be a strictly decreasing function, that, nevertheless, decreases very slowly, such that, e.g., for every $x \leq y$ in $[1, M]$, we have $\frac{f(y)}{f(x)} \geq \frac{f(M)}{f(1)} > \left(1 - \frac{1}{2k}\right)^{k-1}$.

Together with the definition of W (see the pseudocode of Algorithm $\sigma^*(x)$), we obtain that $W \geq 2k$.

Because σ^* is the IFD, we have that $\forall x \leq W$, $\nu_p(x) = \nu_p(W)$. Let us develop this expression:

$$\begin{aligned}
\nu_p(W) &= f(x) \sum_{\ell=1}^k C(\ell) P(x, \ell-1) \\
&= f(x) \sum_{\ell=1}^k C(\ell) \binom{k-1}{\ell-1} \sigma^*(x)^{\ell-1} (1 - \sigma^*(x))^{k-\ell} \\
&= f(x) \sum_{\ell=1}^k C(\ell) \binom{k-1}{\ell-1} \left(1 - \alpha f(x)^{-\frac{1}{k-1}}\right)^{\ell-1} \alpha^{k-\ell} f(x)^{-\frac{k-\ell}{k-1}} \\
&= \sum_{\ell=1}^k C(\ell) \binom{k-1}{\ell-1} \left(1 - \alpha f(x)^{-\frac{1}{k-1}}\right)^{\ell-1} \alpha^{k-\ell} f(x)^{\frac{\ell-1}{k-1}}
\end{aligned}$$

The last line above is a polynomial equation in $f(x)^{\frac{1}{k-1}}$ of degree at most $k-1$. Note that this polynomial is not a constant. Indeed, since $C \neq C_{exc}$ and $C(1) = C_{exc}(1) = 1$, there must exist $\ell \geq 2$, such that $C(\ell) \neq 0$. Hence, the $\ell-1$ 'st coefficient of the polynomial is non-zero, implying that this polynomial is not a constant. It follows that there are at most $k-1$ of values for $f(x)$ that would satisfy this equation (recall that f is strictly decreasing). This yields a contradiction since the equation is supposed to be true for every site $x \leq W$, and $W \geq 2k$. This concludes the proof of Theorem 3.6. \square

3.5 Conclusion

3.5.1 Discussion and Future Directions

This chapter focuses on the mechanism design challenges in the basic dispersal game inspired by animal behavior. Our focus is on congestion policies, where there is no control over the utility associated with a site, and only the price associated with congestion can vary. Indeed, for natural scenarios in particular, it makes sense that evolution would impact the competition levels but the values of sites would be determined by the environment. The definition of congestion policies neglects, however, a factor that might play a significant role in several scenarios, namely, the cost incurred when in visiting a site x (e.g., the energetic cost consumed while traveling to x). Studying the more general model that takes into account this extra cost is left for future work.

The coverage measure can find relevance in contexts of animal foraging. This is because effective group coverage can indirectly contribute to the fitness of individuals, especially when the group is in competition with other groups over the same set of resources [38]. Indeed, in this case, the consumption of many resources by conspecifics reduces the fitness of individuals in the competing group and vice versa. It is therefore of interest to develop a more comprehensive game-theoretic perspective that would integrate into the individual fitness both the competition inside the group, and the competition between groups [25]. A plausible insight that can be learned from our results is that aggressive behavior between

conspecifics, which appears to be wasteful or even harmful from the group’s perspective, can in fact be effective on the group level, as it allows for better coverage. For further discussion on the interpretation of our results in the context of animals see Section 3.5.2.

Our definition of coverage assumes that a single player in a site x suffices to consume the full utility $f(x)$. In natural settings, e.g., in animal foraging scenarios, this assumption might be too strong, and it is of interest to study relaxations of it. One possibility is to assume that there is a maximum capacity of consumption per individual. In some other cases, a more complex definition of coverage might be appropriate. We leave the study of such generalizations for future work.

Although we formally discuss only the one-shot scenario, we wish to stress that the ESS, by definition, is stable when played repeatedly in an infinite population. We did not, however, investigate other forms of repetition, which are left for future work.

Finally, game theory in computer science typically studies the notion of Nash Equilibrium as the main concept to capture stability. This notion was originally designed for games with small number of players. When considering games with thousands or even millions of players, it becomes less plausible to assume that all players are rational, and one may seek for stronger notions of stability. ESS, which was developed in the evolutionary game theoretical community, appears to be a good candidate to capture stability in large populations. We believe that this notion should be applied more in game theory works that correspond to large distributed settings

3.5.2 Informal Discussion on Animal Dispersal

From the perspective of the group, consuming a large amount of food by all members together can indirectly increase the fitness of the group members and hence become significant for their survival. For example, when multiple species compete over the same patched resource, a thorough consumption of patches by one species implies less food remaining for the other. In this respect, it would be interesting to experiment on the interplay between two similar species that compete over the same set of resources, but differ in their level of aggressiveness toward conspecifics [79]. In cases where there is no direct contact between the species (e.g., when they feed at different times of the day), it may appear that the more aggressive species would be inferior as it induces unnecessary waste of energy and risks of injury. However, our results suggest that perhaps it is the aggressive species that would turn out to be the superior one. Indeed, its aggressive behavior incurs higher collisions costs, which may drive individuals to better cover the resources, on the expense of the more peaceful species.

During foraging, both selfish individuals and collaborating groups (e.g., ants or bees) share the basic challenge of balancing the need to exploit the more promising resources, while trying to avoid overlaps (or collisions) with conspecifics [29, 38]. This tradeoff, however, is manifested in the two cases in a manner that appears to be very different: While collaborative groups aim to optimize the tradeoff by setting the parameters in a somewhat “centralized” manner, the competing dynamics is governed by a selection process yielding an evolutionary stable strategy which need not be optimal for individuals. As suggested here, despite these differences, when colliding individuals are punished

severely, the two perspectives yield similar strategies.

Finally, the alignment in behavior between collaborating and competing individuals, as reported here, may suggest a modest contribution to the theory of evolution of eusocial species. The currently dominated theory proposes that eusocial species have evolved from self motivated individuals due to kin selection [43]. selection []. An issue that is discussed less is that a shift in the “motivation focus” can potentially be accompanied by a shift in function and hence, perhaps, also in structure [61]. Indeed, in principle, while evolving from self-interested to altruistic, an individual which had evolved in a way that was tailored to its own needs, may now need to adapt to the collective needs. As these needs are on different scales, they may not necessarily be akin [30], and hence a long evolutionary process may be required in order to bridge this gap. In this respect, that seemingly cooperative behavior can arise in competing systems, as demonstrated here, suggests that the behavioral transition, at least in the context of dispersal or forage, may have been smoother than what one might expect.

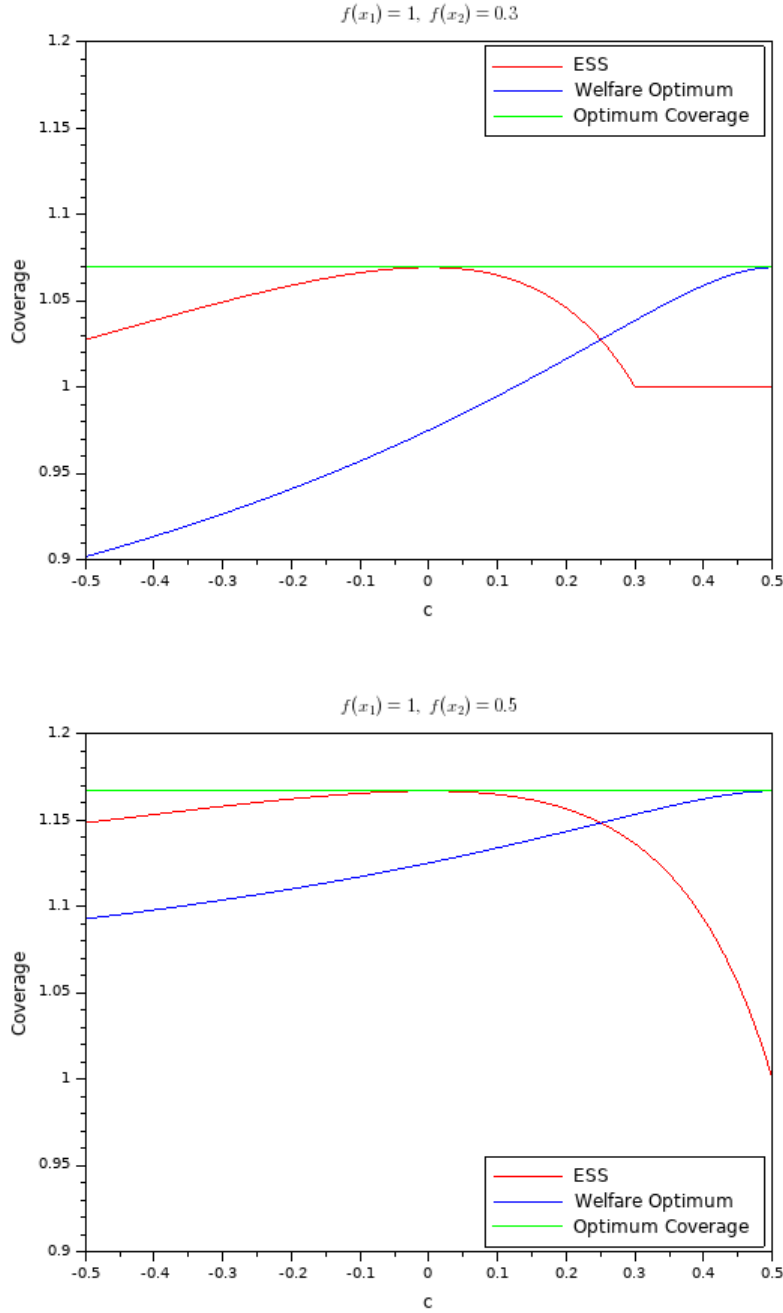


Figure 3.1 – Coverage as a function of the extent of competition (parametrized by c) for the case where two players compete over two sites x_1 and x_2 . In the top figure (respectively, the bottom figure) we consider the case where $f(x_1) = 1$ and $f(x_2) = 0.3$ (respectively, $f(x_2) = 0.5$). The X-axis represents a range of congestion functions C_c , defined as $C_c(1) = 1$ and $C_c(2) = c$, where $c \in [-0.5, 0.5]$. The case $c = 0$ corresponds to the exclusive function and the case $c = 0.5$ corresponds to the sharing function. The Y-axis represents the coverage. The red line corresponds to the coverage of the ESS, the green corresponds to the optimum coverage, taken over all symmetric strategies, and the blue corresponds to the coverage of the symmetric strategy that maximizes the individual payoffs.

Chapter 4

Foraging of Leptos Bats : Models and Simulations

4.1 Introduction

The goal of this chapter is to study the foraging behavior of nectar-feeding bats of the species *Leptonycteris yerbabuenae* (*Leptos* for close friends). During the spring, a few hundred thousands of these small bats, weighing around 20 grams for a size of 8 cm, dwell in a cave in the Sonoran desert in Mexico. Each night, they leave the cave after sunset and spread around a 70 km radius in order to find cacti flowers from which they feed. Each bat will need to visit several hundreds flowers in order to find the 15 mL of nectar that they need to survive (almost as much as their own weight). They are therefore facing an interesting mathematical problem : how to efficiently explore the area in order to maximize the amount of nectar that they can drink during the night?

This is a perfect example of a situation where many agents are faced with a strategic and competitive task, of which several aspects can be accurately modelled using game theory. However, because of the inherent complexity of the real-world setting, we will be faced with a trade-off between the realism of the model we design and the complexity of its analysis. With this in mind, we propose two approaches to analyse this game: one approach will be by the means of a realistic, but scaled down, computer simulation of the behavior of a group of bats in a field of cacti. This is the subject of section 4.3. The second approach is by the means of a more simple, yet insightful, mathematical model of the behavior of a single bat in a field of cacti. This is the subject of section 4.4. Section 4.2 will be dedicated to the description of the field experiments that were conducted on the site to better understand these bats' behavior.

4.1.1 Field Experiments

In may 2016, I went to the Sonoran desert in Mexico to assist a team of zoologists from Tel-Aviv university [95] that study *Leptos* bats. The team was on site for about two months, and I stayed only for about one month. The team also conducted experiments on the same site the year before and the year after my participation. Section 4.2 is a brief description of the different experiment that were conducted there, and the main conclusions that they helped formulate. Specifically, the team recorded GPS data about the bats' trajectories and about the cacti positions, video and audio recordings of the way

bats visit cacti flowers, data about the number of flowers and the quantity and quality of nectar produced by cacti.

These quantitative data greatly improved our global knowledge about the *Leptos* species, and allowed us to design the models presented in sections 4.3 and 4.4, as well as formulate some hypotheses about the bats mental states that lead to this behavior, such as the amount of data that each bat memorizes about a particular field of cacti.

4.1.2 Computer Simulations

In section 4.3, we present a simulation framework which aims to model this specific situation as accurately as possible, but also allows to explore potential situations that are not currently encountered in the field. Several compromises also had to be made in the design of the simulations model. For computational reasons, we limited the simulations to a small group of bats (up to about a hundred bats) and a reasonably-sized cactus field (up to 10000 cacti), but we believe these restriction to impact the results in a very limited way. We also had to decide of a behavioral model for the bats. We picked the simplest model that could provide results consistent with real-field data. Our behavioral model is strongly inspired by [88]. The simulations run in “real-time” and have an accurate representation of space.

The main results that came out of these simulations are the following. On many aspects, the simulations results were very close to the recorded field data, which suggests that our simple behavioral model bears at least some ressemblance with the actual bat behavior. In particular, it suggests that the bats might have a very good memory of the positions and quality of the cacti they visit. In addition, the simulations offers an explanation to why the bats often come back to the same cacti during the night. This type of territoriality might be explained simply by a reinforcement mechanism similar to our behavioral model. Another strong point in favor of these hypotheses is that our model is evolutionarily stable, at least empirically. Indeed, the behavioral parameters used in the simulations are adjusted by an evolutionary algorithm that simulates natural selection and mutations over many generations. Another interesting similarity between the simulations and the field data is that the bats collective behavior is very effective at extracting as much nectar as possible from the cactus field. This situation is reminiscent of chapter 3 : it seems that strong competition at an individual level allows for a great efficiency at the group level. Lastly, the simulations allowed us to explore speculative situations where the number of bats per cactus would be much lower or higher than the natural setting. It demonstrated that the increased or decreased competition in these settings would have a strong impact on the bats’ equilibrium behavior, and on their individual fitness. This suggests that the number of bats in the field is in balance with the number of cacti so that every bat has just enough food to survive and bear children.

4.1.3 Theoretical model

In section 4.4, we design and study a mathematical model inspired by the bats’ foraging behavior. We start with a very simple model and incrementally make it more realistic. The first model features a single bat with a perfect knowledge of the cactus field. This situation is a rather simple yet interesting optimisation problem, and we describe a polynomial

algorithm which finds an optimal strategy for the bat.

The next step is to deprive the bat of its perfect knowledge, and instead give it a perfect memory, which transform the problem into an online optimisation problem. This problem is much harder, and we cannot quite give a result as precise as before. Instead, we prove that an optimal algorithm for this problem must proceed in two phases, an *exploration* phase followed by an *exploitation* phase. This is a standard dichotomy in the literature about similar games, and to see this concept naturally appear in such a mathematical model is an additional argument to study this dichotomy. The key ingredient for an optimal algorithm is therefore to decide when to switch from the first phase to the second. In general, this problem can be solved by a methodical analysis of the whole tree of possible futures depending on our present decisions, but such an analysis leads to an algorithm of exponential time complexity.

Therefore we later turn our attention on a particular case, in which the cacti parameters are chosen uniformly at random. In this case, we are able to give an upper bound on the efficiency of algorithms, and we also give a polynomial-time algorithm that achieves a near-optimal average payoff. This is yet another argument showing that a simple behavior model can display a lot of similarities with the actual bats' behavior, in particular concerning its overall efficiency.

4.2 Field Experiments

4.2.1 Description of the experiments

We describe below the different experiments that were conducted in the Sonoran desert, and the type of data collected.

GPS data Using special nets, made with very light strings designed to catch bats or birds without harming them, we captured many bats. Among the captured individuals, we selected bats that were fit to carry the GPS sensor: we made sure that they weighed 20 grams or more, were not pregnant nor lactating. We released the bats that did not meet these criteria. The GPS “backpack” contains a GPS sensor that record its position (latitude, longitude, and altitude) every few seconds in a memory card. It also contains a “beeper” that emits shortwave radio signals every minute or so. This signal can be picked up from a few hundreds meters to a kilometer and allows for a easier recovery of the device. The whole “backpack” weighs 4 grams and was glued on the bats' backs using special glue, as shown in Figure 4.1. It stays attached to the bat for a few days until the glue weakens and the backpack falls on the ground. It then needs to be found in order to recover the data. We used two strategies in order to obtain a good quantity and quality of data. The first strategy is to capture bats at their feeding site, which makes recovery of the device more likely (we can find it back if it falls on the feeding site or in the cave, but not on the way between). However it makes capture more difficult because there is a lower density of bats flying at the feeding site. The capture rate using this method was a few bats per night, sometimes none, using up to three nets. We also captured bats close to their cave, which allows us to capture several dozens in a single night with a

single net, but the backpacks glued onto these bats must fall in the cave if we are to recover them. Combining these methods, we were able to recover about 25 devices out of the 50 that were used. To the best of my knowledge, this study was the first to collect GPS data from *Leptotis* bats.

Video data A few night-vision cameras were installed¹ on top of poles close to cacti with open flowers, in the hope of recording visits from bats to these flowers. Many visits were recorded.

Audio data Microphones were placed² directly on cacti in order to record sounds and ultrasounds emitted by bats in a radius of a few meters. The microphones were sometimes coupled with the cameras to get a synchronised audio and video recording of visit events – see Figure 4.2.

Tent experiments A few captured bats were released inside a tent along with a sawed off cactus head³ with flowers. This allowed us to take close-up photographs of visit events – see Figure 4.3. We also recorded the sounds made by bats in presence of other bats, which is useful for comparison with recordings from the field data.

Cactus data We closely monitored two dozen cacti for several weeks. We recorded the number of flowers open on each night, the number of buds remaining, the quantity of nectar in some flowers at different hours during the night, and the sugar concentration of the nectar. Half of these cacti were covered with nets to prevent visits from bats – see Figure 4.4.

Map We flew a drone during the day to map the feeding site and collected the GPS coordinates of 680 cacti.

4.2.2 Qualitative results

By confronting the GPS data on the bats trajectories with the GPS positions of the cacti that we observed in the field, we obtained data about the precise sequence of visits that a bat makes each night, over several nights. We were able to determine that a bat visits about 150 to 600 cacti per night, and that some cacti are much more visited than others. By looking at the precise distribution of the number of visits per cactus, we could conclude that the bats do not visit cacti uniformly at random, nor do they pick cacti based only on the distance to the cactus. We also observed that the bats generally come back to the same area from one night to another, even after being captured⁴ (we could have expected

¹ This sensitive equipment must be placed after sunset and removed before sunrise because they cannot survive the high temperatures of the day.

² See footnote 1

³ Since the saguaro cactus is a protected species, we obtained an authorization from the state of Sonora to cut some cacti heads for the need of the experiment, as the branches that we sawed off will likely not survive.

⁴ We even caught a bat twice. We could figure out that the bat was already captured and equipped with a GPS device because of a rectangular area of glue-induced hair loss on its back.



Figure 4.1 – A *Leptos* bat with a GPS device glued to its back, suspended from the ceiling of a tent. Photograph by Stefan Greif.

that they would choose to avoid the area of their capture by fear of being captured again). This suggests that either the bat acquires knowledge about the area that is useful for subsequent nights, or some kind of territoriality is taking place at the feeding sites.

From the video and audio data, we could observe that the duration of a bat's visit to a cactus is usually very short: less than a second inside a flower, and rarely more than ten seconds in the vicinity of a particular cactus. It also seems that the bats tend to avoid each other in the field, sometimes using repulsive calls. These calls were recorded and compared to similar calls recorded in the tent.

The cacti experiments revealed that the number of flowers on each cactus is very variable, from zero to a dozen in a given night. There is also a strong correlation between the number of open flowers in a given night and the number of open flowers the next night on a same cactus. This also suggests that knowledge about these cacti would be useful to a bat seeking to maximize the amount of food she gets in a night. The rate of nectar production and the sugar concentration seem to fluctuate by a factor of two between flowers, but do not seem correlated to individual cacti. And finally, the fact that a cactus is visited or not, by a bat or by a human, does not seem to have an impact on the number of flowers, the number of buds, or the nectar it produces.

4.3 Simulation Results

4.3.1 Experimental Framework

Based on the qualitative results from the previous paragraph, we designed simulations to gain more insight about the bats behavior. To simulate the feeding site, we assign a value to each cactus, which represents the number of open flowers on this cactus. For simplification, we assume that each flower produces the same quality and quantity of



Figure 4.2 – A Saguaro cactus to which is attached a microphone, with a camera on a pole next to it. The yellow rope is about 2 meters high. Photograph by myself.



Figure 4.3 – A Leptotis bat milliseconds before visiting a Saguaro flower in a tent. Photograph by Stefan Greif.

nectar. The probability law for the number of flowers can be chosen as we please. For the cacti positions, we either draw them uniformly at random on a square of chosen dimension, or we use the actual positions measured in the field.

To model the bats behavior, we designed the following model. Each bat remembers the positions of each individual cactus, and the total amount of nectar that she drank from it. For a given cactus i , we denote this quantity by \mathbf{nectar}_i . After each visit, each bat chooses which cactus to visit next, based only on \mathbf{nectar}_i , and its distance d_i to each cactus i . Precisely, the probability p_i of visiting cactus i is equal to

$$p_i = \frac{f_i}{\sum_{j=1}^n f_j}, \quad \text{with} \quad f_i = \frac{1 + \alpha \mathbf{nectar}_i}{d_i^\beta},$$

where α and β are the parameters of this behavior. We call α the *learning parameter* and β the *spatial parameter*. For both parameters, there is a trade-off between *exploration* and *exploitation*. A higher value of α implies that the bat is more likely to return quickly to a cactus which contained a good quantity of nectar, providing a good exploitation of the cacti discovered. Conversely, a lower value of α allows for more exploration of the field, and diversification of cacti visited. For β , a higher value implies that the bat is less likely to visit cacti that are more distant, therefore decreasing the time between visits and increasing exploitation potential. Instead, a lower value of β increases the potential to visit cacti that are more distant, hence better exploration.

4.3.2 Simulations

The basic unit block of a simulation is a *night*: a defined duration in which a number n of bats will go and visit a field of m cacti. Typically we choose the ratio $\frac{m}{n}$ to be around 30,



Figure 4.4 – A Saguaro cactus covered by a fine net to prevent visits from bats. Photograph by myself.

because it matches the actual ratio from field data. The simulation of a night involves many other parameters, described thereafter.

The night duration is usually chosen to be around 3 to 6 hours, to reflect the typical duration recorded in the field data. For simplicity, we assume that all bats stay in the field for exactly this duration, while in reality, some bats may stay for only 3 hours and others for 6 hours during the same night.

The bats speed is chosen to be 5 meters per second. We assume that the bats only move in straight lines at a constant speed, while in reality, their movement is more erratic and can vary in speed. However an average speed of 5 meters per second seems consistent with the field data.

Starting distance: the bats start the night at the cave, which can be far from the feeding site. We have evidence of feeding sites ranging from a few hundred meters from the cave up to 70 kilometers away. This distance can be adjusted, and we typically fix it at 20 kilometers.

Drink limit: for realism, we introduce a limit to how much a bat can drink during a single visit. This limit is affected by the duration of the visit (less than a second), and the speed at which bats digest the nectar (they can drink up to 20 mL in a night). We fix this limit at 0.2 mL per visit.

Flower distribution: based on the field data, the distribution of flowers among cacti is best approximated by a geometric distribution with parameter 0.5. Precisely, it means that half of the cacti have no flower, a quarter have 1 flower, one eighth have 2 flowers, etc. We allow up to 10 flowers on a single cactus.

We can run simulations that span only a night, or run evolutionary simulations. The evolutionary simulations model random mutations and natural selection in order to determine what set of parameters produces the most effective bat behavior. In the first generation of bats, the learning and spatial parameters are distributed according to a normal law of chosen mean and variance. At the end of each night, we remove the 10% of bats that obtained the less amount of nectar, and replace them with “offsprings” of the 10% most successful bats, thus obtaining the next generation. The offsprings inherit the parameters of their parent, slightly altered by a random perturbation that models mutations. We can then obtain data about the values of each parameter and the performances of every bat from every generation.

4.3.3 Outcome of the Simulations

In this section, we present a small selection of the most significative results from the simulations. The field setting used in these results is a virtual recreation of a portion of the real cactus field in the Sonoran desert, which contains 680 cacti. A map of this field is drawn in figure 4.5.

The main purpose of the evolutionary simulations is to find a set of parameters that approximate a population equilibrium, and then be able to compare this equilibrium with

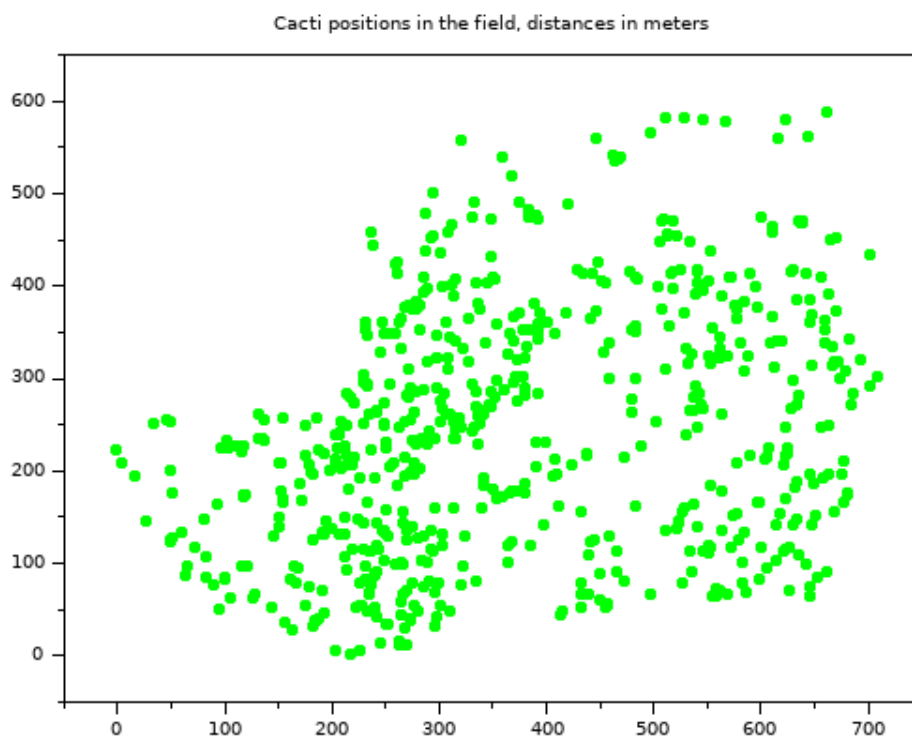


Figure 4.5 – The cacti positions measured in the field and used in the simulations.

the behaviors encountered in nature. Studying convergence towards the equilibrium is also of interest. Since we have no idea a priori about what the equilibrium parameters are, it is important to select a starting point which is as neutral as possible. Since the learning parameter α can be either positive or negative, we chose a starting configuration in which the learning parameters of the bats were selected at random according to a normal distribution centered on zero. This would allow the population to converge either on a positive or a negative value for this parameter. The exploration parameter β , however, can only be positive, therefore we chose a normal distribution centered on 2, with a rather large variance.

After running many tests, and adjusting the relevant parameters, we found that the population always seems to converge towards a single point (α, β) , with α always positive and β approximately between 2 and 4. There were a few instances in which the population was temporarily divided in two distinct sub-populations with little variation within a sub-population, but these situations always eventually collapse into a single coherent population. There were no instances in which the learning parameter would converge towards a negative value. These facts seem to indicate that there is almost always a single point of equilibrium which correspond to a dominant strategy. As expected, the speed of convergence is most affected by the proximity of the starting point to the equilibrium, and the rate of mutation allowed between generations. A smaller mutation rate will lead to a population that has smaller variations within itself, but a slower convergence. If the mutation rate is even too slow, it might not be enough to have a big enough impact to dominate random noise, and the population will not converge at all, instead behaving like a random walk. On the other side, a bigger mutation rate will lead to a less coherent population, but a faster convergence.

Figures 4.6 and 4.7 represent the evolution of a population of 23 bats over 40 generations from a single simulation. For this simulation, we choose the parameters that allow a fast and clear convergence, in a setting that most closely matches the natural setting. The dots represent the average value in the population, and the error bars length are equal to the standard deviation within the population.

The value of β being within the range $[2, 4]$ can also be put in perspective with the concepts of Lévy flights and optimal random searches, (see, e.g [93]), although our setting is a bit different from the classical setting of Lévy flights.

Another interesting factor to study is the efficiency of the bats in exploiting the field of cacti. In figure 4.8, we represented the number of visits received by cacti depending on their number of flowers. This figure clearly indicates that the collective behavior of the bats correctly identifies the most productive cacti, in fact the correlation coefficient between the number of flowers and the number of visits is about 0.866 (in this simulation). Moreover, in figure 4.9, we represented the percentage of nectar drunk over the total nectar produced in the field for 6 populations of bats of various sizes. It shows in particular that 23 bats are enough to collect about 96% of the nectar, which is quite impressive given that it corresponds to a ratio of cacti to bats of approximately 30. Furthermore, in a group of only 12 bats, which corresponds to about 55 cacti per bat, an efficiency of around 85% is still reached. This seems to demonstrate that even the simple behavior implemented in

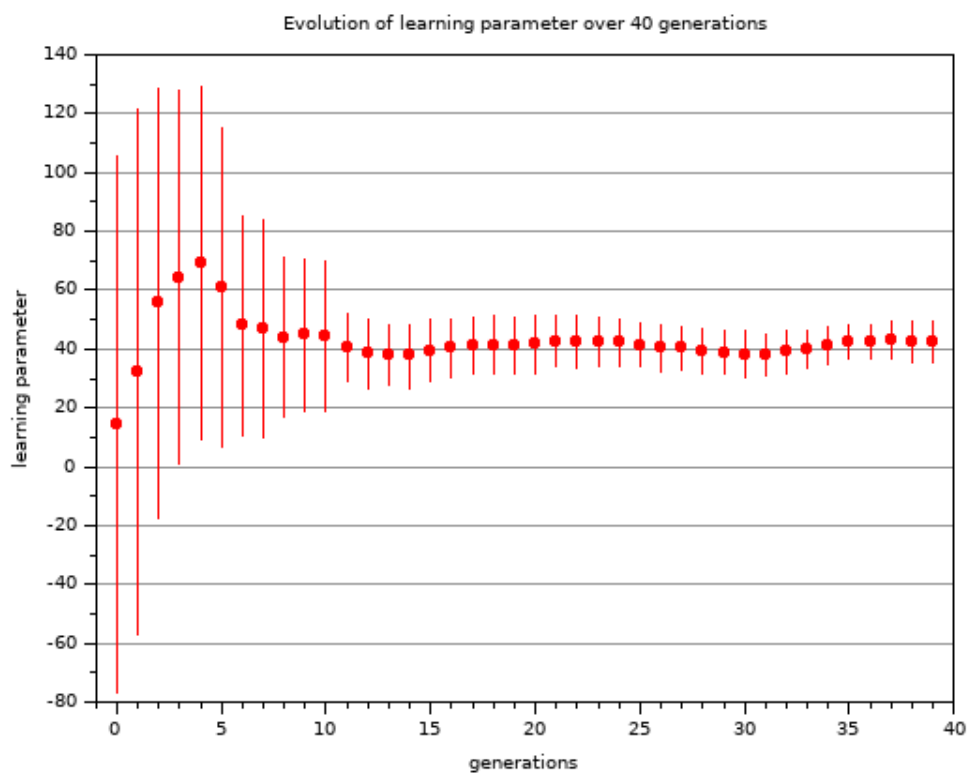


Figure 4.6 – Evolution of the learning parameter in a population of 23 bats over 40 generations.

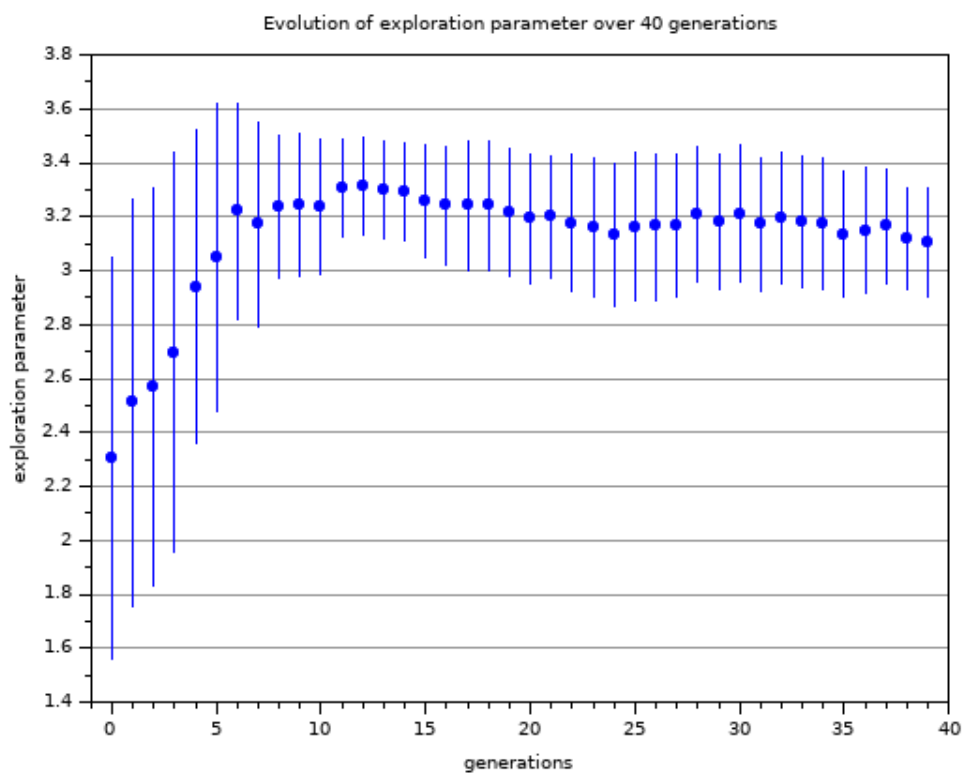


Figure 4.7 – Evolution of the exploration parameter in a population of 23 bats over 40 generations.

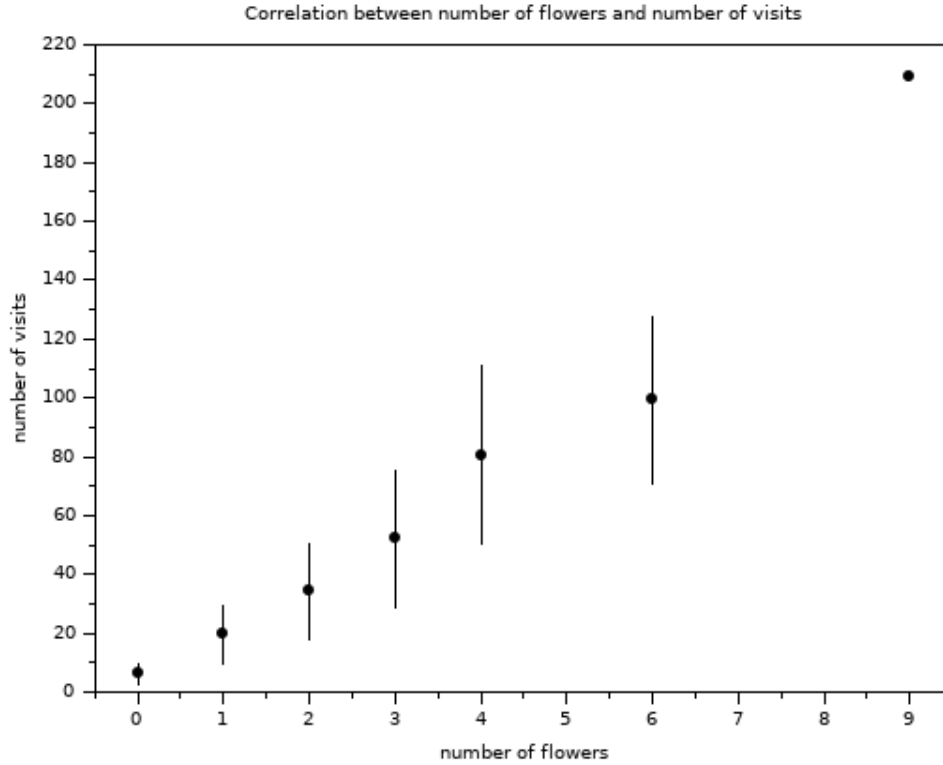


Figure 4.8 – Number of visits received by each cactus, depending on their number of flowers.

the simulations is able to achieve very high group efficiencies, provided only that the bats have a good memory of the field and are able to evaluate distances accurately.

Lastly, we want to study the impact of increased competition on the equilibrium. For this purpose, we represented in figure 4.10 the equilibrium value of the learning parameter in different groups of bats of various sizes. The figure shows that when competition increases, i.e when there are more bats per cactus, the equilibrium value of α also increases. We interpret this result as showing that when competition increases, it becomes more beneficial for the bats to favor exploitation over exploration. Hence, when competition increases, the average number of cacti visited by a single bat tends to decrease, as the number of cacti visited more than once. With a high enough competition, some kind of territoriality even emerges within the population, as demonstrated by the fact that on most cacti, over 80% of the visits are made by a single bat.

The impact of competition on the exploration parameter is however not clear to us, as we failed to demonstrate a significant correlation between the two. It seems that the optimal value of β is always in the same range, between 2 and 4, and seems more impacted by the characteristics of the field itself rather than the number of bats.

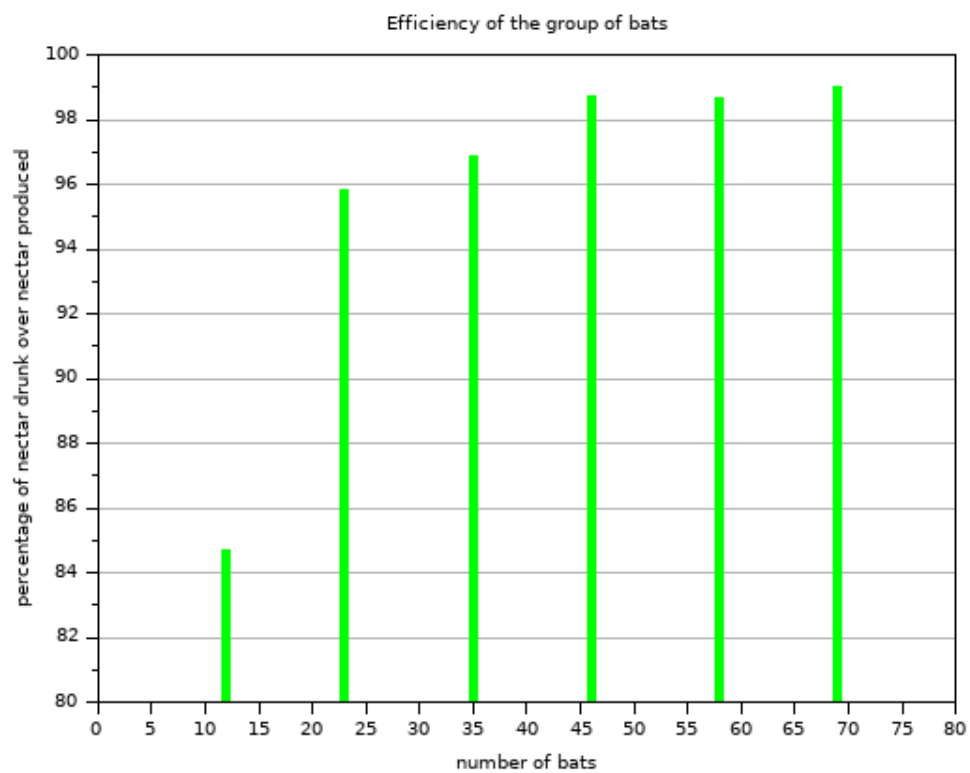


Figure 4.9 – Proportion of nectar drunk over nectar available in the field, for 6 groups of bats of different sizes, demonstrating the efficiency of the group.

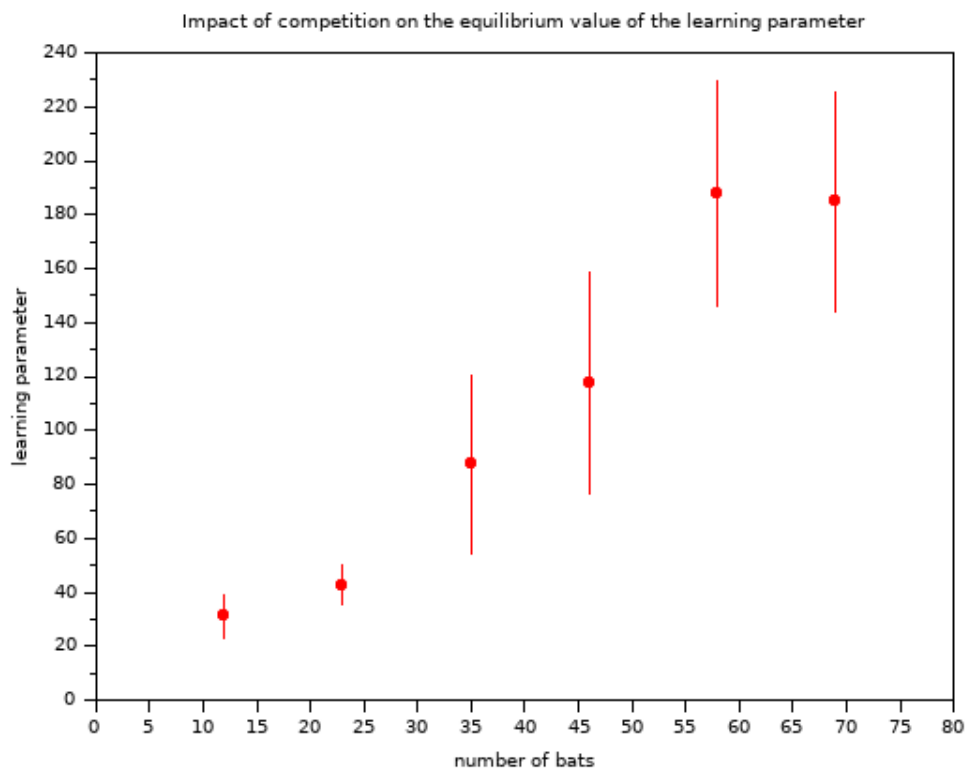


Figure 4.10 – The equilibrium value of the learning parameter in populations of various sizes, demonstrating the impact of competition on the bats behavior.

4.4 Analytical Results

In this section, we design and study a mathematical model inspired from the Leptos bats. This model is the result of a tradeoff between its relevance to the real bats, measured by its overall realism, and its mathematical interest, which requires the model to be neither too simple nor too complex. Our approach was therefore to start from the simplest model and incrementally make it more realistic.

4.4.1 Setting and notations

There are a few assumptions we can make in order to make the problem analytically tractable. First, it seems that the bats have no difficulty in finding the cacti from which they feed. Indeed, most of these cacti are huge Saguaro cacti measuring up to 12 meters, scattered around the desert in packs of hundreds of cacti. There is evidence that the bats are able to recognize a pack of cacti from altitudes of a few hundred meters. Within a pack, the cacti are a few meters apart from each other, and the bats are able to navigate easily around them, without ever colliding with a cactus, thanks to their echolocating abilities. For these reasons, we will simply assign an integer from 1 to m to each cactus and assume that the bats fly directly from one cactus to another, hence ignoring the spatial distribution of cacti.

The difficulty of the problem is not to find the cacti, but to find *good* ones. Indeed, not every cactus is identical: some of them have many flowers (up to a dozen or even more), while others have a single flower or none. Even different flowers are not identical, some of them produce more nectar than others, or produce nectar with a higher sugar concentration. We want to take all of these differences into account to assign a positive value to each cactus, which represents the amount of nectar (or energy) that it produces during a given time period. For a cactus i , we note r_i this value, and call it the *filling rate* of the cactus, as it represents the speed at which the cactus' flowers fill with nectar as time passes. We also assign a non-negative value b_i to each cactus, which represents the amount of nectar initially available from it.

Notations In the remainder of the chapter, we will denote the integer interval $\{1, \dots, n\}$ by $[n]$.

We will analyse a single bat behavior, and we model it as follows. The bat has to visit exactly T cacti during the night. The bat's strategy simply amounts to the order in which she wants to visit the cacti, and is therefore represented as a vector $\sigma \in [m]^T$ of T elements from $[m]$. We call σ a *sequence of visits*: σ_1 is the index of the first cactus visited, σ_2 is the index of the second cactus visited, etc. We note $S = [m]^T$ the set of such sequences. For a given sequence σ of visits, let

$$\hat{\sigma} = \{i \in [m] \mid \exists t \in [T], \sigma_t = i\}$$

be the set of cacti that are visited at least once. For a cactus $i \in [m]$, we note $t_i(\sigma)$ the last time that this cactus is visited in σ , defined as:

$$t_i(\sigma) = \begin{cases} \max\{t \in [T] \mid \sigma_t = i\} & \text{if } i \in \hat{\sigma} \\ 0 & \text{otherwise} \end{cases}$$

The bat's objective is to maximize the amount of nectar that she can drink in T visits, we call this quantity the *utility* of the sequence, and we denote it by $U(\sigma)$. Hence, we have

$$U(\sigma) = \sum_{i \in \hat{\sigma}} (b_i + r_i t_i(\sigma)) \quad (4.1)$$

In some proofs, it is more convenient to use the following equivalent expression for $U(\sigma)$:

$$U(\sigma) = \sum_{i \in \hat{\sigma}} b_i + \sum_{t=1}^T r_{\sigma_t} (t - t_{\sigma_t}(\sigma_1, \dots, \sigma_{t-1})) \quad (4.2)$$

To summarize, the game's parameters are: the number of visits T , the number of cacti m , the vector of cacti filling rates $R = (r_i)_{1 \leq i \leq m} \in (\mathbb{R}_+^*)^m$ and the vector of initial quantity of nectar $B = (b_i)_{1 \leq i \leq m} \in (\mathbb{R}_+)^m$. We note such a game $\Gamma(T, m, R, B)$. We note $\mathbf{Opt}(T, m, R, B)$ the set of optimal sequences for the game $\Gamma(T, m, R, B)$, i.e., $\mathbf{Opt}(T, m, R, B) = \{\sigma \in S \mid U(\sigma) = \max_{\sigma' \in S} U(\sigma')\}$.

4.4.2 Model with complete information

We are first interested in the simplest model, in which the bat knows the values of vectors R and B in advance. In this setting, the bat can plan all her visits from the beginning, and predict the utility associated with any sequence of visits. Therefore, the problem simply amounts to finding $\arg \max_{\sigma \in S} U(\sigma)$. Without loss of generality, we will assume that the cacti are ordered in decreasing order of filling rates, i.e., $i < j \implies r_i \geq r_j$. This will make the statements of some results clearer.

Let us start by stating a few intuitive lemmas about optimality conditions. The proof technique is always the same: we start with a sequence satisfying a given condition, and exhibit another sequence with a larger payoff, therefore proving that the first sequence is not optimal. The first two lemmas address the case where $m \leq T$.

Lemma 4.1. *If $m \leq T$, then an optimal sequence must visit every cactus exactly once during the last m visits.*

Proof. Let σ be a sequence and assume that there exists $i \in [m]$ such that $\forall t \in \{(T - m + 1), \dots, T\}, \sigma_t \neq i$. By the pigeonhole principle, there must be a cactus visited at least twice during the last m visits, i.e., $\exists t_1 < t_2 \in \{(T - m + 1), \dots, T\}, \sigma_{t_1} = \sigma_{t_2}$. Let us consider the sequence σ' defined by:

$$\sigma'_t = \begin{cases} i & \text{when } t = t_1 \\ \sigma_t & \text{otherwise} \end{cases}$$

From this definition, it follows that

- $\hat{\sigma}' = \hat{\sigma} \cup \{i\}$
- $t_i(\sigma') = t_1 > t_i(\sigma)$
- $\forall k \in \hat{\sigma} \setminus \{i\}, t_k(\sigma') = t_k(\sigma)$

Therefore,

$$U(\sigma') - U(\sigma) = \begin{cases} r_i(t_1 - t_i(\sigma)) & \text{when } i \in \hat{\sigma} \\ r_i t_1 + b_i & \text{otherwise} \end{cases}$$

In both cases, $U(\sigma') - U(\sigma) > 0$, which proves that σ is not optimal. \square

Lemma 4.2. *If $m \leq T$, in every optimal sequence, the last m cacti visited must be arranged in increasing order of filling rates.*

Proof. Let σ be a sequence for the game $\Gamma(T, m, R, B)$ and assume there exist $t_1, t_2 \in \{(T - m + 1), \dots, T\}$ such that $t_1 < t_2$ and $r_{\sigma_{t_1}} > r_{\sigma_{t_2}}$. Let $i = \sigma_{t_1}$ and $j = \sigma_{t_2}$. Let us consider the sequence σ' defined by:

$$\sigma'_t = \begin{cases} j & \text{when } t = t_1 \\ i & \text{when } t = t_2 \\ \sigma_t & \text{otherwise} \end{cases}$$

From this definition, it follows that:

- $\hat{\sigma}' = \hat{\sigma}$,
- $\forall k \in \hat{\sigma} \setminus \{i, j\}, t_k(\sigma') = t_k(\sigma)$.

Furthermore, if there exists $t'_1 > t_2$ such that $\sigma_{t'_1} = i$, by lemma 4.1, σ is not optimal. Otherwise we have $t_i(\sigma') = t_2$. By the same reasoning, if there exists $t'_2 > t_1$ such that $\sigma_{t'_2} = j$, by lemma 4.1, σ is not optimal. Otherwise we have $t_j(\sigma') = t_1$.

If $t_i(\sigma') = t_2$ and $t_j(\sigma') = t_1$, then $U(\sigma') - U(\sigma) = r_i(t_2 - t_1) + r_j(t_1 - t_2) = (r_i - r_j)(t_2 - t_1) > 0$, which proves that σ is not optimal. \square

Thanks to these two lemmas, we are able to completely determine the optimal sequences when $m \leq T$. Indeed, every sequence σ satisfying the conditions of lemmas 4.1 and 4.2 achieves the optimal payoff of $U(\sigma) = \sum_{i=1}^m (b_i + r_i(T - i + 1))$. We now turn our attention to the other case, which is $T < m$.

Lemma 4.3. *If $T \leq m$, then an optimal sequence must visit every cactus at most once.*

Proof. Let σ be a sequence and assume there exist $t_1, t_2 \in [T]$ such that $t_1 < t_2$ and $\sigma_{t_1} = \sigma_{t_2}$. It follows that $\text{card}(\hat{\sigma}) < T \leq m$, and therefore there must be an unvisited cactus $i \in [m] \setminus \hat{\sigma}$.

Let us consider the sequence σ' defined by:

$$\sigma'_t = \begin{cases} i & \text{when } t = t_1 \\ \sigma_t & \text{otherwise} \end{cases}$$

From this definition, it follows that

- $\hat{\sigma}' = \hat{\sigma} \cup \{i\}$,

- $t_i(\sigma') = t_1$,
- $\forall k \in \hat{\sigma}, t_k(\sigma') = t_k(\sigma)$.

Therefore, $U(\sigma') - U(\sigma) = r_i t_1 + b_i > 0$, which proves that σ is not optimal. \square

Lemma 4.4. *If $T \leq m$, in every optimal sequence, all cacti must be arranged in increasing order of filling rates.*

Proof. Let σ be a sequence for the game $\Gamma(T, m, R, B)$ and assume there exist $t_1, t_2 \in [T]$ such that $t_1 < t_2$ and $r_{\sigma_{t_1}} > r_{\sigma_{t_2}}$. Let $i = \sigma_{t_1}$ and $j = \sigma_{t_2}$.

Let us consider the sequence σ' defined by:

$$\sigma'_t = \begin{cases} j & \text{when } t = t_1 \\ i & \text{when } t = t_2 \\ \sigma_t & \text{otherwise} \end{cases}$$

From this definition, it follows that:

- $\hat{\sigma}' = \hat{\sigma}$,
- $\forall k \in \hat{\sigma} \setminus \{i, j\}, t_k(\sigma') = t_k(\sigma)$.

Furthermore, if $t_i(\sigma') \neq t_2$ or $t_j(\sigma') \neq t_1$, then by lemma 4.3, σ is not optimal. Otherwise, $U(\sigma') - U(\sigma) = r_i(t_2 - t_1) + r_j(t_1 - t_2) = (r_i - r_j)(t_2 - t_1) > 0$, which proves that σ is not optimal. \square

As a consequence of lemmas 4.3 and 4.4, when $T \leq m$, an optimal sequence is entirely determined by its support, which is necessarily of size T . We will use this fact later. First we turn our attention to the particular case where all cacti are initially empty and have different filling rates. We call this simplified model the *vacant model*. In this particular case, we can show that there is a single optimal sequence, as stated in theorem 4.5.

Theorem 4.5. *Under the assumptions*

1. $T \leq m$,
2. $\forall i \in [m], b_i = 0$,
3. $\forall i, j \in [m], i < j \implies r_i > r_j$,

the only optimal sequence σ^ is defined by: $\forall t \in [T], \sigma_t^* = T + 1 - t$.*

Proof. Let Γ be a game satisfying the theorem's assumptions, and σ be an optimal sequence for Γ . Let us prove that $\sigma = \sigma^*$. By lemma 4.3, $\hat{\sigma}$ and $\{1, \dots, T\}$ have the same cardinal. Let us assume that they are different. It follows that there exist two cacti $i \in \{1, \dots, T\} \setminus \hat{\sigma}$ and $j \in \hat{\sigma} \setminus \{1, \dots, T\}$.

Let us consider the sequence σ' defined by:

$$\sigma'_t = \begin{cases} i & \text{when } t = t_j(\sigma) \\ \sigma_t & \text{otherwise} \end{cases}$$

From this definition, and since j must appear only once in $\widehat{\sigma}$ by lemma 4.3, it follows that:

- $\widehat{\sigma}' = (\widehat{\sigma} \setminus \{j\}) \cup \{i\}$,
- $t_i(\sigma') = t_j(\sigma)$
- $\forall k \in \widehat{\sigma} \setminus \{j\}, t_k(\sigma') = t_k(\sigma)$.

Furthermore, we have $i < j$ by definition, which implies $r_i > r_j$ from the hypotheses, and therefore, $U(\sigma') - U(\sigma) = (r_i - r_j)t_j(\sigma) > 0$. Which proves that σ is not optimal. By contradiction, it implies that we must have $\widehat{\sigma} = \{1, \dots, T\}$. Then, by lemma 4.4, it follows that $\sigma = \sigma^*$. \square

We can now turn our attention back to the general model, in which the cacti are not necessarily initially empty, i.e $B \neq 0$. Note that the sequence σ^* defined in Theorem 4.5 is not necessarily optimal in this general setting. Indeed, consider the following example. There are only two cacti and a single visit is allowed, i.e $m = 2$ and $T = 1$. The first cactus has a filling rate of 1 and initial quantity of 0. The second cactus has a filling rate of 0 and an initial quantity of 2. If one applies the sequence σ from Theorem 4.5, one visits cactus 1 and gets a utility of 1, while visiting cactus 2 gives a utility of 2.

Also note that Lemmas 4.3 and 4.4 still hold in this setting. Therefore, the task of finding a optimal sequence is reduced to finding a set of T cacti. Indeed, any optimal sequence will visit exactly T cacti, according to lemma 4.3, and any optimal sequence whose support are these T cacti will visit them in the same order, according to lemma 4.4. It follows that it is possible to find an optimal sequence by an exhaustive search in the set of all subsets of $\{1, \dots, m\}$ with T elements, which we note $\mathcal{P}_T(\{1, \dots, m\})$. However this approach will lead to an inefficient algorithm, since the number of such subsets is $\binom{m}{T}$. The goal of the following paragraph is to provide an algorithm that finds an optimal sequence in $O(mT^2)$. Before describing the algorithm, we prove the following lemma.

Theorem 4.6. *Let $2 \leq T \leq m$ and $(R, B) \in (\mathbb{R}_+^*) \times (\mathbb{R}_+)$. Then for every $\sigma \in \mathbf{Opt}(T-1, m, R, B)$, there exist $\sigma' \in \mathbf{Opt}(T, m, R, B)$ such that $\widehat{\sigma} \subset \widehat{\sigma}'$.*

Proof. Since this proof requires the construction of many sequences, we designed figure 4.11 to help the reader visualize the relations between them.

Let $\sigma \in \mathbf{Opt}(T-1, m, R, B)$ and $\sigma^0 \in \mathbf{Opt}(T, m, R, B)$. We construct the sequence σ' iteratively, starting from σ_0 . If $\widehat{\sigma} \subset \widehat{\sigma}^0$, then we are done. Otherwise, let $i \in \widehat{\sigma} \setminus \widehat{\sigma}^0$. We show that we can form a sequence $\sigma^1 \in \mathbf{Opt}(T, m, R, B)$ with $\widehat{\sigma} \cap \widehat{\sigma}^1 = (\widehat{\sigma} \cap \widehat{\sigma}^0) \cup \{i\}$. We define recursively a sequence of rounds and cacti as follows:

$$\begin{array}{ccccccc}
& & t_2 & & t_1 & & t_3 \\
\sigma & : & \sigma_1 & \sigma_2 & \textcolor{red}{i}_1 & \sigma_4 & \textcolor{red}{i} & \sigma_6 & \textcolor{red}{i}_2 & \sigma_8 \\
\sigma^0 & : & \sigma_1^0 & \sigma_2^0 & \textcolor{red}{i}_2 & \sigma_4^0 & \textcolor{red}{i}_1 & \sigma_6^0 & \textcolor{red}{i}_3 & \sigma_8^0 & \sigma_9^0 \\
\sigma^1 & : & \sigma_1^0 & \sigma_2^0 & \textcolor{red}{i}_1 & \sigma_4^0 & \textcolor{red}{i} & \sigma_6^0 & \textcolor{red}{i}_2 & \sigma_8^0 & \sigma_9^0 \\
\sigma^2 & : & \sigma_1 & \sigma_2 & \textcolor{red}{i}_2 & \sigma_4 & \textcolor{red}{i}_1 & \sigma_6 & \textcolor{red}{i}_3 & \sigma_8
\end{array}$$

Figure 4.11 – Visualization of the different sequences defined in the proof of theorem 4.6. This is an example where $T = 9, l = 3, t_1 = 5, t_2 = 3$ and $t_3 = 7$.

- $t_1 = t_i(\sigma)$ and $i_1 = \sigma_{t_1}^0$;
- for all $k \geq 1$, if $i_k \notin \widehat{\sigma}$, we stop here;
- otherwise, we define $t_{k+1} = t_{i_k}(\sigma)$ and $i_{k+1} = \sigma_{t_{k+1}}^0$.

Because $\mathbf{card}(\widehat{\sigma}) = T - 1$, the process must stop at a step $l \leq T - 1$, and find a cactus i_l such that $i_l \in \widehat{\sigma^0} \setminus \widehat{\sigma}$.

We define a sequence σ^1 of size T as follows:

- $\sigma_{t_1}^1 = i$
- for all $k \in \{2, \dots, l\}, \sigma_{t_k}^1 = i_{k-1}$
- for all $t \notin \{t_k \mid k \in \{1, \dots, l\}\}, \sigma_t^1 = \sigma_t^{S_0}$.

From this definition, it follows that:

- $\widehat{\sigma^1} = (\widehat{\sigma^0} \cup \{i\}) \setminus \{i_l\}$
- $t_i(\sigma^1) = t_1$
- $\forall k \in \{1, \dots, l-1\}, t_{i_k}(\sigma^1) = t_{i_{k+1}}$
- $\forall j \notin \{i, i_1, \dots, i_l\}, t_j(\sigma^1) = t_j(\sigma^0)$

Furthermore, from the definitions of i_k and t_k , we know that for all $k \in \{1, \dots, l\}, t_{i_k}(\sigma^0) = t_k$. Therefore, we obtain the following equation.

$$U(\sigma^0) - U(\sigma^1) = b_{i_l} - b_i + \sum_{k=1}^l t_k r_{i_k} - t_1 r_i - \sum_{k=2}^l t_k r_{i_{k-1}}$$

And because σ^0 is optimal, we obtain the following inequality.

$$b_{i_l} + \sum_{k=1}^l t_k r_{i_k} \geq b_i + t_1 r_i + \sum_{k=2}^l t_k r_{i_{k-1}} \quad (4.3)$$

Let us define another sequence σ^2 of size $T - 1$ as follows.

$$\sigma_t^2 = \begin{cases} i_k & \text{when } t = t_k, \text{ for all } k \in \{1, \dots, l\} \\ \sigma_t & \text{otherwise} \end{cases}$$

It follows that $U(\sigma) - U(\sigma^2) = b_i + t_1 r_i + \sum_{k=2}^l t_k r_{i_{k-1}} - b_{i_l} - \sum_{k=1}^l t_k r_{i_k}$, and because σ is optimal, we obtain the following inequality.

$$b_i + t_1 r_i + \sum_{k=2}^l t_k r_{i_{k-1}} \geq b_{i_l} + \sum_{k=1}^l t_k r_{i_k} \quad (4.4)$$

Combining inequalities 4.3 and 4.4, we obtain that

$$b_i + t_1 r_i + \sum_{k=2}^l t_k r_{i_{k-1}} = b_{i_l} + \sum_{k=1}^l t_k r_{i_k}$$

And therefore $U(\sigma^1) = U(\sigma^0)$, which proves that σ^1 is optimal. As advertised, we found a sequence $\sigma^1 \in \mathbf{Opt}(T, m, R, B)$ such that $\hat{\sigma} \cap \hat{\sigma}^1$ contains one more element than $\hat{\sigma} \cap \hat{\sigma}^0$. By applying this processus recursively, we will eventually obtain a sequence $\sigma' \in \mathbf{Opt}(T, m, R, B)$ such that $\hat{\sigma} \subset \hat{\sigma}'$. \square

Thanks to theorem 4.6, we can find an optimal sequence following the simple algorithm 1. We note $\text{INSERT}(i, S)$ a function that inserts a cactus i in an ordered sequence S such that the resulting sequence is ordered (in increasing order of filling rates), which takes $O(T)$ operations. We note $\text{UTILITY}(S)$ a function computing the utility of a sequence S , which takes $O(T)$ operations as well (see equation 4.1). Finding the sequence of maximal utility requires $O(m)$ operations. Overall, because of the loops, the algorithm requires $O(mT^2)$. The theorem 4.6 guaranties that the sequence S computed by this algorithm is optimal. Therefore, in this model with perfect information, we are always able to find an optimal sequence, in polynomial time at most.

4.4.3 Model with incomplete information

We now turn our attention to the following more complex and more realistic setting, where the bat is only provided with incomplete information. The model makes the following assumptions : first, we limit our analysis to games such that $B = 0$ and $T \leq m$. The main difference with the previous model is that the vector R is initially unknown to the bat. It is instead chosen according to a distribution \mathcal{D} , which the bat knows from the beginning. Furthermore, when the bat visits a cactus, she learns its filling rate and remembers it. We will assume that the distribution \mathcal{D} is such that the rates of individual cacti are independent and identically distributed.

```

begin
   $S \leftarrow \emptyset$ ;
  for  $t \leftarrow 1$  to  $T$  do
    for  $i \leftarrow 1$  to  $m$  do
       $S_i \leftarrow \text{INSERT}(i, S)$ ;
       $U_i \leftarrow \text{UTILITY}(S_i)$ ;
    end
     $j = \arg \max_{i \in [m]} U_i$ ;
     $S \leftarrow S_j$ 
  end
end

```

Algorithm 1: An algorithm finding an optimal sequence in $O(mT^2)$ when $T \leq m$

In this situation, a strategy of the bat can be more complex than a simple sequence of moves, and can be described as an online algorithm A whose input is the distribution \mathcal{D} , the sequence of previously visited cacti, and their filling rates; and whose output is the next cactus to be visited. The algorithm has a fundamental choice to make at each round, which is either to revisit a cactus whose rate is already known, or to visit a new cactus, whose rate is unknown but follows a known distribution. In the case that the algorithm chooses to visit a new cactus, its rate will be independent from any previous choices of the algorithm and from any known information at that point. Since we assumed that the number of cacti m is not less than the number of rounds T , it will always remain possible to visit a new cactus at any point during the game. Therefore the number m of cacti is not really relevant to the algorithm and we can simply think as if cacti were "created" the first time they are visited.

Such an algorithm A might be randomized, therefore we define its utility on a given vector R as follows:

$$U(A, R) = \sum_{\sigma \in S} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma, R) \quad (4.5)$$

Where $\sigma_{A,R}$ is the random variable corresponding to the sequence of visits obtained when executing algorithm A on the instance R . More interestingly, the utility of an algorithm on a given distribution \mathcal{D} is defined by:

$$U(A, \mathcal{D}) = \sum_{R \in \text{supp}(\mathcal{D})} \mathbb{P}_{\mathcal{D}}[R] U(A, R) \quad (4.6)$$

Where $\text{supp}(\mathcal{D})$ denotes the support of distribution \mathcal{D} , and $\mathbb{P}_{\mathcal{D}}[R]$ is the probability of R according to this distribution.

Assuming the distribution \mathcal{D} has a finite support, at each time step, the tree of all possible outcomes is finite. Indeed, it is bounded in depth by the number T of visits and in degree by $T - 1 + \text{card}(\text{supp}(\mathcal{D}))$, where $T - 1$ branches correspond to visited cacti that can be revisited, and $\text{card}(\text{supp}(\mathcal{D}))$ branches correspond to the possible outcomes of visiting a new cactus. Therefore, it is always possible to implement a min-max algorithm

that will maximize the expected utility. However such an algorithm must explore the whole tree, which yields a time complexity of $O(T^{T+\text{card}(\text{supp}(\mathcal{D}))})$. As this complexity is quite impractical, we will try to find a more simple optimal algorithm. Thanks to Theorem 4.7, we can restrict our investigation to a class of algorithms which proceed in two phases. During the first phase, only new cacti are visited and during the second, the algorithm only visits known cacti. We say that there is an *exploration* phase followed by an *exploitation* phase. Since we know from section 4.4.2 what sequence of visits maximizes the utility of the second phase, we can even assume that the second phase follows this optimal sequence. Furthermore, it should be clear that no algorithm in which the first phase is shorter than $\frac{T}{2}$ can be optimal, therefore we will assume that the first phase lasts for at least $\frac{T}{2}$ rounds. Let us now prove what we just claimed.

Definition 4.1. *Let σ be a sequence. We define the following quantities:*

- $\tau_1(\sigma)$ is the last time step in which a cactus is visited for the first time in σ ,
- $\tau_2(\sigma)$ is the first time step in which a cactus is visited for the second time in σ ,
- $\tau_3(\sigma)$ is the first time step greater than $\tau_2(\sigma)$ in which a cactus is visited for the first time in σ ,
- $\tau_4(\sigma)$ is the time step in which $\sigma_{\tau_2(\sigma)}$ is visited for the first time in σ .

Formally:

$$\begin{aligned}\tau_1(\sigma) &= \max\{t \in [T] \mid \forall t' < t, \sigma_{t'} \neq \sigma_t\} \\ \tau_2(\sigma) &= \min\{t \in [T] \mid \exists t' < t, \sigma_{t'} = \sigma_t\} \\ \tau_3(\sigma) &= \min\{t \in \{\tau_2(\sigma), \dots, T\} \mid \forall t' < t, \sigma_{t'} \neq \sigma_t\} \\ \tau_4(\sigma) &= \min\{t \in [T] \mid \sigma_t = \sigma_{\tau_2(\sigma)}\}\end{aligned}$$

We adopt the convention that $\min(\emptyset) = +\infty$ and $\max(\emptyset) = -\infty$.

Theorem 4.7. *Let \mathcal{D} be a probability distribution such that the rates of different cacti are i.i.d and A be an optimal online algorithm for this distribution, i.e it maximizes the utility $U(A, \mathcal{D})$. Let R be an instance such that $\mathbb{P}_{\mathcal{D}}[R] > 0$. Then we must have $\tau_1(\sigma_{A,R}) < \tau_2(\sigma_{A,R})$ with probability 1.*

Proof. Let \bar{r} be the expected value of the filling rate of a random cactus according to \mathcal{D} .

For convenience, we will use the following “light” notations:

- $\tau_1, \tau_2, \tau_3, \tau_4$ instead of $\tau_1(\sigma), \tau_2(\sigma), \tau_3(\sigma), \tau_4(\sigma)$,
- i instead of σ_{τ_2} ,
- j instead of σ_{τ_3} .

Function $A'(A, \sigma, T)$

Input: An algorithm A , a sequence σ and an integer T .

Output: The next cactus to visit, according to the algorithm A' , when the sequence σ describes the sequence of previous visits, and T is the maximum number of visits.

begin

$t \leftarrow \text{card}(\sigma) + 1;$

$\tau_1 \leftarrow \tau_1(\sigma);$

$\tau_2 \leftarrow \tau_2(\sigma);$

$\tau_3 \leftarrow \tau_3(\sigma);$

$\tau_4 \leftarrow \tau_4(\sigma);$

$i \leftarrow \sigma_{\tau_2};$

if $t = \tau_2$ **and** $\tau_2 < \tau_1$ **then**

 | **return** *new cactus*

else if $t = \tau_3$ **and** $r_i(\tau_2 - \tau_4) \geq \bar{r}\tau_2$ **then**

 | $\sigma' \leftarrow (\sigma_1, \dots, \sigma_{\tau_2});$

 | **return** $A(\sigma', T)$

else

 | **return** $A(\sigma, T)$

end

end

end

Algorithm 2: The A' algorithm used in the proof of theorem 4.7

We define an algorithm A' (see algorithm 2). To get a better understanding of A' , we introduce a partition of the set S of sequences into three sets S_1 , S_2 and S_3 , defined as follows.

$$\begin{aligned} S_1 &= \{\sigma \in S \mid \tau_1(\sigma) < \tau_2(\sigma)\} \\ S_2 &= \{\sigma \in S \mid (\tau_2 < \tau_1) \wedge (r_i(\tau_2 - \tau_4) < \bar{r}\tau_2)\} \\ S_3 &= \{\sigma \in S \mid (\tau_2 < \tau_1) \wedge (r_i(\tau_2 - \tau_4) \geq \bar{r}\tau_2)\} \end{aligned}$$

We also introduce two functions $\phi : S_2 \rightarrow S$ and $\psi : S_3 \rightarrow S$, defined as follows.

$$\begin{aligned} \phi(\sigma)_t &= \begin{cases} k & \text{when } t = \tau_2 \\ \sigma_t & \text{otherwise} \end{cases} \\ \psi(\sigma)_t &= \begin{cases} k & \text{when } t = \tau_2 \\ \sigma_{\tau_2} & \text{when } t = \tau_3 \\ \sigma_t & \text{otherwise} \end{cases} \end{aligned}$$

Where k is a *new* cactus.

It is now reasonably easy to verify that the definition of A' implies:

$$\sigma_{A',R} = \begin{cases} \sigma_{A,R} & \text{when } \sigma_{A,R} \in S_1 \\ \phi(\sigma_{A,R}) & \text{when } \sigma_{A,R} \in S_2 \\ \psi(\sigma_{A,R}) & \text{when } \sigma_{A,R} \in S_3 \end{cases}$$

It is now sufficient to prove the following two claims to conclude.

Claim 1: For every instance R ,

$$\sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] U(\phi(\sigma), R) \geq \sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma, R)$$

Furthermore, the inequality becomes strict if $\mathbb{P}[\sigma_{A,R} \in S_2] > 0$.

Claim 2: For every instance R ,

$$\sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] U(\psi(\sigma), R) \geq \sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma, R)$$

Furthermore, the inequality becomes strict if $\mathbb{P}[\sigma_{A,R} \in S_3] > 0$.

Proof of claim 1. Let σ be a sequence in S_2 . The definition of function ϕ implies that $U(\phi(\sigma)) - U(\sigma) = r_k\tau_2 - r_i(\tau_2 - \tau_4)$. Since k is a *new* cactus generated by A' in time step τ_2 , its rate r_k is a random variable independent from τ_2 , and $\mathbb{E}[r_k] = \bar{r}$. For the sake of brevity, we will abusively write $U(\phi(\sigma)) - U(\sigma) = \bar{r}\tau_2 - r_i(\tau_2 - \tau_4)$, which is positive by definition of S_2 . Therefore:

$$\begin{aligned} \sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] (U(\phi(\sigma), R) - U(\sigma, R)) &= \sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] (\bar{r}\tau_2 - r_i(\tau_2 - \tau_4)) \\ &\geq 0 \end{aligned}$$

And the inequality become strict if there is a sequence $\sigma \in S_2$ such that $\mathbb{P}[\sigma_{A,R} = \sigma] > 0$.

Proof of claim 2. Let σ be a sequence in S_3 . The definition of function ψ implies that $U(\psi(\sigma)) - U(\sigma) = r_i\tau_3 + r_k\tau_2 - r_i\tau_2 - r_j\tau_3 = r_i(\tau_3 - \tau_2) + r_k\tau_2 - r_j\tau_3$.

As before, since k is a *new* cactus generated by A' in time step τ_2 , its rate r_k is a random variable independent from τ_2 , and $\mathbb{E}[r_k] = \bar{r}$. Furthermore, by definition, $j = \sigma_{\tau_3}$ is a *new* cactus, generated by algorithm A in round τ_3 . Therefore, r_j is independent from τ_3 and $\mathbb{E}[r_j] = \bar{r}$. It follows that:

$$\sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] (U(\psi(\sigma), R) - U(\sigma, R)) = \sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] (r_i - \bar{r})(\tau_3 - \tau_2)$$

By definition of S_3 , we know that $r_i > \bar{r}$, and by definition, $\tau_3 > \tau_2$. Therefore claim 1 is true.

Conclusion By definition,

$$\begin{aligned} U(A, R) &= \sum_{\sigma \in S} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma) \\ &= \sum_{\sigma \in S_1} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma) + \sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma) + \sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma) \end{aligned}$$

and

$$\begin{aligned} U(A', R) &= \sum_{\sigma \in S} \mathbb{P}_{A',R}[\sigma] U(\sigma) \\ &= \sum_{\sigma \in S_1} \mathbb{P}[\sigma_{A,R} = \sigma] U(\sigma) + \sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] U(\phi(\sigma)) + \sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] U(\psi(\sigma)) \end{aligned}$$

Therefore:

$$\begin{aligned} U(A', R) - U(A, R) &= \\ &= \sum_{\sigma \in S_2} \mathbb{P}[\sigma_{A,R} = \sigma] (U(\phi(\sigma)) - U(\sigma)) + \sum_{\sigma \in S_3} \mathbb{P}[\sigma_{A,R} = \sigma] (U(\psi(\sigma)) - U(\sigma)) \end{aligned}$$

Using the two previous claims, this implies that $U(A', R) \geq U(A, R)$ for every instance R , and $U(A', R) > U(A, R)$ if $\mathbb{P}[\sigma_{A,R} \in S_2 \cup S_3] > 0$. But since A is optimal, we must have $U(A', R) \leq U(A, R)$, therefore $\mathbb{P}[\sigma_{A,R} \in S_1 \cup S_2] = 0$, which is equivalent to $\mathbb{P}[\sigma_{A,R} \in S_1] = 1$, and therefore $\tau_1(\sigma_{A,R}) < \tau_2(\sigma_{A,R})$ with probability 1. \square

Corollary 4.8. *Let \mathcal{D} be a probability distribution such that the rates of different cacti are i.i.d and A be an optimal online algorithm for this distribution. Let R be an instance such that $\mathbb{P}_{\mathcal{D}}[R] > 0$. Then every cactus is present at most twice in $\sigma_{A,R}$, with probability 1.*

Proof. By theorem 4.7, we know that the algorithm A always proceed in two phases: an exploration phase followed by an exploitation phase. During the exploitation phase, the algorithm does not learn any new information about the cacti, since it only visit cacti whose filling rate is already known. Therefore, at the end of the exploration phase, the algorithm is faced by the very problem solved in section 4.4.2. Therefore, by lemma 4.3, each cactus will be visited at most once during the exploitation phase, which is at most twice in total. \square

4.4.4 Uniform distribution

In this section, we study in more details the particular case of the uniform distribution: we assume that according to distribution \mathcal{D} , every cactus has a filling rate chosen independently and uniformly at random in $[0, 1]$. We want to understand how effective the optimal online algorithm can be compared to an ideal situation where every cactus visited has a rate of 1. For this purpose, we define the competitive ratio $\rho(A)$ of the algorithm A as:

$$\rho(A) = \frac{U(A, \mathcal{D})}{U_{max}}, \text{ where } U_{max} = \sum_{t=1}^T t = \frac{T(T+1)}{2}$$

We introduce the distinction between *adaptive algorithms*, in which the switching time between the exploration and the exploitation phases is determined during the execution, and *rigid algorithms*, for which the switching time is determined before the algorithm execution. This section deals only with the latter, in which the switching time is independent from the rates of the generated cacti. The two main results of this section are an upper and a lower bound on the asymptotic competitive ratio of an optimal rigid algorithm, when T goes to infinity.

Let us start with the following very useful lemma, concerning the order statistics of independent uniform distributions.

Lemma 4.9. *Let $(X_i)_{1 \leq i \leq n}$ be n independent random variables following a uniform distribution over $[0, 1]$. Let $X_{(i)}$ be the i -th order statistic of these variables, i.e, the value of the i -th variable when they are arranged in increasing order. Then $\mathbb{E}[X_{(i)}] = \frac{i}{n+1}$.*

Proof. Let x be a real number in $[0, 1]$. We define the following random variable: $N(x) = \sum_{i=1}^n \mathbb{1}_{\{X_i < x\}}$, which counts how many of the X_i are lower than x . The variables $\mathbb{1}_{\{X_i < x\}}$ are independent and follow a Bernoulli distribution of parameter x , therefore $N(x)$ follows a binomial law of parameters n and x . Furthermore, by definition, we have:

$$\mathbb{P}[X_{(i)} \geq x] = \mathbb{P}[N(x) \leq i] = \sum_{k=1}^i \binom{n}{k} x^k (1-x)^{n-k}$$

Now, since $X_{(i)}$ are positive variables, we can compute their expected value as follows:

$$\begin{aligned} \mathbb{E}[X_{(i)}] &= \int_0^1 \mathbb{P}[X_{(i)} \geq x] dx \\ &= \sum_{k=1}^i \binom{n}{k} \int_0^1 x^k (1-x)^{n-k} dx \end{aligned}$$

By an integration by parts, we have for every $k \in \{1, \dots, n-1\}$,

$$\int_0^1 x^k (1-x)^{n-k} dx = \frac{k}{n-k+1} \int_0^1 x^{k-1} (1-x)^{n-k+1} dx$$

Which, by induction, gives:

$$\int_0^1 x^k (1-x)^{n-k} dx = \frac{1}{\binom{n}{k}(n+1)}$$

And therefore,

$$\mathbb{E}[X_{(i)}] = \sum_{k=1}^i \frac{1}{n+1} = \frac{i}{n+1}$$

□

Upper bound Now, let us study the performance of the optimal rigid algorithm, which we call $A(\alpha, T)$, defined as follows.

Definition 4.2. *The behavior of $A(\alpha, T)$ is the following:*

1. *For the first $\lfloor \alpha T \rfloor$ rounds, explore unvisited cacti.*
2. *For the remaining $T - \lfloor \alpha T \rfloor$ rounds, apply the optimal offline algorithm described in section 4.4.2.*

Let σ be the random sequence of visits resulting from the execution of algorithm $A(\alpha, T)$. We will compute its utility using equation 4.2, which we will rewrite as follows. For convenience, we use the notation $\tau(\sigma, t) = t_{\sigma_t}(\sigma_1, \dots, \sigma_{t-1})$, i.e, $\tau(\sigma, t)$ corresponds to the time of the previous visit of the cacti visited in round t .

$$\begin{aligned} U(\sigma) &= U_1(\sigma) + U_2(\sigma) \\ \text{with } U_1(\sigma) &= \sum_{t=1}^{\alpha T} r_{\sigma_t}(t - \tau(\sigma, t)) \\ \text{and } U_2(\sigma) &= \sum_{t=\alpha T+1}^T r_{\sigma_t}(t - \tau(\sigma, t)) \end{aligned}$$

Therefore, the algorithm's competitive ratio is

$$\rho(\alpha, T) = \frac{\mathbb{E}[U(\sigma)]}{U_{max}} = \frac{2\mathbb{E}[U(\sigma)]}{T(T+1)}$$

and its asymptotic competitive ratio is

$$\rho(\alpha) = \lim_{T \rightarrow +\infty} \rho(\alpha, T)$$

We now prove the following technical lemma, which we will use in the proof of the upper bound on $\rho(\alpha)$.

Lemma 4.10. *Let σ be the random sequence resulting from the execution of algorithm $A(\alpha, T)$. For all $i \in \{\alpha T + 1, \dots, T\}$, let g_i be the function from sequences to sequences that swaps the cacti visited in rounds $\tau(\sigma, i)$ and $T + 1 - i$, i.e.,*

- $(g_i(\sigma))_{T+1-i} = \sigma_{\tau(\sigma, i)},$
- $(g_i(\sigma))_{\tau(\sigma, i)} = \sigma_{T+1-i},$
- and for all $j \in [T] \setminus \{\tau(\sigma, i), T + 1 - i\}, (g_i(\sigma))_t = \sigma_t$

Let us define the function $g = g_{\alpha T + 1} \circ \dots \circ g_T$.

Then, $U_2(g(\sigma)) \geq U_2(\sigma)$, and for all $t \in \{\alpha T + 1, \dots, T\}$,

- $\tau(g(\sigma), t) = T + 1 - t,$
- $g(\sigma)_t = \sigma_t.$

Proof. We proceed by induction, showing that every step g_i can only increase the value of U_2 .

Base case: The first step is applying function g_T , which swaps the first cactus visited and the cactus visited in round $\tau(\sigma, T)$. The reader can refer to figure 4.12 for a visualization of the two sequences σ and $g_T(\sigma)$. By definition of $A(\alpha)$, every cactus is visited at most twice in σ , once before αT and once after $\alpha T + 1$. There are three cases to distinguish:

Case 0: $\sigma_1 = \sigma_T$, i.e., $\tau(\sigma, T) = 1$. In this case we have $g_T(\sigma) = \sigma$.

Case 1: σ_1 is not revisited. In this case,

- $\tau(g_T(\sigma), T) = 1 < \tau(\sigma, T)$, and
- $\tau(g_T(\sigma), t) = \tau(\sigma, t)$ for all $t \in \{\alpha T + 1, \dots, T - 1\}$,

and therefore $U_2(g_T(\sigma)) > U_2(\sigma)$.

Case 2: σ_1 is revisited, i.e., there exist $t^* \in \{\alpha T + 1, \dots, T - 1\}$ such that $\sigma_{t^*} = \sigma_1$ and $\tau(\sigma, t^*) = 1$. In this case,

- $\tau(g_T(\sigma), T) = \tau(\sigma, t^*) = 1,$
- $\tau(g_T(\sigma), t^*) = \tau(\sigma, T) > 1,$
- $\tau(g_T(\sigma), t) = \tau(\sigma, t)$ for all $t \in \{\alpha T + 1, \dots, T - 1\} \setminus \{t^*\},$

and therefore,

$$\begin{aligned} U_2(g_T(\sigma)) - U_2(\sigma) &= r_{\sigma_T}(\tau(\sigma, T) - 1) + r_{\sigma_{t^*}}(1 - \tau(\sigma, T)) \\ &= (r_{\sigma_T} - r_{\sigma_{t^*}})(\tau(\sigma, T) - 1) \end{aligned}$$

Which is positive since $r_{\sigma_T} \geq r_{\sigma_{t^*}}$, by lemma 4.4.

$$\begin{array}{rcccccccc}
\text{round} & : & 1 & 2 & \cdots & \tau(\sigma, T) & \cdots & \alpha T & \alpha T + 1 & \cdots & T \\
\sigma & : & \sigma_1 & \sigma_2 & \cdots & \sigma_T & \cdots & \sigma_{\alpha T} & \sigma_{\alpha T + 1} & \cdots & \sigma_T \\
g_T(\sigma) & : & \sigma_T & \sigma_2 & \cdots & \sigma_1 & \cdots & \sigma_{\alpha T} & \sigma_{\alpha T + 1} & \cdots & \sigma_T
\end{array}$$

Figure 4.12 – Visualization of the two sequences σ and $g_T(\sigma)$ defined in the first part of the proof of lemma 4.10.

Induction step Now, let i be an integer in $\{\alpha T + 1, \dots, T - 1\}$. For convenience, we note $\sigma^{i+1} = g_{i+1} \circ \dots \circ g_T(\sigma)$ and $\sigma^i = g_i(\sigma^{i+1})$. The reader can refer to figure 4.13 for a visualization of the two sequences σ^{i+1} and σ^i . Let us compare $U_2(\sigma^i)$ and $U_2(\sigma^{i+1})$. The function g_i swaps the cacti visited in rounds $\tau(\sigma^{i+1}, i)$ and $T + 1 - i$. By induction, for all $j \in \{i + 1, \dots, T\}$, we have $\tau(\sigma^{i+1}, j) = T + 1 - j$. Therefore $\tau(\sigma^{i+1}, i) \notin \{1, \dots, T - i\}$. We can now distinguish three cases.

Case 0: $\tau(\sigma^{i+1}, i) = T + 1 - i$, i.e, $\sigma_{T+1-i}^{i+1} = \sigma_i^{i+1}$. In this case we have $\sigma^i = \sigma^{i+1}$.

Case 1: σ_{T+1-i}^{i+1} is not revisited in σ^{i+1} . In this case,

- $\tau(\sigma^i, i) = T + 1 - i < \tau(\sigma^{i+1}, i)$, and
- $\tau(\sigma^i, t) = \tau(\sigma^{i+1}, t)$ for all $t \in \{\alpha T + 1, \dots, T\} \setminus \{i\}$,

and therefore $U_2(\sigma^i) > U_2(\sigma^{i+1})$.

Case 2: σ_{T+1-i}^{i+1} is revisited in σ^{i+1} , i.e, there exist $t^* \in \{\alpha T + 1, \dots, i\}$ such that $\sigma_{t^*}^{i+1} = \sigma_{T+1-i}^{i+1}$ and $\tau(\sigma^{i+1}, t^*) = T + 1 - i$. In this case,

- $\tau(\sigma^i, i) = \tau(\sigma^{i+1}, t^*) = T + 1 - i$,
- $\tau(\sigma^i, t^*) = \tau(\sigma^{i+1}, i) > T + 1 - i$,
- $\tau(\sigma^i, t) = \tau(\sigma^{i+1}, t)$ for all $t \in \{\alpha T + 1, \dots, T\} \setminus \{i, t^*\}$,

and therefore,

$$\begin{aligned}
U_2(\sigma^i) - U_2(\sigma^{i+1}) &= r_{\sigma_i^{i+1}}(\tau(\sigma^{i+1}, i) - (T + 1 - i)) + r_{\sigma_{t^*}^{i+1}}(1 - \tau(\sigma^{i+1}, i)) \\
&= (r_{\sigma_i^{i+1}} - r_{\sigma_{t^*}^{i+1}})(\tau(\sigma^{i+1}, i) - 1)
\end{aligned}$$

Which is positive since $r_{\sigma_i^{i+1}} \geq r_{\sigma_{t^*}^{i+1}}$, by lemma 4.4. \square

We can now state and prove the following upper bound on the asymptotic competitive ratio of an optimal rigid algorithm.

$$\begin{array}{cccccccccccccccc}
\text{round} & : & 1 & \cdots & T-i & T+1-i & \cdots & \tau(\sigma^{i+1}, i) & \cdots & \alpha T & \alpha T+1 & \cdots & T \\
\sigma^{i+1} & : & \sigma_T & \cdots & \sigma_{i+1} & \sigma_{T+1-i}^{i+1} & \cdots & \sigma_i & \cdots & \sigma_{\alpha T}^{i+1} & \sigma_{\alpha T+1} & \cdots & \sigma_T \\
\sigma^i & : & \sigma_T & \cdots & \sigma_{i+1} & \sigma_i & \cdots & \sigma_{T+1-i}^{i+1} & \cdots & \sigma_{\alpha T}^{i+1} & \sigma_{\alpha T+1} & \cdots & \sigma_T
\end{array}$$

Figure 4.13 – Visualization of the two sequences σ^{i+1} and σ^i defined in the induction step of the proof of lemma 4.10.

Theorem 4.11. *When \mathcal{D} is the uniform distribution over $[0, 1]$ and when T goes to infinity, the best competitive ratio achieved by a rigid algorithm $A(\alpha, T)$ is at most $\rho \simeq 0.6037$, i.e., for all $\alpha \in [\frac{1}{2}, 1]$, $\rho(\alpha) \leq \rho$.*

Proof. Let σ be the random sequence of visits obtained when executing the algorithm $A(\alpha)$. The proof proceeds as follows: first, we obtain a majoration of $\rho(\alpha, T)$ by a function $f(\alpha, T)$. Then we compute the limit $f(\alpha) = \lim_{T \rightarrow +\infty} f(\alpha, T)$. And finally we compute $M = \sup_{\frac{1}{2} \leq \alpha \leq 1} f(\alpha)$, which is an upper bound of the asymptotic competitive ratio $\rho(\alpha)$.

For simplicity, we assume that αT is an integer, i.e., we study the algorithm $A' = A\left(\frac{\lfloor \alpha T \rfloor}{T}\right)$. We will see further in the proof that $\rho(\alpha)$ is continuous in α . Since $\lim_{T \rightarrow +\infty} \frac{\lfloor \alpha T \rfloor}{T} = \alpha$, it is clear that this assumption will not affect the final result.

During the first phase of the algorithm, only new cacti are visited, therefore, for all $t \in \{1, \dots, \alpha T\}$, $\mathbb{E}[r_{\sigma_t}] = \frac{1}{2}$ and $\tau(\sigma, t) = 0$, and we can explicit $\mathbb{E}[U_1(\sigma)]$ as follows:

$$\begin{aligned}
\mathbb{E}[U_1(\sigma)] &= \sum_{t=1}^{\alpha T} \mathbb{E}[r_{\sigma_t}(t - \tau(\sigma, t))] \\
&= \sum_{t=1}^{\alpha T} \mathbb{E}[r_{\sigma_t}]t \\
&= \frac{1}{2} \sum_{t=1}^{\alpha T} t \\
&= \frac{\alpha T(\alpha T + 1)}{4}
\end{aligned}$$

Unfortunately, we cannot use the same method to compute $\mathbb{E}[U_2(\sigma)]$, since we do not know explicitly how the algorithm selects cacti to be revisited. However, thanks to lemma 4.10, we can obtain an upper bound. Indeed, we know that $U_2(\sigma) \leq U_2(g(\sigma))$, and

we can bound the value of $U_2(g(\sigma))$ as follows:

$$\begin{aligned} U_2(g(\sigma)) &= \sum_{t=\alpha T+1}^T r_{g(\sigma)_t}(t - \tau(g(\sigma), t)) \\ &= \sum_{t=\alpha T+1}^T r_{\sigma_t}(2t - T - 1) \end{aligned}$$

Furthermore, using lemmas 4.4 and 4.9, we know that $\mathbb{E}[r_{\sigma_t}] \leq \frac{(\alpha-1)T+t}{\alpha T+1}$, therefore

$$\begin{aligned} \mathbb{E}[U_2(g(\sigma))] &\leq \sum_{t=\alpha T+1}^T \frac{((\alpha-1)T+t)(2t-T-1)}{\alpha T+1} \\ &\leq \frac{2}{\alpha T+1} \sum_{t=\alpha T+1}^T t^2 + \frac{(2\alpha-3)T-1}{\alpha T+1} \sum_{t=\alpha T+1}^T t + \frac{(1-\alpha)T^2 + (1-\alpha)T}{\alpha T+1} \sum_{t=\alpha T+1}^T 1 \\ &\leq \frac{2(2(1-\alpha^3)T^3 + 3(1-\alpha^2)T^2 + (1-\alpha)T)}{6(\alpha T+1)} + \frac{((2\alpha-3)T-1)((\alpha+1)T+1)(1-\alpha)T}{2(\alpha T+1)} \\ &\quad + \frac{((1-\alpha)T^2 + (1-\alpha)T)(1-\alpha)T}{\alpha T+1} \\ &\leq \frac{(-10\alpha^3 + 15\alpha^2 - 6\alpha + 1)T^3 + (-3\alpha^2 + 3\alpha)T^2 + (\alpha-1)T}{6(\alpha T+1)} \end{aligned}$$

Putting the previous inequalities together, we obtain:

$$\begin{aligned} \mathbb{E}[U(\sigma)] &= \mathbb{E}[U_1(\sigma)] + \mathbb{E}[U_2(\sigma)] \\ &\leq \frac{\alpha T(\alpha T+1)}{4} + \frac{(-10\alpha^3 + 15\alpha^2 - 6\alpha + 1)T^3 + (-3\alpha^2 + 3\alpha)T^2 + (\alpha-1)T}{6(\alpha T+1)} \\ &\leq \frac{(-17\alpha^3 + 30\alpha^2 - 12\alpha + 2)T^3 + 6\alpha T^2 + (5\alpha - 2)T}{12(\alpha T+1)} \end{aligned}$$

Therefore:

$$\rho(\alpha) \leq \frac{-17\alpha^3 + 30\alpha^2 - 12\alpha + 2}{6\alpha}$$

We note f the rational function

$$f : \alpha \in [\frac{1}{2}, 1] \mapsto \frac{-17\alpha^3 + 30\alpha^2 - 12\alpha + 2}{6\alpha} \quad (4.7)$$

We want to find $\max_{\alpha \in [\frac{1}{2}, 1]} f(\alpha)$, therefore we compute the derivative of f :

$$f'(\alpha) = \frac{-17\alpha^3 + 15\alpha^2 - 1}{3\alpha^2} \quad (4.8)$$

We find that f' has a single root in the interval $[\frac{1}{2}, 1]$, which is approximately equal to 0.7875, and corresponds to a maximum. Substituting this value in $f(\alpha)$ gives: $\max_{\alpha \in [\frac{1}{2}, 1]} f(\alpha) \simeq 0.6037$. \square

Lower bound We now want to obtain a lower bound on the asymptotic competitive ratio of rigid algorithms. For this purpose, we define a family of rigid algorithms depending on two parameters, that we note α and β .

Definition 4.3. *We define the class of algorithms $A(\alpha, \beta)$, parametrized by $\alpha \in [\frac{1}{2}, 1]$ and $\beta \in [(1 - \alpha), \alpha]$. On any instance R , the behavior of algorithm $A(\alpha, \beta)$ is the following:*

1. *For the first $\lfloor \alpha T \rfloor$ rounds, explore unvisited cacti.*
2. *For the remaining $T - \lfloor \alpha T \rfloor$ rounds, revisit the cacti with highest rate among the cacti previously opened before round $\lfloor \beta T \rfloor$, in increasing order of filling rates.*

As before, we note σ the random sequence of visits resulting from the execution of $A(\alpha, \beta)$, and we define the competitive ratio of the algorithm as

$$\rho(\alpha, \beta, T) = \frac{\mathbb{E}[U(\sigma)]}{U_{max}} = \frac{2\mathbb{E}[U(\sigma)]}{T(T+1)}$$

and its asymptotic competitive ratio as

$$\rho(\alpha, \beta) = \lim_{T \rightarrow +\infty} \rho(\alpha, \beta, T)$$

Theorem 4.12. *When \mathcal{D} is the uniform distribution over $[0, 1]$ and when T goes to infinity, the best competitive ratio achieved by an $A(\alpha, \beta)$ algorithm is $\rho \simeq 0.5404$. It is achieved for the following choices of parameters: $\alpha \simeq 0.8703$ and $\beta \simeq 0.3442$.*

Proof. For simplicity, we assume that αT and βT are both integers, i.e, we study the algorithm $A' = A\left(\frac{\lfloor \alpha T \rfloor}{T}, \frac{\lfloor \beta T \rfloor}{T}\right)$. We will see further in the proof that $\rho(\alpha, \beta)$ is continuous in α and β . Since $\lim_{T \rightarrow +\infty} \frac{\lfloor \alpha T \rfloor}{T} = \alpha$ and $\lim_{T \rightarrow +\infty} \frac{\lfloor \beta T \rfloor}{T} = \beta$, it is clear that this assumption will not affect the final result.

In order to compute $\mathbb{E}[U(\sigma)]$, we define the following partition of $\hat{\sigma}$, the set of visited cacti:

$$\hat{\sigma} = \Sigma_1 + \Sigma_2 + \Sigma_3$$

with

$$\begin{aligned} \Sigma_1 &= \{i \in \hat{\sigma} \mid t_i(\sigma) \leq \beta T\} \\ \Sigma_2 &= \{i \in \hat{\sigma} \mid \beta T + 1 \leq t_i(\sigma) \leq \alpha T\} \\ \Sigma_3 &= \{i \in \hat{\sigma} \mid \alpha T + 1 \leq t_i(\sigma) \leq T\} \end{aligned}$$

i.e, Σ_1 is the set of cacti that are visited before round βT and not selected by the algorithm to be revisited, Σ_2 is the set of cacti visited between rounds $\beta T + 1$ and αT , which are not revisited by definition, and Σ_3 is the set of cacti visited before round βT and selected by the algorithm to be revisited after round αT .

To compute $\mathbb{E}[U(\sigma)]$, we use equation 4.1:

$$\mathbb{E}[U(\sigma)] = \sum_{i \in \Sigma_1} \mathbb{E}[r_i t_i(\sigma)] + \sum_{i \in \Sigma_2} \mathbb{E}[r_i t_i(\sigma)] + \sum_{i \in \Sigma_3} \mathbb{E}[r_i t_i(\sigma)]$$

To compute the first sum, we make use of the fact that cacti are selected to be revisited independently of their visit time, therefore for $i \in \Sigma_1$, the variables r_i and $t_i(\sigma)$ are also independent, and $\mathbb{E}[t_i(\sigma)] = \frac{\beta T + 1}{2}$. To obtain $\mathbb{E}[r_i]$, we use lemma 4.9, with $n = \beta T$. Indeed, the cacti in Σ_1 are the $\beta T - (T - \alpha T)$ cacti with the worst rates among cacti visited before round βT . These observations lead to the following expression:

$$\begin{aligned} \sum_{i \in \Sigma_1} \mathbb{E}[r_i t_i(\sigma)] &= \sum_{i \in \Sigma_1} \mathbb{E}[r_i] \mathbb{E}[t_i(\sigma)] \\ &= \frac{\beta T + 1}{2} \sum_{i \in \Sigma_1} \mathbb{E}[r_i] \\ &= \frac{\beta T + 1}{2} \sum_{i=1}^{(\alpha + \beta - 1)T} \frac{i}{\beta T + 1} \\ &= \frac{1}{2} \sum_{i=1}^{(\alpha + \beta - 1)T} i \\ &= \frac{((\alpha + \beta - 1)T + 1)(\alpha + \beta - 1)T}{4} \end{aligned}$$

To compute the second sum, we simply remark that these cacti are not selected by the algorithm, and therefore their rates are simply distributed uniformly in $[0, 1]$, independently of their visit time. Therefore, $\mathbb{E}[r_i t_i(\sigma)] = \frac{1}{2} \mathbb{E}[t_i(\sigma)]$ for every $i \in \Sigma_2$. Therefore:

$$\begin{aligned} \sum_{i \in \Sigma_2} \mathbb{E}[r_i t_i(\sigma)] &= \frac{1}{2} \sum_{i \in \Sigma_2} \mathbb{E}[t_i(\sigma)] \\ &= \frac{1}{2} \sum_{t=\beta T+1}^{\alpha T} t \\ &= \frac{((\alpha + \beta)T + 1)(\alpha - \beta)T}{4} \end{aligned}$$

To compute the third sum, we use lemma 4.9 again. This time, the cacti rates are not independent of their visit time. However, for every $t \in \{\alpha T + 1, \dots, T\}$, we know that the cactus visited in round t has the $(\beta T + t - T)$ -th smallest filling rate among the first βT cacti visited: the last cactus visited has the best rate, the second to last has the second

best rate, etc. Therefore:

$$\begin{aligned}
\sum_{i \in \Sigma_3} \mathbb{E}[r_i t_i(\sigma)] &= \sum_{t=\alpha T+1}^T \left(\frac{(\beta-1)T+t}{\beta T+1} \right) t \\
&= \frac{(\beta-1)T}{\beta T+1} \sum_{t=\alpha T+1}^T t + \frac{1}{\beta T+1} \sum_{t=\alpha T+1}^T t^2 \\
&= \frac{(\beta-1)T((\alpha+1)T+1)(1-\alpha)T}{2(\beta T+1)} \\
&\quad + \frac{1}{\beta T+1} \left(\frac{(1-\alpha^3)T^3}{3} + \frac{(1-\alpha^2)T^2}{2} + \frac{(1-\alpha)T}{6} \right)
\end{aligned}$$

Using the formula for square pyramidal numbers.

Putting the three sums together, and reducing to the same denominator, we obtain the following expressions for $\mathbb{E}[U(\sigma)]$ and $\rho(\alpha, \beta, T)$:

$$\mathbb{E}[U(\sigma)] = \frac{(-4\alpha^3 + 6\alpha\beta^2 + 6\alpha^2 - 6\beta^2 - 6\alpha\beta + 9\beta - 2)T^3 + (6\alpha\beta - 3\beta + 3)T^2 + (4\alpha - 1)T}{12(\beta T + 1)}$$

$$\rho(\alpha, \beta, T) = \frac{(-4\alpha^3 + 6\alpha\beta^2 + 6\alpha^2 - 6\beta^2 - 6\alpha\beta + 9\beta - 2)T^2 + (6\alpha\beta - 3\beta + 3)T + 4\alpha - 1}{6(T+1)(\beta T + 1)}$$

We can now obtain an expression of $\rho(\alpha, \beta) = \lim_{T \rightarrow +\infty} \rho(\alpha, \beta, T)$:

$$\rho(\alpha, \beta) = \frac{-4\alpha^3 + 6\alpha\beta^2 + 6\alpha^2 - 6\beta^2 - 6\alpha\beta + 9\beta - 2}{6\beta}$$

It remains to find $M = \sup_{\alpha, \beta} \rho(\alpha, \beta)$. For this purpose, we compute the partial derivatives of $\rho(\alpha, \beta)$:

$$\begin{aligned}
\frac{\partial \rho}{\partial \alpha}(\alpha, \beta) &= \frac{-2\alpha^2 + \beta^2 + 2\alpha - \beta}{\beta} \\
\frac{\partial \rho}{\partial \beta}(\alpha, \beta) &= \frac{2\alpha^3 + 3\alpha\beta^2 - 3\alpha^2 - 3\beta^2 + 1}{3\beta^2}
\end{aligned}$$

And we find that there is a single maximum satisfying $\alpha \in [\frac{1}{2}, 1]$ and $\beta \in [1 - \alpha, \alpha]$. It is achieved with the values $\alpha \simeq 0.8703$ and $\beta \simeq 0.3442$. Finally, we obtain the value of $M \simeq 0.5404$. \square

4.5 Conclusion and future research directions

This chapter can briefly be described as an attempt to build stronger bridges between ethology (the study of animal behavior), and algorithmic game theory. In section 4.3,

we used computer simulations to gain some insight about a particular ecosystem. This approach offers some advantages, mostly a reduced cost compared with a field expedition, and the possibility to be more speculative about the setting, and explore alternate possibilities. In section 4.4, we drew inspiration from ethology to design an interesting optimisation problem, which in turn reinforced some of our theories about the actual bats behavior.

For future research, we envision another incrementation on our analytical model, which would possess the following characteristics. We would like to model competition between bats by the means of a stochastic process which would cause the cacti to randomly empty, as if visited by another bat. We believe that such a model would be at the same time much more realistic and might be analytically tractable. More precisely, we would try to show that in such a model, the optimal algorithms feature some periodicity, which would be reminiscent of the flight patterns of the actual bats.

About the simulations, it is of course always possible to make them more realistic and more complex on certain aspects, but in doing so we run two risks. First, there is the risk to overfit the computer model to the field data, and thus lose a part of its explicative and predictive power. Secondly, it is important to keep in mind that the core object of study is the bat mind, of which very little is known. To make the model more realistic, it might be necessary to make more assumptions about the way these bats actually think, and such hypotheses might be wrong, leading to a model that is, in fact, less realistic.

Chapter 5

Conclusion

As we presented in the introduction, this thesis is an attempt to successfully apply algorithmic game theory to network and population games, while maintaining what we find the most interesting characteristic of such games, which is the fact that every player only possess a small fraction of the information that is present in the whole population. The three main chapters of the thesis reflect quite well the chronology of our understanding of the problems that seem intrinsically linked to this enterprise.

The approach presented in the second chapter is very abstract and very general. It showed that it is possible to model games with many players and massively imperfect information without sacrifices to simplifying assumptions. Furthermore, we were able to show that the standard concept of Nash equilibrium, although refined into the more specific trembling-hand equilibrium (or the sequential equilibrium), is still valid to analyze such games. However, when we tried to apply this theoretical framework to more concrete examples, we found that the sheer complexity of this abstract model made any attempt to apply standard analytical methods to specific instances very difficult, if not impossible. In fact, we were not able to analyze any concrete game beyond the very trivial ones.

In Chapter 3, we tried to overcome these difficulties by analyzing a specific game using a specific framework. More precisely, we studied a dispersal game, using the concepts of Ideal Free Distribution (IFD) and Evolutionary Stable Strategy (ESS), and introducing the concept of *coverage*. Furthermore, we adopted a point of view that is leaning on the field of *mechanism design*, sometimes called reverse game theory. Indeed, the goal of this approach is to find the parameters of the game studied that maximize some function of the outcomes. In our case, we wanted to maximize the coverage. This analysis led to interesting findings : we showed that the ideal free distribution is always an ESS in this game, we found the sharing policy that maximizes the coverage, and we called it the exclusive policy because of its characteristics. After this analysis, we feel that this dispersal game is well understood, both on its game theoretic aspect and on its mechanism design aspect. However, we found it difficult to generalize the results of this analysis to any broader class of games. Both the method and the findings seem to be closely tied to the specific game that we chose to analyze. In summary, we encountered a problem opposite to the conclusion about our first chapter : our model was a bit too specific to be widely applicable.

The fourth chapter is quite different of the previous ones on a few aspects. First, it is

the fruit of an international and interdisciplinary effort conducted over at least four years. Second, it encompasses many approaches at solving the same problem. The problem studied is to try to understand the foraging behavior of the *Leptos* bats while staying on their seasonal residence in the Sonoran desert. The approaches we used went from the very practical, the field trip and experiments, to the very abstract, a mathematical model of the situation, and included multi-purpose computer simulations. The simulations were very successful at recreating some aspects of the natural setting, which led us to have some confidence that our assumptions were reasonably close to the “truth”. This confidence allowed us to extrapolate on these results, speculate on some of the causes of the phenomena that we observed, and make some predictions about what we might observe if the proportion of bats to cacti were to change in the future, such as the increase or decrease of territorial behavior, and the increase or decrease of fitness for the bats. We also hope that these insights will be helpful in designing new experiments that will improve further our understanding of *Leptos* bats. The last section of the fourth chapter is a bit of a long shot. Trying to link two fields as wildly different as biology and mathematics is not an easy task. However, we believe that a lot of good science can result from these interactions. The mathematical model introduced here is interesting in itself, but developed further, it could offer a strong support to theories about the actual functioning of a bat’s brain. Such approaches have already been successful in the study of ants [26, 27, 29, 31].

Over the three chapters of this thesis, we have striven to solve the problems of many-agent game theory using diverse methods, mainly different mathematical models and computer simulations, with reasonable success but not without the feeling that the question was a bit too difficult to be attacked only by theoretical tools. Indeed, with these tools, we still have to make a new abstract model from scratch for every game that we want studied. Therefore we feel that we haven’t yet found the correct method to analyse such games. Since, in their complexity, these games share some aspects with complex systems; mainly the fact that the outcome of the game is often the result of a sum of many insignificant decisions by multiple agents, and that the equilibrium can be seen as an emergent property; we would be interested to try analyzing the matter with the tools of complex system theory. Another interesting research direction would be to incorporate to the analysis some tools from the era of big data, such as machine learning, neural networks or deep learning. Studying the behavior of artificial neural networks trained to play games has recently had some major implications in the study of the game of go, as the artificial intelligence AlphaGo, developed by DeepMind technologies, has reached and exceeded human-level at the game in 2016. We could imagine that training a neural network to play games inspired from real-world problems could lead to a better understanding of these games. This could simplify the design of efficient nudges, or allow the anticipation of the players strategies, or change of strategies.

Bibliography

- [1] Ittai Abraham, Danny Dolev, and Joseph Y. Halpern. Distributed protocols for leader election: A game-theoretic perspective. In *27th International Symposium on Distributed Computing (DISC)*, pages 61–75, 2013.
- [2] Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236, 2011.
- [3] Yehuda Afek, Yehonatan Ginzberg, Shir Landau Feibish, and Moshe Sulamy. Distributed computing building blocks for rational agents. In *Symposium on Principles of Distributed Computing (PODC)*, pages 406–415. ACM, 2014.
- [4] Yehuda Afek, Shaked Rafaeli, and Moshe Sulamy. Cheating by duplication: Equilibrium requires global knowledge. Technical report, arXiv, 2017.
- [5] Susanne Albers and Matthias Hellwig. Online makespan minimization with parallel schedules. *CoRR*, abs/1304.5625, 2013.
- [6] Noga Alon, László Babai, and Alon Itai. A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem. *J. Algorithms*, 7(4):567–583, 1986.
- [7] Chen Avin, Avi Cohen, Pierre Fraigniaud, Zvi Lotker, and David Peleg. Preferential attachment as a unique equilibrium. In *The Web Conference (WWW)*, New-York, 2018. ACM.
- [8] Leonid Barenboim and Michael Elkin. *Distributed Graph Coloring: Fundamentals and Recent Developments*. Synthesis Lectures on Distributed Computing Theory. Morgan & Claypool Publishers, 2013.
- [9] Sayan Bhattacharya, Sungjin Im, Janardhan Kulkarni, and Kamesh Munagala. Coordination mechanisms from (almost) all scheduling policies. In *Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12-14, 2014*, pages 121–134, 2014.
- [10] Dietrich Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12(1):258–268, 1968.

- [11] Dietrich Braess, Anna Nagurney, and Tina Wakolbinger. On a paradox of traffic planning. *Transportation science*, 39(4):446–450, 2005.
- [12] C. Cannings Broom, M. and G. Vickers. Multi-player matrix games. *Bulletin of Mathematical Biology*, 59(3):931–952, 1997.
- [13] Mark Broom and Jan Rychtár. *Game-theoretical models in biology*. CRC Press, 2013.
- [14] Maciej Bukowski and Jacek Miekisz. Evolutionary and asymptotic stability in symmetric multi-player games. *International Journal of Game Theory*, 33(1):41–54, 2004.
- [15] A. Calvó-Armengol and J. de Martí Beltran. Information gathering in organizations: Equilibrium, welfare and optimal network structure. *Journal of the European Economic Association*, 7:116–161, 2009.
- [16] A. Calvó-Armengol, J. de Martí Beltran, and Prat A. Communication and influence. *Theoretical Economics*, 10:649–690, 2015.
- [17] Subir K Chakrabarti. Equilibrium in Behavior Strategies in Infinite Extensive Form Games with Imperfect Information. *Economic Theory*, 2(4):481–494, 1992.
- [18] Xi Chen and Xiaotie Deng. Settling the complexity of two-player nash equilibrium. In *Foundations of Computer Science, 2006. FOCS’06. 47th Annual IEEE Symposium on*, pages 261–272. IEEE, 2006.
- [19] Chris Cosner, Juan Dávila, and Salomé Martínez. Evolutionary stability of ideal free nonlocal dispersal. *Journal of Biological Dynamics*, 6(2):395–405, 2012.
- [20] Ross Cressman and Vlastimil Křivan. Migration dynamics for the ideal free distribution. *The American Naturalist*, 168(3):384–397, 2006.
- [21] Ross Cressman and Vlastimil Křivan. The ideal free distribution as an evolutionarily stable state in density-dependent population games. *Oikos*, 119(8):1231–1242, 2010.
- [22] Artur Czumaj and Berthold Vöcking. Tight bounds for worst-case equilibria. *ACM Trans. Algorithms*, 3(1):4:1–4:17, February 2007.
- [23] Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.
- [24] Dina K.N. Dechmann, Silke L. Heucke, Luca Giuggioli, Kamran Safi, Christian C. Voigt, and Martin Wikelski. Experimental evidence for group hunting via eavesdropping in echolocating bats. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1668):2721–2728, 2009.

- [25] Avigdor Eldar. Social conflict drives the evolutionary divergence of quorum sensing. *Proceedings of the National Academy of Sciences*, 108(33):13635–13640, 2011.
- [26] Yuval Emek, Tobias Langner, David Stolz, Jara Uitto, and Roger Wattenhofer. How many ants does it take to find the food? *Theor. Comput. Sci.*, 608:255–267, 2015.
- [27] Yuval Emek, Tobias Langner, Jara Uitto, and Roger Wattenhofer. Solving the ANTS problem with asynchronous finite state machines. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pages 471–482, 2014.
- [28] Ky Fan. Fixed-Point and Minimax Theorems in Locally Convex Topological Linear Spaces. *PNAS*, 38(2):121–126, 1952.
- [29] Ofer Feinerman and Amos Korman. The ANTS problem. *Distributed Computing*, 30(3):149–168, 2017.
- [30] Ofer Feinerman and Amos Korman. Individual versus collective cognition in social insects. *Journal of Experimental Biology*, 220(1):73–82, 2017.
- [31] Ehud Fonio, Yael Heyman, Lucas Boczkowski, Aviram Gelblum, Adrian Kosowski, Amos Korman, and Ofer Feinerman. A locally-blazed ant trail achieves efficient collective navigation despite limited information. *eLife*, 5:e20185, 2016.
- [32] Pierre Fraigniaud, Amos Korman, and Yoav Rodeh. Parallel exhaustive search without coordination. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 312–323, 2016.
- [33] Peter Frazier, David Kempe, Jon Kleinberg, and Robert Kleinberg. Incentivizing exploration. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 5–22. ACM, 2014.
- [34] Stephen Dewitt Fretwell and Henry L. Lucas Jr. On territorial behavior and other factors influencing habitat distribution in birds. *Acta biotheoretica*, 19(1):16–32, 1969.
- [35] Stephen Dewitt Fretwell and Henry L Lucas. On territorial behavior and other factors influencing habitat distribution in birds. *Acta biotheoretica*, 19(1):16–36, 1969.
- [36] Drew Fudenberg and David Levine. Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games. *Journal of Economic Theory*, 31(2):251–268, 1983.
- [37] Andrea Galeotti, Christian Ghiglino, and Francesco Squintani. Strategic information transmission networks. *J. Economic Theory*, 148(5):1751–1769, 2013.
- [38] Luc-Alain Giraldeau and Thomas Caraco. *Social foraging theory*. Princeton University Press, 2018.

- [39] Irving L. Glicksberg. A Further Generalization of the Kakutani Fixed Point Theorem with Application to Nash Equilibrium Points. *Proceedings of the AMS*, 3(1):170–174, 1952.
- [40] Chaitanya S Gokhale and Arne Traulsen. Evolutionary games in the multiverse. *Proceedings of the National Academy of Sciences*, 107(12):5500–5504, 2010.
- [41] Thomas Gregor, Koichi Fujimoto, Noritaka Masaki, and Satoshi Sawai. The onset of collective behavior in social amoebae. *Science*, 328(5981):1021–1025, 2010.
- [42] J. Hagenbach and F. Koessler. Strategic communication networks. *Review of Economic Studies*, 77(3):1072–1099, 2011.
- [43] William D. Hamilton. Narrow roads of geneland. freeman. *Freeman*, 1996.
- [44] Christopher Harris. Existence and Characterization of Perfect Equilibrium in Games of Perfect Information. *Econometrica: Journal of the Econometric Society*, 53(3):613–628, 1985.
- [45] Thomas T Hills, Peter M Todd, David Lazer, A David Redish, Iain D Couzin, Cognitive Search Research Group, et al. Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, 19(1):46–54, 2015.
- [46] Matthew O. Jackson, Brian W. Rogers, and Yves Zenou. Networks: An economic perspective. Technical report, arXiv, 2016.
- [47] Matthew O Jackson and Yves Zenou. Games on networks. In *Handbook of game theory with economic applications*, volume 4, pages 95–163. Elsevier, 2015.
- [48] Martyn Kennedy and Russell D Gray. Can ecological theory predict the distribution of foraging animals? a critical analysis of experiments on the ideal free distribution. *Oikos*, pages 158–166, 1993.
- [49] Martyn Kennedy and Russell D. Gray. Can ecological theory predict the distribution of foraging animals? a critical analysis of experiments on the ideal free distribution. *Oikos*, 68(1):158–166, 1993.
- [50] Jon Kleinberg and Sigal Oren. Mechanisms for (mis) allocating scientific credit. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 529–538. ACM, 2011.
- [51] Amos Korman and Yoav Rodeh. Parallel search without coordination. In *Proceedings of the 24th International Colloquium on Structural Information and Communication Complexity, SIROCCO 2017, Porquerolles, France, June 19-22, 2017*, 2017.
- [52] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413. Springer, 1999.

- [53] Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the “wisdom of the crowd”. *Journal of Political Economy*, 122(5):988–1012, 2014.
- [54] David M Kreps and Robert Wilson. Sequential Equilibria. *Econometrica: Journal of the Econometric Society*, pages 863–894, 1982.
- [55] Fabian Kuhn, Thomas Moscibroda, and Roger Wattenhofer. Radio network clustering from scratch. In *12th European Symposium on Algorithms (ESA)*, LNCS 3221, pages 460–471. Springer, 2004.
- [56] Harold W. Kuhn. Extensive Games and the Problem of Information. In *Contributions to the Theory of Games*, volume II (Annals of Mathematics Studies, 28), pages 193–216. Princeton University Press, 1953.
- [57] Carlton E Lemke and Joseph T Howson, Jr. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12(2):413–423, 1964.
- [58] Nathan Linial. Locality in distributed graph algorithms. *SIAM J. Comput.*, 21(1):193–201, 1992.
- [59] Michael Luby. A Simple Parallel Algorithm for the Maximal Independent Set Problem. *SIAM J. Comput.*, 15(4):1036–1053, 1986.
- [60] Yishay Mansour, Aleksandrs Slivkins, and Vasilis Syrgkanis. Bayesian incentive-compatible bandit exploration. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 565–582. ACM, 2015.
- [61] John M. McNamara and Alasdair I. Houston. Integrating function and mechanism. *Trends in Ecology & Evolution*, 24(12):670 – 675, 2009.
- [62] Linda Mealey. The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences*, 18(3):523–541, 1995.
- [63] Rodrigo A Medellin, Marina Rivero, Ana Ibarra, J Antonio de la Torre, Tania P Gonzalez-Terrazas, Leonora Torres-Knoop, and Marco Tschapka. Follow me: foraging distances of *leptonycteris yerbabuenae* (chiroptera: Phyllostomidae) in sonora determined by fluorescent powder. *Journal of Mammalogy*, 99(2):306–311, 2018.
- [64] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.
- [65] Douglas W Morris. Shadows of predation: habitat-selecting consumers eclipse competition between coexisting prey. *Evolutionary Ecology*, 17(4):393–422, 2003.
- [66] Thomas Moscibroda and Roger Wattenhofer. Maximal independent sets in radio networks. In *24th Symposium on Principles of Distributed Computing (PODC)*, pages 148–157. ACM, 2005.

- [67] Moni Naor and Larry J. Stockmeyer. What can be computed locally? *SIAM J. Comput.*, 24(6):1259–1277, 1995.
- [68] John Nash. Non-cooperative games. ph. d. dissertation in mathematics, princeton university, 1950. 1950.
- [69] John F. Nash. Non-Cooperative Games. *The Annals of Mathematics*, 54(2):286–295, 1951.
- [70] John F Nash et al. Equilibrium Points in n-Person Games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.
- [71] AJ Nicholson. An outline of the dynamics of animal populations. *Australian Journal of Zoology*, pages 9–65, 1954.
- [72] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007.
- [73] Martin J Osborne and Ariel Rubinstein. *A course in game theory*. MIT press, 1994.
- [74] Christos H Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and system Sciences*, 48(3):498–532, 1994.
- [75] Yiangos Papanastasiou, Kostas Bimpikis, and Nicos Savva. Crowdsourcing exploration. *Management Science*, 2017.
- [76] Geoffrey A Parker. Searching for mates. *Behavioural ecology: an evolutionary approach*, 1:214–244, 1978.
- [77] Geoffrey A Parker. Evolutionary stable strategies. *Behavioural ecology: an evolutionary approach*, pages 30–61, 1984.
- [78] David Peleg. *Distributed Computing: A Locality-Sensitive Approach*. Discrete Mathematics and Applications. SIAM, Philadelphia, 2000.
- [79] Noa Pinter-Wollman, Tamar Dayan, David Eilam, and Noga Kronfeld-Schor. Can aggression be the force driving temporal separation between competing common and golden spiny mice? *Journal of Mammalogy*, 87(1):48–53, 2006.
- [80] H Ronald Pulliam and Thomas Caraco. Living in groups: is there an optimal group size. *Behavioural ecology: an evolutionary approach*, 2:122–147, 1984.
- [81] H Ronald Pulliam and Brent J Danielson. Sources, sinks, and habitat selection: a landscape perspective on population dynamics. *The American Naturalist*, 137:S50–S66, 1991.

- [82] Nicanor Quijano and Kevin M Passino. The ideal free distribution: Theory and engineering application. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):154–165, 2007.
- [83] Roy Radner. Collusive Behavior in non-Cooperative epsilon-Equilibria of Oligopolies with Long but Finite Lives. *Journal of Economic Theory*, 22:136–154, 1980.
- [84] Robert W Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.
- [85] Reinhard Selten. Spieltheoretische behandlung eines oligopolmodells mit nachfragetragheit: Teil i: Bestimmung des dynamischen preisgleichgewichts. *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics*, 12(H. 2):301–324, 1965.
- [86] Reinhard Selten. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International journal of game theory*, 4(1):25–55, 1975.
- [87] J Maynard Smith and George R Price. The logic of animal conflict. *Nature*, 246(5427):15, 1973.
- [88] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge university press, 1982.
- [89] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics reports*, 446(4-6):97–216, 2007.
- [90] Bernhard Thomas. On evolutionarily stable sets. *Journal of Mathematical Biology*, 22(1):105–115, 1985.
- [91] Tom Tregenza. Building on the ideal free distribution. *Adv. Ecol. Res*, pages 253–307, 1995.
- [92] Adrian Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 416–425. IEEE, 2002.
- [93] Gandimohan M Viswanathan, Sergey V Buldyrev, Shlomo Havlin, MGE Da Luz, EP Raposo, and H Eugene Stanley. Optimizing the success of random searches. *nature*, 401(6756):911, 1999.
- [94] J Von Neumann and O Morgenstern. Theory of games and economic behavior. 1944.
- [95] Yossi Yovel. The bat lab. <http://www.yossiyovel.com/>, 2019. [Online; accessed 6-March-2019].