RESOURCE ALLOCATION AND DECISION MAKING FOR GROUPS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Warut Suksompong
August 2018

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Tim Roughgarden) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Ariel Procaccia)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Gregory Valiant)

Approved for the Stanford University Committee on Graduate Studies

Abstract

Social choice theory concerns the design and analysis of methods for aggregating possibly conflicting individual preferences to reach a collective decision in a satisfactory manner. The discipline dates from Condorcet's voting paradox in the 18th century, and its paradigm was significantly influenced by Arrow's seminal work in the 1950s. While classical social choice theory focuses on the existence and non-existence of aggregation methods that satisfy certain axioms, over the past two decades computer scientists have studied the discipline from a computational perspective, leading to an active research area of computational social choice. This dissertation presents results of the classical flavor as well as those applying complexity concepts and algorithmic techniques.

In the first part, we consider resource allocation settings. We depart from the usual framework in which each recipient of a bundle of items is represented by a single preference, and assume instead that each interested party consists of agents who may have different preferences on the items. We study the problem of assigning a small subset of indivisible items to a group of agents so that the subset is *agreeable* to all agents, and derive bounds on the size of such a set under various informational and computational assumptions. We also investigate fairness guarantees in the allocation of indivisible items among groups. While the problem is more challenging than in the traditional setting with individual agents, we show that it is nevertheless possible to obtain positive results using asymptotic analysis and approximation.

In the second part, we study decision making problems, where our goal is to choose the "best" alternatives from a given set of alternatives. We demonstrate that the bipartisan set, a method for selecting the best alternatives of a tournament proposed in the 1990s, is the unique method that satisfies a number of desirable properties. In addition, we consider issues related to choosing winners in sports competitions. We show that if tournament organizers have the freedom to determine the bracket of a single-elimination tournament, they can often make their favorite player win the tournament. We also tackle the problem of scheduling *asynchronous* round-robin tournaments, in which all games take place at different times, and propose schedules that perform well with respect to measures concerning the quality and fairness of the tournament.

Acknowledgments

Pursuing a PhD has been a tremendously fun and fulfilling experience. I am truly grateful to the many people who have made this experience possible for me.

First and foremost, I would like to thank my advisor Tim Roughgarden. Tim has given me the freedom to pursue my research interests while at the same time constantly providing me with excellent advice. In addition, his research insights and clarity of thought, his passion for teaching and writing, and his investment in his students will always be an inspiration for me. It has been an enormous privilege to pursue a PhD under his guidance.

Next, I would like to express my thanks to Felix Brandt. Felix was kind enough to host me in Munich every year during my PhD, and working with him has had a profound influence on my research. I am also inspired by his intellectual curiosity, his attention to detail, and his dedication.

I was fortunate to have Charles Leiserson as my master's thesis advisor at MIT. Charles' patience was invaluable in my early days of doing research, and working on my master's thesis under his supervision inspired me to continue my studies in graduate school.

I wish to thank Ariel Procaccia and Gregory Valiant for kindly serving on my reading and orals committee. I would like to extend my gratitude to Yinyu Ye for chairing my orals committee, and Ashish Goel and Fuhito Kojima for serving on the committee. Moreover, I would like to thank Ashish Goel and Virginia Vassilevska Williams for having me as a rotation student during my first year at Stanford.

I am grateful to Ian Kash and Peter Key for hosting me at Microsoft Research, and Xiaohui Bei and Ariel Procaccia for mentoring me during research visits. These opportunities greatly enriched my PhD experience and led to new collaborations.

In addition, I have had the fortune of working with many excellent coauthors over the years, including Xiaohui Bei, Florian Brandl, Felix Brandt, Markus Brill, Sebastian Csar, Tomer Ezra, Michal Feldman, Guangda Huzhang, Ian Kash, Peter Key, Michael Kim, Charles Leiserson, Pasin Manurangsi, Hoon Oh, Ariel Procaccia, Tim Roughgarden, Christian Saile, Tao Schardl, Hans Georg Seedig, Erel Segal-Halevi, Rik Sengupta, and Virginia Vassilevska Williams. I have learned a great deal from all of them.

Thanks to Megan Harris, Ruth Harris, Jam Kiattinant, and Jay Subramanian, who were always ready to help with administrative matters at Stanford. I would also like to acknowledge generous support from a Stanford Graduate Fellowship.

Thanks to all of my friends, near and far, for being there for me. Special thanks to Josh Alman, Sorathan Chaturapruek, Vaggos Chatziafratis, Ofir Geri, Anthony Kim, Michael Kim, Kostas Kollias, Weihao Kong, Panupong Pasupat, Lisa Sauermann, Okke Schrijvers, Vatsal Sharan, Joshua Wang, and Hongyang Zhang for making my time at Stanford so pleasant. In particular, I would like to thank Ofir and Anthony for fun conversations in the office, Vaggos, Weihao, and Vatsal for enjoyable times TAing together, Sorathan and Panupong for listening to my practice talks, Kostas for his encouragement, Okke for his advice and good humor, and Lisa for her kindness and enthusiasm.

Lastly, I wish to express my deepest gratitude to my family for their unconditional love and unwavering support throughout my life. Without them I would not have made it this far.

Contents

Abstract **iv**

Acknowledgments **v**

1 Introduction **1**

1.1 Motivation: Social Choice Theory 1

1.2 Overview of the Dissertation 2

1.2.1 Part I: Resource Allocation 2

1.2.2 Part II: Decision Making 6

1.3 Prerequisites 8

1.4 Bibliographic Notes 9

1.4.1 Excluded Works 9

2 Preliminaries **11**

2.1 Fair Division 11

2.2 Tournaments and Tournament Solutions 13

2.2.1 Tournaments 13

2.2.2 Tournament Solutions 14

2.3 Probabilistic Bounds 15

I Resource Allocation **17**

3 Computing a Small Agreeable Set of Indivisible Items **18**

3.1 Introduction 18

3.2 Preliminaries 21

3.3 Worst-Case Bounds 23

3.3.1 General Worst-Case Bound 23

3.3.2 Matching the Worst-Case Bound 26

3.3.3 Computing Small Necessarily Agreeable Sets . . . 30
3.4 Efficient Approximation . . . 34
3.4.1 General Preferences . . . 34
3.4.2 Additive Utilities . . . 37
3.5 Conclusion and Future Work . . . 42

4 Asymptotic Existence of Fair Divisions for Groups **43**
4.1 Introduction . . . 43
4.2 Preliminaries . . . 45
4.3 Asymptotic Existence and Non-Existence of Fair Divisions . . . 46
4.3.1 Existence . . . 46
4.3.2 Non-Existence . . . 50
4.4 Truthful Mechanism for Approximate Envy-Freeness . . . 52
4.5 Conclusion and Future Work . . . 53

5 Approximate Maximin Shares for Groups of Agents **55**
5.1 Introduction . . . 55
5.2 Preliminaries . . . 56
5.3 Two Groups of Agents . . . 56
5.3.1 Large Number of Agents: No Possible Approximation . . . 57
5.3.2 Approximation via Modified Round-Robin Algorithm . . . 57
5.3.3 Approximation via Maximin Partitions . . . 60
5.3.4 Experimental Results . . . 62
5.4 Several Groups of Agents . . . 62
5.4.1 Positive Results . . . 63
5.4.2 Negative Results . . . 64
5.5 Conclusion and Future Work . . . 65

6 Democratic Fair Allocation of Indivisible Goods **67**
6.1 Introduction . . . 67
6.2 Preliminaries . . . 68
6.3 Two Groups with Binary Valuations . . . 69
6.4 Two Groups with General Valuations . . . 71
6.5 Three or More Groups . . . 73
6.5.1 Approximate Envy-Freeness . . . 73
6.5.2 Approximate Maximin Share . . . 76
6.6 Conclusion and Future Work . . . 77

II Decision Making 78

7 On the Structure of Stable Tournament Solutions 79
7.1 Introduction . . . 79
7.2 Stable Sets and Stable Choice Functions . . . 80
7.3 Generators of Stable Choice Functions . . . 84
7.4 Tournament Solutions . . . 86
7.5 Stable Tournament Solutions and Their Generators . . . 88
7.6 Local Reversal Symmetry . . . 89
7.6.1 Disproving Stability . . . 90
7.6.2 Most Discriminating Stable Tournament Solutions . . . 92
7.7 Conclusion and Future Work . . . 96

8 Who Can Win a Single-Elimination Tournament? 98
8.1 Introduction . . . 98
8.2 Preliminaries . . . 100
8.3 Structural Results . . . 101
8.4 Single-Elimination Winners and Tournament Solutions . . . 104
8.4.1 The Strength of Kings . . . 106
8.5 Generative Models for Tournaments . . . 107
8.5.1 Generalizing the Condorcet Random Model . . . 110
8.6 Conclusion and Future Work . . . 112

9 Scheduling Asynchronous Round-Robin Tournaments 114
9.1 Introduction . . . 114
9.2 Preliminaries . . . 115
9.3 Even Number of Players . . . 118
9.4 Odd Number of Players . . . 120
9.5 Conclusion and Future Work . . . 124

A Omitted Proofs from Chapter 3 126
A.1 Direct Proof of Theorem 3.3.1 for Two Agents . . . 126
A.2 NP-hardness of Balanced 2-Partition . . . 127

B Omitted Proofs from Chapter 4 129
B.1 Proof of Theorem 4.3.1 . . . 129

C Examples for Chapter 7 **134**
C.1 Examples for Remark 7.5.1 134
C.2 Examples for Remark 7.5.3 134
C.3 Examples for Remark 7.6.13 135
C.4 Examples for Remark 7.6.14 136

Bibliography **139**

List of Tables

5.1 Values of the best possible approximation ratio for the maximin share 56
5.2 Experimental results for maximin share approximation 63

6.1 Summary of democratic fairness guarantees. 68

9.1 Summary of the results on asynchronous round-robin tournaments 116

List of Figures

3.1 Graph in the proof of Theorem 3.3.1 25

7.1 Visualization of stability 82
7.2 Logical relationships between choice-theoretic properties 83
7.3 Local reversals of a tournament 90
7.4 Construction in the proof of Theorem 7.6.3 91
7.5 Illustration of the proof of Theorem 7.6.9 95

9.1 The circle design for $n = 10$ 117
9.2 The circle design for $n = 11$ 117
9.3 Two schedules with rest difference index 1 for $n = 6$ 119
9.4 The schedule in Theorem 9.4.4 for $n = 5$ 121
9.5 Two schedules for $n = 7$ 121

C.1 Tournament T_4 135
C.2 Tournament T_7 137

Chapter 1

Introduction

1.1 Motivation: Social Choice Theory

The high-level motivation of most of the work presented in this dissertation is from social choice theory, so it seems appropriate that we begin our exposition there. Given a set of agents with individual and possibly conflicting preferences, how can we combine the preferences to reach a collective decision in a satisfactory manner? This is, of course, a very fundamental question that has been studied for a long time in mathematics, economics, and political science. In the setting usually considered in social choice theory, there is a set of agents, each of whom has a preference over the alternatives in a given set. The type of alternatives can vary according to the application. For example, in voting the alternatives are political candidates, in resource allocation the alternatives are different ways to allocate the resources, and in coalition formation the alternatives are coalition structures. Our goal is to aggregate the preferences of the agents to reach a collective decision, whether that means selecting the winning political candidate, allocating the resources among the agents, or deciding on the coalition structure.

Social choice theory dates from Condorcet's voting paradox in the 18th century, when Marquis de Condorcet made the simple yet at first sight perhaps surprising observation that when each agent specifies a linear preference ranking over the alternatives, the majority relation between the alternatives can contain a cycle [51]. Indeed, assume that there are three alternatives a, b, c, the first agent prefers a to b to c, the second agent prefers b to c to a, and the third agent prefers c to a to b. Then a majority (two out of three) of the agents prefer a to b. Likewise, a majority of the agents prefer b to c, and a majority of the agents prefer c to a. Condorcet's example shows that, somewhat paradoxically, the agents' linear preferences do not always admit a *Condorcet winner*, an alternative that a majority of the agents prefer to any other alternative. In other words, the majority relation between the alternatives does not necessarily suggest a clear winning alternative. While much of the work in social choice theory in the 18th and 19th centuries focused on analyzing specific examples and methods, the paradigm was drastically shifted in the middle of the 20th century. The shift was due to the seminal work of the Nobel laureate Kenneth Arrow in 1950, who established a sweeping statement that applies to all possible aggregation methods, whether previously proposed or not, at once. In particular, Arrow showed that when there are at least three alternatives, any aggregation method that satisfies a short list of intuitively reasonable properties must be a *dictatorship*, meaning that there is a "dictator" whose preference the aggregation rule always directly outputs without taking into account the preferences of the remaining agents [9].

Since Arrow's work, social choice theorists mainly focused on proving the existence and non-existence of aggregation methods satisfying certain properties, often referred to as "axioms". While it is clear that an aggregation rule should not be used if it has a number of undesirable properties, the computational cost of implementing the rule is another crucial consideration. Indeed, a rule that satisfies all of the required axioms and yet takes an exorbitant amount of time to implement would be of limited practical use. This concern was taken up by computer scientists, who over the past two decades have studied the discipline from a computational perspective, leading to an active research area of computational social choice. Work in this area often uses techniques from theoretical computer science, particularly algorithms and complexity theory, to investigate the computational aspects of aggregation problems. This dissertation presents results of the classical flavor as well as those applying complexity concepts and algorithmic techniques. For detailed surveys of social choice theory and computational social choice, we refer to excellent books by Arrow et al. [10, 11] and Brandt et al. [38].

1.2 Overview of the Dissertation

We now present a bird's-eye view of this dissertation. We begin in **Chapter 2 (Preliminaries)** by introducing concepts and definitions that are used across multiple chapters of the dissertation. In particular, we provide the basic setting and notions of fair division and give a brief introduction to tournaments and tournament solutions. We also state a number of probabilistic bounds that we use in our analyses.

The rest of the dissertation is divided into two parts, as we outline below.

1.2.1 Part I: Resource Allocation

In the first part of the dissertation, we focus on resource allocation problems. These problems are extremely commonplace in our daily lives, from allocating school supplies to children and course slots in universities to students, to allocating machine processing time to users and kidneys to kidney transplant patients. Indeed, one of the fundamental problems in economics is how to allocate scarce resources in the best possible way. As the 2012 Nobel laureate in Economics Alvin Roth said in his Nobel lecture, "understanding who gets what, and how and why, is still very much a work in progress" [124].

While there is a vast literature on resource allocation, the literature almost always assumes that each interested party consists of a single agent, or a group of agents represented by a single preference. However, this assumption is too restrictive for many practical situations. For instance, if we divide items among families, it could be that different members of a family have different values for the items—perhaps the father has high value for a television, while the daughter finds it outdated and consequently has little value for it. Another example is a university that needs to divide its resources among competing groups of agents, in this case departments. The agents in each group have different and possibly misaligned interests—the professors who perform theoretical research may prefer more whiteboards and open space in the department building, whereas those who engage in experimental work are more likely to prefer laboratories. These situations cannot be modeled by traditional resource allocation settings in which each recipient of a bundle of items is represented by a single preference. Part I of this dissertation investigates group resource allocation settings, where agents share the same set of items even though they have varying preferences.[1]

Before we investigate a model where we allocate resources among several groups of agents, we first consider in **Chapter 3 (Computing a Small Agreeable Set of Indivisible Items)** a model in which there is only a single group of agents. In particular, all of the agents belong to the same group, and they are collectively allocated the same subset of indivisible items. The items are treated as public goods within the group, meaning that every agent derives full utility from all items in the subset. Our goal is to make this subset "agreeable" to all agents, where one can think of agreeability as a minimal desirability condition: An agent might be able to find other subsets of items that she personally prefers to the allocated set, but the allocated set is still acceptable for her, and she can agree with its allocation to the group. The definition of agreeability that we use is that every agent likes the allocated set of items at least as much as the complement set. Without further constraints, the problem described so far would be trivial. Indeed, since only one group is present, there is nothing stopping us from allocating the whole set of items to the group, which clearly everyone would find agreeable. We therefore impose a constraint on the size of the allocated subset: we want this subset to be as small as possible.

When there is one agent, we may already need to include half of the items in an agreeable set in the case that the agent values all items equally. What happens if there are more agents? At first sight, it might seem that if the agents have very different preferences from one another, the size of the smallest agreeable set would increase by a nontrivial amount. However, and quite surprisingly, this size increases by only about half an item per additional agent, even in the worst case; there is also an example of agents' preferences for which this bound is exactly attained. We therefore have

[1] Despite the applicability of group resource allocation, the only previous works we are aware of that deal with group settings are recent works by Segal-Halevi and Nitzan [133, 134], which study the allocation of *divisible* resources such as cake or land. In contrast, in this dissertation we focus exclusively on the allocation of *indivisible* resources such as houses and cars.

a tight worst-case bound on the minimum size of an agreeable set for any number of agents and items. Nevertheless, in many instances we can find an agreeable set that is much smaller than this worst-case bound. For example, if there is a particular item that all agents prefer to the rest of the items combined, then it suffices to only allocate this item. This yields an agreeable set that is much smaller than the worst-case bound, which is already half of the items even for a single agent. With this observation in mind, we address the question of how well we can approximate the size of the smallest agreeable set for any given instance in time polynomial in the input size. We show that for arbitrary preferences, we cannot do much better in the worst case than the trivial algorithm that always outputs the whole set of items. On the other hand, if we impose an additional structure of the preferences, it becomes possible to obtain a significantly better approximation.

In the remaining chapters of Part I, we consider allocating items to multiple groups of agents. The objective that we are interested in is *fairness*: we want all agents to feel that they receive a fair share of the resources. There is a rich and beautiful theory of fair division that goes back several decades and has been studied in mathematics, economics, and more recently in computer science.[2] A fair division problem that has received a fair amount of attention in popular literature is *cake cutting*. The aim of cake cutting is to fairly divide a cake, which is a metaphor for any divisible resource, between two or more agents. In fact, fair division has several more applications beyond cake cutting. Some of these applications, including sharing apartment rent, splitting taxi fare, and distributing household tasks, are implemented on the website Spliddit (`www.spliddit.org`), which has attracted tens of thousands of users since its launch in 2014 [72].

To reason about fairness, we must define what it means for an allocation of resources to be *fair*. Several notions of fairness have been proposed. One of the oldest and most important notions is *envy-freeness*, which means that each agent values her set of items at least as much as any other agent's set of items. In other words, no agent should envy another agent. Clearly this requirement cannot always be satisfied: if there are one item and two agents who both value the item positively, then no matter what we do, the agent who does not get the item will envy the other agent. So the notion is often relaxed to *envy-freeness up to one good (EF1)*, which means that if an agent envies another agent, then there must be some item in the second agent's bundle such that if we remove it from the bundle, the first agent no longer envies the second agent. EF1 can always be satisfied for any number of agents with arbitrary monotonic valuations [101]. Another fairness notion is called *maximin share fairness*, which means that each agent receives at least her *maximin share*. The maximin share of an agent is defined as the maximum utility that the agent can guarantee by dividing the items into n parts and obtaining the worst part, where n denotes the number of agents. Maximin share fairness can always be satisfied for two agents but not for three or more agents. However, when agents have additive valuations, there always exists an allocation that gives every agent a constant fraction of her maximin share [120].

[2] See the survey by Thomson [155] and the references therein for an introduction to the subject.

Given the fact that an envy-free allocation does not always exist, one positive result that we could still hope to show is that such an allocation exists "most of the time". This is precisely what Dickerson et al. [54] did—they studied the asymptotic existence and non-existence of envy-free allocations in the traditional setting with one agent per group. As an example, assume that agents have additive utilities, and each agent's utility is drawn uniformly and independently at random from the interval $(0, 1)$. If the number of items (m) is less than the number of agents (n), then any allocation leaves some agent with no item, so no allocation can be envy-free. Dickerson et al. showed that even if m is slightly larger than n, an envy-free allocation is still unlikely to exist. However, as soon as m is larger than n by a logarithmic factor, there exists an envy-free allocation with high probability. In **Chapter 4 (Asymptotic Existence of Fair Divisions for Groups)**, we generalize these results to the group setting, assuming that all groups contain the same number of agents. If the number of items is less than the *total* number of agents, the probability that an envy-free allocation exists is bounded away from 1. As we mentioned earlier, this regime is not interesting in the setting with one agent per group since no envy-free allocation can exist. In contrast, in the group setting, envy-free allocations can exist as long as the number of items is at least the number of *groups*. Likewise, we show that when the number of items exceeds the total number of agents by at least a logarithmic factor, an envy-free allocation exists with high probability. Our results are robust in the sense that they hold no matter whether the agents are distributed into several groups of small size or few groups of large size.

Our results in Chapter 4 circumvent the potential non-existence of fair allocations by using asymptotic analysis. Another way to obtain positive results is to use approximation. If we only allocate items between two agents, it is known that a simple "cut-and-choose" algorithm gives both agents their maximin share: We let the first agent partition the items into two bundles that are as equal as possible in her view, and let the second agent choose the bundle that she prefers. One might wonder whether this guarantee still holds if we have two *groups* of agents. That is, with two groups of agents, can we still give every agent her maximin share, or at least a fraction thereof? In **Chapter 5 (Approximate Maximin Shares for Groups of Agents)**, we provide a complete answer to this question. It turns out that the answer depends on the number of agents in the two groups. Specifically, suppose that the two groups consist of n_1 and n_2 agents, where we assume without loss of generality that $n_1 \geq n_2$. We show that if $n_2 = 1$, or if $(n_1, n_2) = (2, 2)$ or $(3, 2)$, then we can give a fraction of the maximin share to every agent. Crucially, this fraction depends only on the number of agents and not on the number of items. In other words, the guarantee does not worsen as we increase the number of items. In all of the remaining cases, there exists an instance in which all agents have positive maximin share, but any allocation leaves some agent with no utility. This implies that there is no hope of obtaining any positive approximation in these cases.

In light of the results in Chapters 4 and 5, which hold either asymptotically or for certain numbers of agents, a natural question is whether there is any hope of obtaining worst-case guarantees that hold for any number of agents. **Chapter 6 (Democratic Fair Allocation of Indivisible Goods)** answers this question in the affirmative. For two groups with any number of agents, we establish the existence of an allocation that is EF1 for at least half of the agents in each group; the bound $1/2$ is also tight. Now, one might be tempted to think that if we relax EF1 to, say, EF2 or EF3, where we define these notions analogously to EF1 but with two or three goods allowed to be removed,[3] then we would be able to satisfy a larger fraction of the agents. This is, however, not true: the bound $1/2$ remains tight for EFc for any constant c. Nevertheless, we would still like to satisfy more than half of the agents; ideally we want to give all agents some nontrivial guarantee. It turns out that this is also possible, provided that we allow the fairness notion to depend on the number of agents. In particular, for two groups of agents with additive valuations, there always exists an allocation that is envy-free up to $n-1$ goods for all agents, where n denotes the total number of agents in the two groups. Like the guarantees in Chapter 5, this guarantee does not depend on the number of items; it does not get worse as we have more items to allocate.

1.2.2 Part II: Decision Making

In the second part of the dissertation, we turn our attention to decision making problems, where our goal is to choose the "best" alternatives from a given set of alternatives. We focus in particular on settings where there is a *dominance relation* between the alternatives, meaning that for any two distinct alternatives, one dominates the other. In graph theory, this structure is known as a *tournament*.[4] A tournament could arise, for example, from a sports competition in which every pair of players play each other once and there is no tie. A common interpretation in social choice theory is that the tournament represents the majority relation of the agents' preferences. For instance, the preferences in Condorcet's voting paradox mentioned in Section 1.1 give rise to a tournament in which a dominates b, b dominates c, and c dominates a. If there are an odd number of agents, the majority relation consists of no ties and therefore constitutes a tournament.

Given a tournament, how should we determine the "best" alternatives, or the "winners" of the tournament? The first method that probably comes to many people's minds, which is also the method often used in sports competitions, is to select the alternatives with the highest out-degree, that is, the alternatives that dominate the highest number of other alternatives. This set of alternatives is known in social choice theory as the *Copeland set*. The Copeland set is an example of a *tournament solution*, i.e., a method for choosing the winners of a given tournament. We provide several insights that improve our understanding of tournament solutions in **Chapter 7 (On the Structure of Stable Tournament Solutions)**. A common way to evaluate tournament solutions is to specify a set of properties that we desire, and verify whether each tournament solution satisfies the properties. Four properties that have been considered in the literature are monotonicity,

[3] See Section 2.1 for formal definitions.

[4] Not to be confused with formats for organizing sports competitions such as single-elimination tournaments or round-robin tournaments, which we also study in Chapters 8 and 9 respectively.

composition consistency, stability, and regularity.[5] Monotonicity means that if an alternative is chosen, it should remain chosen when it dominates one more alternative and everything else is unchanged. Composition consistency means that the tournament solution should choose the best alternatives from the best components. Stability says that a set is chosen from two different sets of alternatives exactly when it is chosen from the union of these sets. Finally, regularity states that if all alternatives have equal out-degree, all of them should be chosen.

One tournament solution that we investigate in detail in Chapter 7 is the *bipartisan set*, proposed in 1993 by Laffond et al. [93] and elegantly defined via the support of the unique mixed maximin strategies of the zero-sum game given by the tournament's skew-adjacency matrix. It is known that the bipartisan set satisfies all of the four properties mentioned in the preceding paragraph, so one could naturally ask whether it is the only tournament solution that satisfies these properties. The answer is negative: the trivial tournament solution, which always selects all alternatives regardless of the tournament, also fulfills the four properties. However, the purpose of tournament solutions is to distinguish the best alternatives from the rest of the alternatives; always selecting all alternatives clearly defeats this purpose. We show that among all tournament solutions that fulfill the four properties, the bipartisan set is the unique most discriminating one in the sense that it selects on average the least number of alternatives, where the average is taken over all labeled tournaments of any fixed size. Our statement applies at once to all possible tournament solutions, whether previously considered or not, very much in the spirit of Arrow's seminal result discussed in Section 1.1.

We end the dissertation with two chapters on issues related to choosing winners in sports competitions. A very popular format for organizing sports competitions is a *single-elimination tournament*, also known as a *knockout tournament*. It is used, for example, in the NCAA championships as well as in most tennis, badminton, and snooker competitions. As much as we love single-elimination tournaments, it is also clear that the winner of these tournaments can depend significantly on the bracket. To put it differently, if the tournament organizers have a favorite player who they want to be the winner of the tournament, they could potentially choose a bracket to help this player win. The *tournament fixing problem*, which we study in **Chapter 8 (Who Can Win a Single-Elimination Tournament?)**, formalizes this intuition. Given a set of players, the underlying tournament relation between players (if a pair of players were to meet, who would win?), and the favorite player of the tournament organizers, the tournament fixing problem asks whether there exists a bracket for a balanced single-elimination tournament that makes this player win the tournament. Such a bracket does not always exist; this is obviously the case if the player is not capable of beating any other player in the tournament. Nevertheless, following the manner in which we circumvent the potential non-existence of envy-free allocations in Chapter 4, it could still be true that a manipulation by the tournament organizers is possible "most of the time".

To establish such a statement, we need to specify the distribution over the underlying tournament

[5] See Sections 2.2 and 7.2 for formal definitions.

relations that we consider. A model that has been previously studied is the *Condorcet Random model*. In this model, there exists a linear order of players from strongest to weakest. Most of the time stronger players win against weaker players, but there is an upset probability p with which a weaker player beats a stronger player. It is known from previous work that if $p \in \Omega\left(\sqrt{\log n/n}\right)$, where n denotes the number of players, then with high probability, every player is a single-elimination winner under some bracket.[6] On the other hand, if $p \in o(\log n/n)$, it is unlikely that the weakest player can win the tournament. The reason is that in this regime, the weakest player beats in expectation less than $\log n$ players, but one needs to beat at least $\log n$ players to have any chance of winning a single-elimination tournament, since the tournament consists of $\log n$ rounds. This leaves a gap between $o(\log n/n)$ and $\Omega\left(\sqrt{\log n/n}\right)$. We close this gap by showing that $\Theta(\log n/n)$ is in fact the point at which the transition occurs. In particular, if $p \in \Omega(\log n/n)$, where $\Omega(\cdot)$ hides a sufficiently large constant, then with high probability, every player can win a single-elimination tournament under some bracket. Furthermore, we prove results on significantly larger classes of distributions over the underlying tournament relations.

Finally, in **Chapter 9 (Scheduling Asynchronous Round-Robin Tournaments)**, we consider another very popular format for organizing sports competitions, namely a *round-robin tournament*. In its simplest form, every pair of players play each other once. We study a variant of round-robin tournaments that we call an *asynchronous round-robin tournament*. In contrast to usual round-robin tournaments, in which games involving disjoint sets of players can take place in parallel, all games in an asynchronous round-robin tournament are played at different times. There are several reasons why one might want to organize such a tournament. For example, this format allows spectators to follow all the games live, and it can be carried out with only one venue. We propose schedules for asynchronous round-robin tournaments that perform well with respect to three measures concerning the fairness and quality of the tournament. The three measures that we strive to optimize are the *guaranteed rest time* (the minimum amount of time that the schedule allows players to take a rest between games), the *games-played difference index* (the maximum difference between the number of games played by any two players at any point in the schedule), and the *rest difference index* (the maximum difference between the rest time of the two players before any game). Interestingly, we show that the schedules that perform well depend on whether the number of players is odd or even.

1.3 Prerequisites

Chapter 3 and small parts of some other chapters require basic knowledge of complexity theory and asymptotic ("big O") notation; we refer to the seminal book by Cormen et al. [52] for an introduction

[6] It is important to note the quantifiers here. We are *not* saying that with high probability, there exists a bracket that makes every player win—this is not possible since for every bracket there can be only one winner. Rather, we are saying that with high probability, for every player, there exists a bracket that makes the player win.

to these topics. Apart from that, most of the dissertation should be accessible to anyone familiar with mathematical proofs. In particular, we do not assume prior knowledge of economics on the part of the reader—all revelant economic concepts are introduced in the dissertation—nor do we assume familiarity with other topics in computer science or mathematics.

1.4 Bibliographic Notes

The material presented in this dissertation is based on joint works with Felix Brandt, Markus Brill, Michael P. Kim, Pasin Manurangsi, Hans Georg Seedig, Erel Segal-Halevi, and Virginia Vassilevska Williams. Specifically, Chapter 3 is based on my own work [148] and joint work with Pasin Manurangsi [106, 107], Chapter 4 is based on joint work with Pasin Manurangsi [105], Chapter 5 is based on my own work [153], Chapter 6 is based on joint work with Erel Segal-Halevi [135], Chapter 7 is based on joint work with Felix Brandt, Markus Brill, and Hans Georg Seedig [35], Chapter 8 is based on joint work with Michael P. Kim and Virginia Vassilevska Williams [84], and Chapter 9 is based on my own work [150].

1.4.1 Excluded Works

In addition to the material presented in this dissertation, I have worked on a number of other topics during my PhD studies.[7] While most of these works are closely related to the theme of this dissertation, they are not part of the dissertation. These works include (but are not limited to):

- Fair division[8]
  - Asymptotic existence of fair allocations [108, 149]
  - Fair division under contiguity constraints [151]
  - Query complexity of fair division [119]
  - Truthfulness in the allocation of divisible resources [18]
- Algorithmic mechanism design
  - Black-box transformations in algorithmic mechanism design [152]
  - Pricing cloud resources [83]
  - Pricing multi-unit markets [60]

[7] During my undergraduate and master's studies, I have also worked on topics in combinatorics [53, 145] and parallel algorithms [99, 146, 154].

[8] Unlike the works on fair division presented in Chapters 4–6 of this dissertation which consider group settings, the following works consider traditional resource allocation settings in which resources are allocated among individual agents.

- Tournament solutions
  - Choosing from large random tournaments [127]
- Voting theory
  - Efficiency and strategyproofness in randomized voting [30]
- Coalition formation
  - Stability in hedonic games [147]
- Noncooperative games
  - Equilibrium concepts in two-player zero-sum games [36]

Chapter 2

Preliminaries

This chapter introduces some definitions, notation, and tools that we will use across multiple chapters. Preliminaries specific to a single chapter are presented in the chapter itself.

2.1 Fair Division

In this section, we introduce the basic definitions and notation for the fair division setting, which we will study in Chapters 4–6.

In the fair division setting for groups, there is a set A of agents. The agents are partitioned into k groups $A_1, A_2, \ldots, A_k$ with $n_1, n_2, \ldots, n_k$ agents, respectively. The jth agent of the ith group is denoted by a_{ij}. There is a set $G = \{g_1, g_2, \ldots, g_m\}$ of goods. A *bundle* is a subset of G. Each agent a_{ij} has a nonnegative *utility* $u_{ij}(G')$ for each bundle $G' \subseteq G$; for a good g, we sometimes write $u_{ij}(g)$ instead of $u_{ij}(\{g\})$. For any agent a_{ij}, denote by $u_{ij,\max} := \max_{l=1,2,\ldots,m} u_{ij}(g_l)$ the maximum utility of the agent for any single good. Denote by $\mathbf{u}_{ij} = (u_{ij}(g_1), u_{ij}(g_2), \ldots, u_{ij}(g_m))$ the utility vector of agent a_{ij} for single goods.

A utility function u is said to be

- *monotonic* if $u(G'') \leq u(G')$ for any bundles $G'' \subseteq G' \subseteq G$;
- *additive* if $u(G') = \sum_{g \in G'} u(g)$ for any bundle $G' \subseteq G$;
- *binary* if it is additive and $u(g) \in \{0, 1\}$ for any good $g \in G$.

Every binary utility function is additive, and every additive utility function is monotonic. We assume that agents always have monotonic utility functions; this is a reasonable assumption since agents can always ignore goods that give them negative value.

We will allocate a bundle $G_i \subseteq G$ to each group A_i, where $(G_1, G_2, \ldots, G_k)$ forms a partition of G, i.e., $\bigcup_{i=1}^{k} G_i = G$ and $G_i \cap G_j = \emptyset$ for any $i \neq j$. The goods are treated as public goods within each group, meaning that for every group i, the utility of each agent a_{ij} is $u_{ij}(G_i)$. We refer to a setting with agents partitioned into groups, goods, and utility functions as an *instance*.

We now define the fairness notions considered in this dissertation. We begin by defining what it means for an allocation to be fair for an *individual* agent. We start with envy-freeness, a classical and important fairness notion [69, 157].

Definition 2.1.1. *Given an agent a_{ij} and an integer $c \geq 0$, an allocation is called* envy-free up to c goods (EFc) for a_{ij} *if for every $i' \neq i$ there is a set $H_{i'} \subseteq G_{i'}$ with $|H_{i'}| \leq c$ such that*

$$u_{ij}(G_i) \geq u_{ij}(G_{i'} \backslash H_{i'}).$$

In other words, one can remove the envy of a_{ij} toward group i' by removing at most c goods from the group's bundle.

An EF0 allocation is also known as envy-free.

Clearly, an envy-free allocation does not always exist even in the traditional setting with one agent per group, for example if there are two agents who both have positive value for a single good. However, an EF1 allocation, and therefore an EFc allocation for any $c \geq 1$, always exists [101].

Next, we define the maximin share concepts [47, 120]. While much more recent than envy-freeness, these concepts have received significant attention in the last few years [7, 16, 71, 73, 92, 120].

Definition 2.1.2. *Given an agent a_{ij} and an integer $c \geq 2$, the* 1-out-of-c maximin share (MMS) *of a_{ij} is defined as the maximum, over all partitions of G into c sets, of the minimum of the agent's utilities for the sets in the partition:*

$$MMS^c_{ij}(G) := \max_{G'_1, G'_2, \ldots, G'_c} \min(u_{ij}(G'_1), u_{ij}(G'_2), \ldots, u_{ij}(G'_c)),$$

where $(G'_1, G'_2, \ldots, G'_c)$ is a partition of G.

The 1-out-of-k MMS of an agent, where k is the number of groups, is simply called her MMS *and denoted by $MMS_{ij}(G)$. Any partition of G for which the maximum is attained above in the case $c = k$ is called a* maximin partition *of the agent.*

An allocation $(G_1, G_2, \ldots, G_k)$ is said to be

- MMS-fair for a_{ij}, *if $u_{ij}(G_i) \geq MMS_{ij}(G)$;*
- 1-out-of-c MMS-fair for a_{ij}, *if $u_{ij}(G_i) \geq MMS^c_{ij}(G)$;*
- q-MMS-fair for a_{ij}, *for some $q \in (0, 1)$, if $u_{ij}(G_i) \geq q \cdot MMS_{ij}(G)$;*
- positive-MMS-fair for a_{ij} *if $MMS_{ij}(G) > 0$ implies $u_{ij}(G_i) > 0$.*

Note that MMS-fairness implies q-MMS-fairness for any $q \in (0, 1)$, which in turn implies positive-MMS-fairness. On the other hand, MMS-fairness is implied by envy-freeness, and the MMS of an agent a_{ij} is at most $u_{ij}(G)/k$.

We are now ready to define how we extend individual fairness notions to group fairness notions.

Definition 2.1.3. *For any $h \in [0, 1]$ and any given fairness notion for individual agents, an allocation $(G_1, G_2, \ldots, G_k)$ is said to be h-*democratic fair *if it is fair for at least $h \cdot n_i$ agents in every group A_i. It is said to be* unanimously fair *or simply* fair *if it is 1-democratic fair.*

The term "democratic fairness" already appeared in the work of Segal-Halevi and Nitzan [133]; however, they used it in the narrower sense that at least half of the agents in each group must be satisfied. In our terminology this is called 1/2-*democratic fairness*. Hence, our democratic fairness notion generalizes existing notions of group fairness.

2.2 Tournaments and Tournament Solutions

In this section, we introduce the basic definitions and notation on tournaments and tournament solutions, which we will be concerned with in Chapters 7 and 8.

2.2.1 Tournaments

A *tournament* T is a pair $(A, \succ)$, where A is a set of alternatives, sometimes called the *feasible set*,[1] and $\succ$ is a connex and asymmetric (and thus irreflexive) binary relation on A, usually referred to as the *dominance relation*.[2] Intuitively, $a \succ b$ signifies that alternative a is preferable to alternative b. The dominance relation can be extended to sets of alternatives by writing $X \succ Y$ when $a \succ b$ for all $a \in X$ and $b \in Y$.

For a tournament $T = (A, \succ)$ and an alternative $a \in A$, we denote by

$$\overline{D}(a) = N_{in}(a) = \{x \in A \mid x \succ a\}$$

the *dominators* of a and by

$$D(a) = N_{out}(a) = \{x \in A \mid a \succ x\}$$

the *dominion* of a. We let $out(a) = |N_{out}(a)|$ and $in(a) = |N_{in}(a)|$. When varying the tournament, we will refer to $\overline{D}_{T'}(a)$, $D_{T'}(a)$, $in_{T'}(a)$, and $out_{T'}(a)$ for some tournament $T' = (A', \succ')$. An alternative a is said to *cover* another alternative b if $D(b) \subseteq D(a)$. It is said to be a *Condorcet winner* if it dominates all other alternatives, and a *Condorcet loser* if it is dominated by all other alternatives. The order of a tournament $T = (A, \succ)$ is denoted by $|T| = |A|$. A tournament is

[1] In Chapter 8, A is the set of players.

[2] This means that $x \not\succ x$ for all $x \in A$, and exactly one of $x \succ y$ and $y \succ x$ holds for all $x, y \in A$ such that $x \neq y$.

regular if the dominator set and the dominion set of each alternative are of the same size, i.e., for all $a \in A$ we have $|D(a)| = |\overline{D}(a)|$. It is easily seen that regular tournaments are always of odd order.

2.2.2 Tournament Solutions

A *tournament solution* is a function that maps a tournament to a nonempty subset of its alternatives. We assume that tournament solutions are invariant under tournament isomorphisms. A tournament solution is *trivial* if it returns all alternatives of every tournament.

We now define some common tournament solutions.[3]

- The *top cycle* (TC) is the (unique) smallest set of alternatives such that all alternatives in the set dominate all alternatives not in the set;
- The *uncovered set* (UC), also known as the set of *kings*, contains all alternatives that are not covered by another alternative;
- The *Banks set* (BA) contains all alternatives that are Condorcet winners of inclusion-maximal transitive subtournaments.

For two tournament solutions S and S', we write $S' \subseteq S$, and say that S' is a *refinement* of S and S a *coarsening* of S', if $S'(T) \subseteq S(T)$ for all tournaments T. The following inclusions are well-known:

$$BA \subseteq UC \subseteq TC.$$

To simplify notation, we will often identify a (sub)tournament by its set of alternatives when the dominance relation is clear from the context. For example, for a tournament solution S and a subset of alternatives $X \subseteq A$ in a tournament $T = (A, \succ)$, we write $S(X)$ for $S(T|_X)$, where $T|_X$ is the tournament induced by restricting T to the set of alternatives X.

We consider the following desirable properties of tournament solutions, all of which are standard in the literature.

First, monotonicity requires a chosen alternative to still be chosen when its dominion is enlarged and everything else is unchanged.

Definition 2.2.1. *A tournament solution is* monotonic *if for all $T = (A, \succ)$, $T' = (A, \succ')$, $a \in A$ such that $\succ_{A\setminus\{a\}} = \succ'_{A\setminus\{a\}}$ and for all $b \in A \setminus \{a\}$, $a \succ' b$ whenever $a \succ b$,*

$$a \in S(T) \quad \textit{implies} \quad a \in S(T').$$

Regularity requires that all alternatives be chosen from regular tournaments.

Definition 2.2.2. *A tournament solution is* regular *if $S(T) = A$ for all regular tournaments $T = (A, \succ)$.*

[3] See, e.g., [34, 80, 98, 115] for more thorough treatments of tournament solutions.

Even though regularity is often considered in the context of tournament solutions, it does not possess the normative appeal of other conditions.

Finally, we consider a structural invariance property that is based on components of similar alternatives and, loosely speaking, requires that a tournament solution choose the "best" alternatives from the "best" components. A *component* is a nonempty subset of alternatives $B \subseteq A$ that bear the same relationship to any alternative not in the set, i.e., for all $a \in A \backslash B$, either $B \succ \{a\}$ or $\{a\} \succ B$. A *decomposition* of T is a partition of A into components.

For a given tournament $\tilde{T}$, a new tournament T can be constructed by replacing each alternative with a component. Let $B_1, \dots, B_k$ be pairwise disjoint sets of alternatives and consider tournaments $T_1 = (B_1, \succ_1), T_2 = (B_2, \succ_2), \dots, T_k = (B_k, \succ_k)$, and $\tilde{T} = (\{1, 2, \dots, k\}, \tilde{\succ})$. The *product* of $T_1, T_2, \dots, T_k$ with respect to $\tilde{T}$, denoted by $\prod(\tilde{T}, T_1, T_2, \dots, T_k)$, is the tournament $T = (A, \succ)$ such that $A = \bigcup_{i=1}^{k} B_i$ and for all $b_1 \in B_i, b_2 \in B_j$,

$$b_1 \succ b_2 \quad \text{if and only if} \quad i = j \text{ and } b_1 \succ_i b_2, \text{ or } i \neq j \text{ and } i \tilde{\succ} j.$$

Here, $\tilde{T}$ is called the *summary* of T with respect to the above decomposition.

Definition 2.2.3. *A tournament solution is said to be* composition-consistent *if for all tournaments $T, T_1, T_2, \dots, T_k$ and $\tilde{T}$ such that $T = \prod(\tilde{T}, T_1, T_2, \dots, T_k)$,*

$$S(T) = \bigcup_{i \in S(\tilde{T})} S(T_i).$$

All three tournament solutions introduced above satisfy monotonicity. *TC* and *UC* are regular, while *UC* and *BA* are composition-consistent.

2.3 Probabilistic Bounds

In this section, we present a number of probabilistic bounds that will be useful for our analyses. Our first bound, commonly known as the Chernoff bound, is a fundamental result that gives an upper bound on the probability that a sum of independent random variables is far away from its expected value [49, 77].

Lemma 2.3.1 (Chernoff bound)**.** *Let $X_1, X_2, \dots, X_r$ be independent random variables that take on values in the interval $[0, 1]$, and let $X := X_1 + \dots + X_r$. We have*

$$\Pr[X \geq (1 + \delta)\,\mathbb{E}[X]] \leq \exp\left(\frac{-\delta^2\,\mathbb{E}[X]}{3}\right),$$

and,

$$\Pr[X \leq (1-\delta)\,\mathbb{E}[X]] \leq \exp\left(\frac{-\delta^2\,\mathbb{E}[X]}{2}\right)$$

for every $\delta \geq 0$*.*

The second result, Lévy's inequality, yields a bound on the maximum of partial sums of independent random variables [100].

Lemma 2.3.2 (Lévy's inequality)**.** *Let* $X_1, X_2, \ldots, X_r$ *be independent random variables, each symmetrically distributed around its median, and let* $Y_i := X_1 + X_2 + \cdots + X_i$ *for* $i = 1, 2, \ldots, r$*. For any real number* x*, we have*

$$\Pr\left[\max_{1\leq i\leq r} |Y_i| \geq x\right] \leq 2\Pr[|Y_r| \geq x].$$

Our third result, the Berry-Esseen theorem, states that a sum of a sufficiently large number of independent random variables behaves similarly to a normal distribution [19, 59]. On the surface, this sounds like the central limit theorem. However, the Berry-Esseen theorem relies on a slightly stronger assumption and delivers a more concrete bound.

Lemma 2.3.3 (Berry-Esseen theorem)**.** *Let* $X_1, X_2, \ldots, X_r$ *be* r *independent and identically distributed random variables, each of which has mean* μ*, variance* σ^2*, and third moment*[4] ρ*. Let* $S := X_1 + X_2 + \cdots + X_r$*. There exists an absolute constant* C_{BE} *such that*

$$\left|\Pr[S \leq x] - \Pr_{y\sim\mathcal{N}(\mu r, \sigma^2 r)}[y \leq x]\right| \leq \frac{\rho C_{BE}}{\sigma^3\sqrt{r}}$$

for every $x \in \mathbb{R}$*, where* $\mathcal{N}(\mu r, \sigma^2 r)$ *is the normal distribution with mean* μr *and variance* $\sigma^2 r$*, i.e., its probability density function is*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi r}} e^{\frac{-(x-\mu r)^2}{2\sigma^2 r}}.$$

[4]The third moment of a random variable X is defined as $\mathbb{E}[|X - \mathbb{E}[X]|^3]$.

Part I

Resource Allocation

Chapter 3

Computing a Small Agreeable Set of Indivisible Items

3.1 Introduction

A typical resource allocation problem involves dividing a set of resources among interested agents. We are often concerned with the *efficiency* of the allocation, e.g., achieving high social welfare or ensuring that no other allocation would make every agent better off than the current allocation. Another important issue, which we will deal with at length in Chapters 4–6, is the *fairness* of the allocation: we might want the resulting allocation to be *envy-free* or *proportional*. A common feature of such problems is that one agent's gain is another agent's loss: The setting inherently puts the agents in conflict with one another, and our task is to try to resolve this conflict as best we can according to our objectives.

In this chapter, we consider a variant of the resource allocation problem where instead of the agents being pitted against one another, they belong to one and the same group. We will collectively allocate a subset of items to this group; our goal is to make this subset "agreeable" to all agents. For instance, the agents could be going together on a trip and agreeing on the set of items to put in a shared luggage, or they could be choosing a subset of items as prizes from a team competition that they won together. Agreeability can be thought of as a minimal desirability condition: While an agent may be able to find other subsets of items that she personally prefers, the current subset is still acceptable for her, and she can agree with its allocation to the group. In other words, if the agreeability condition is not met for some agent, then the agent will be unsatisfied and tempted to leave the group. Without further constraints, the problem described so far would be trivial, since we could simply allocate the entire set of items to the agents. We therefore impose a constraint that the allocated subset should be as small as possible. This constraint on size is reasonable in a variety of settings, including in the two given examples. Indeed, in the first example a luggage has limited space, and in the second example the organizers may want some items to be left as prizes for the losing teams.

We define the notion of agreeability based on the fairness notion of envy-freeness. A subset of items is said to be *agreeable* to an agent if the agent likes it at least as much as the complement set. Agreeability, or minor variants thereof, has been considered in the context of fair division, where each group consists of a single agent [13, 25, 27]. For example, Brams et al. [27] calls the property "worth at least 50 percent". In the example of agents going together on a trip, a subset of items that they take is agreeable if they like it no less than the complement subset of items left at home. Put differently, based on the set of items chosen, every agent would rather go on the trip than stay at home. Similarly, for agents taking items as prizes from a team competition, if the competition is between two teams and a subset of items is not agreeable to some agent in the winning team, we will have an undesirable situation where the agent envies the losing team that takes the remaining items.

This chapter initiates the study of agreeability in resource allocation. First, in Section 3.3, we establish upper bounds on the size of the smallest agreeable set, both when the algorithm has access to the agents' full preferences and when the algorithm only has access to the agents' preferences on single items. In addition, we present algorithms that compute agreeable sets whose size matches the worst-case bounds under both assumptions.

In Section 3.3.1, we derive a tight upper bound on the number of items that may need to be included in an agreeable set, for any number of agents and items. Remarkably, even though agents may have vastly differing and perhaps conflicting preferences, the number of extra items that we might need to choose in order to accommodate all of them is surprisingly small, i.e., half an item per additional agent (Theorem 3.3.1). Our result holds under a very weak assumption that preferences are *monotonic*, meaning that an agent cannot be worse off whenever an item is added to her set. Interestingly, to establish this result we make use of Kneser's conjecture, a combinatorial result whose proof by Lovász [103] gave rise to the field of topological combinatorics.

In Section 3.3.2, we turn our attention to the question of whether we can efficiently compute an agreeable set whose size matches the worst-case bound given in Section 3.3.1. We answer the question in the affirmative for the cases of two and three agents. To this end, we make the assumption that preferences are *responsive*, meaning that an agent cannot be worse off when an item is added to her set or replaced by another item that she weakly prefers to the original item. While responsiveness is stronger than monotonicity, it is still a generalization of additivity, a very common assumption on preferences in resource allocation problems. We present polynomial-time algorithms that compute an agreeable subset whose size matches the upper bound when there are two or three agents (Theorems 3.3.3 and 3.3.5).

In Section 3.3.3, we assume that the algorithm only has access to the agents' ordinal preferences on single items rather than subsets of items. Models of this type offer the advantage that the associated algorithms are often simple to implement and the agents do not need to give away or even determine their entire preferences; such models have therefore received widespread attention [13, 25, 89]. With only the ordinal preferences on single items at its disposal, however, in most cases the algorithm cannot tell whether a certain subset is agreeable to an agent or not. Nevertheless, by assuming that preferences are responsive, we can extend preferences on single items to partial preferences on subsets. This allows us to deduce that certain subsets are always agreeable as long as the full responsive preferences are consistent with the rankings over single items; we call such subsets *necessarily agreeable*. Denoting by m the number of items, we show using results from discrepancy theory that for any constant number of agents, there exists a necessarily agreeable subset of size $m/2 + O(\log m)$, and such a subset can be found in polynomial time (Theorem 3.3.6). Furthermore, we establish the tightness of this bound by showing that even with three agents, there exist preferences for which every necessarily agreeable subset has size $m/2 + \Omega(\log m)$ (Theorem 3.3.9).

Next, in Section 3.4, we investigate the problem of computing an agreeable subset of approximately optimal size for any given instance, as opposed to one whose size matches the worst-case bound over all instances with the same number of agents and items. We tackle the problem using two models for representing preferences that are well-studied in the literature, and exhibit computationally efficient algorithms for finding an agreeable set of approximately optimal size in each of them. Moreover, in both of the models we show that our approximation ratios are asymptotically tight.

In Section 3.4.1, we consider general preferences using the value oracle model, where the preferences of the agents are represented by utility functions and the algorithm is allowed to query the utility of any agent for any subset of items. We exhibit an efficient approximation algorithm with approximation ratio $O(m/\log m)$ in this model (Theorem 3.4.1). While this may not seem impressive, especially in light of the observation that the trivial algorithm which always outputs the entire set of items already achieves approximation ratio $O(m)$, we show that our approximation ratio is in fact the best we can hope for. In other words, there does not exist a polynomial-time algorithm with approximation ratio $o(m/\log m)$, even when there is only a single agent (Theorem 3.4.2).

In Section 3.4.2, we assume that the agents' preferences are represented by additive utility functions. Additivity provides a reasonable tradeoff between simplicity and expressiveness; it is commonly assumed in the literature, especially in recent work [7, 26, 48, 105, 153]. We show that under additive valuations, it is NP-hard to decide whether there exists an agreeable set containing exactly half of the items, even where there are only two agents (Theorem 3.4.3). On the other hand, using results on covering integer programs, we demonstrate the existence of an $O(\log n)$-approximation algorithm for computing an agreeable set of minimum size (Theorem 3.4.8). Moreover, we show that this approximation factor is tight: For any constant $\delta > 0$, it is NP-hard to approximate the problem to within a factor of $(1-\delta)\ln n$ (Theorem 3.4.7).

3.2 Preliminaries

We consider n agents, numbered $1, 2, \ldots, n$, who will be collectively allocated a subset of the set $S = \{x_1, x_2, \ldots, x_m\}$ of m indivisible items. Denote by $\mathcal{S}$ the set of all subsets of S. Each agent i is endowed with a preference relation $\succeq_i$, a reflexive, complete, and transitive ordering over $\mathcal{S}$. Let $\succ_i$ denote the strict part and $\sim_i$ the indifference part of the relation $\succeq_i$. For items x and y, we will sometimes abuse notation and write $x \succeq y$ to mean $\{x\} \succeq \{y\}$. An agent i is said to *strongly prefer* a set T_1 to T_2 if $T_1 \succ_i T_2$, and *weakly prefer* T_1 to T_2 if $T_1 \succeq_i T_2$.

We assume in this chapter that preferences are monotonic, i.e., an agent cannot be worse off when an item is added to her set. Monotonicity is a natural assumption in a wide range of situations. In particular, it implies free disposal of items—every item is considered to be of nonnegative value to all agents.

Definition 3.2.1. *A preference $\succeq$ on $\mathcal{S}$ is* monotonic *if $T \cup \{x\} \succeq T$ for all $T \subseteq S$.*

Note that if $x \in T$, then $T \cup \{x\} \succeq T$ always holds, so we only need to check the condition when $x \in S \backslash T$.

We are now ready to define the central notion of this chapter.

Definition 3.2.2. *A subset $T \subseteq S$ is said to be* agreeable *to agent i if $T \succeq_i S \backslash T$.*

When the set of agents considered is clear from the context, we will sometimes refer to a set that is agreeable to all agents simply as an agreeable set. Since preferences are monotonic, the whole set S is agreeable to every agent, so an agreeable set always exists for any number of agents.[1] Agreeability to an agent also implies that the agent does not strictly prefer any subset of the complement to the current set. That is, we have $T \succeq_i U$ for any $U \subseteq S \backslash T$.

Another property of preferences that we will consider is responsiveness, which says that an agent cannot be worse off whenever an item is added to her set or replaced by another item that she weakly prefers to the original item. While stronger than monotonicity, responsiveness is still a reasonable assumption in many settings.[2]

Definition 3.2.3. *A preference $\succeq$ on $\mathcal{S}$ is* responsive *if it satisfies the following two conditions:*

- *$\succeq$ is monotonic;*
- *$(T \backslash \{y\}) \cup \{x\} \succeq T$ for all $T \subseteq S$ and $x, y \in S$ such that $x \succeq y$, $x \notin T$ and $y \in T$.*

[1] If preferences are not monotonic, an agreeable set might not exist, e.g., if there are two agents with strict preferences, and one agent's preference is exactly the opposite of the other agent's preference.

[2] For a comprehensive treatment of properties concerning the ranking of sets of objects, we refer to a survey by Barberà et al. [15].

If we have access to the complete preference of an agent, we can check whether a subset is agreeable to the agent simply by comparing it to its complement. When we only have access to the agent's preference on single items, however, most of the time we cannot tell whether a given subset is agreeable or not. Nevertheless, if we assume that the agent's preference is responsive, we can sometimes deduce that a certain subset is agreeable only by looking at the agent's preference on single items. The following definition captures this intuition. In general, we use $\succeq$ to denote a preference on $\mathcal{S}$ and $\succeq^{sing}$ to denote a preference on the single items in S.

Definition 3.2.4. *Fix a preference $\succeq^{sing}$ on the single items in S. A subset $T \subseteq S$ is said to be* necessarily agreeable *with respect to $\succeq^{sing}$ if $T \succeq S \backslash T$ for any responsive preference $\succeq$ on $\mathcal{S}$ consistent with $\succeq^{sing}$.*

For the sake of brevity, we say that a subset of items is necessarily agreeable to an agent if it is necessarily agreeable with respect to the preference on single items of the agent.

We now make a connection to the model where every agent has a cardinal utility for each subset of items. A *utility function* u is a function that maps any subset of items to a nonnegative real number. Since each agent's preference is reflexive, complete, and transitive, there is a utility function $u_i : \mathcal{S} \to \mathbb{R}_{\geq 0}$ for agent i such that for any $T_1, T_2 \subseteq S$, we have $T_1 \succeq_i T_2$ if and only if $u_i(T_1) \geq u_i(T_2)$. Moreover, since we consider monotonic preferences, we have $u_i(T_1) \leq u_i(T_2)$ for any $T_1 \subseteq T_2$. We assume that $u_i(\emptyset) = 0$ for all i. A utility function u is said to be *additive* if $u(T_1 \cup T_2) = u(T_1) + u(T_2)$ for any disjoint subsets T_1, T_2, and *subadditive* if $u(T_1 \cup T_2) \leq u(T_1) + u(T_2)$ for any T_1, T_2. Any monotonic additive function is also subadditive. Subadditive utility functions have been extensively studied in the literature [20, 62].

When the preferences of the agents are given by subadditive utility functions, a subset that an agent regards as agreeable also gives the agent a utility of at least half of the agent's utility for the whole set S. Indeed, for any agreeable subset T we have

$$f(S) = f(T \cup (S \backslash T)) \leq f(T) + f(S \backslash T) \leq 2f(T),$$

which implies that $f(T) \geq f(S)/2$. Hence an agreeable subset also gives a 2-approximation of the welfare to an agent when the agent's utility function is subadditive.

We end this section by giving a characterization of necessarily agreeable subsets, which will be used multiple times in the chapter. Similar statements have been shown by Aziz et al. [13] and Brams et al. [27], although our treatment differs slightly in dealing with ties.

Proposition 3.2.5. *Fix a preference $\succeq^{sing}$ on the single items in S with*

$$x_1 \succeq^{sing} x_2 \succeq^{sing} \cdots \succeq^{sing} x_m.$$

Let $T \subseteq S$, and define $I_k = \{x_1, x_2, \ldots, x_k\}$ for all $k = 1, 2, \ldots, m$.

If $|I_k \cap T| \geq k/2$ for all $k = 1, 2, \ldots, m$, then T is necessarily agreeable with respect to $\succeq^{sing}$. The converse also holds if the preference $\succeq^{sing}$ is strict.

Proof. Assume first that $|I_k \cap T| \geq k/2$ for all $k = 1, 2, \ldots, m$. Since $|I_m \cap T| \geq m/2$, we have that $|T| \geq |S \backslash T|$. Let $T' \subseteq T$ be the subset consisting of the $|S \backslash T|$ items of T with the smallest indices.

Define a bijective function $f : T' \to S \backslash T$ as follows: Given the item $x_k \in T$ with the smallest index for which $f(x_k)$ is not yet defined, we define $f(x_k)$ to be the item in $S \backslash T$ with the smallest index that has not occurred in the range of f so far. Since $|I_k \cap T| \geq k/2$ for all $k = 1, 2, \ldots, m$, the function f maps each item x_k to another item x_l with $l > k$. The definition of responsiveness implies that for any responsive preference $\succeq$ on $\mathcal{S}$ consistent with $\succeq^{sing}$, it holds that $T' \succeq S \backslash T$. Since any responsive preference is also monotonic, we have $T \succeq S \backslash T$, which implies that T is necessarily agreeable with respect to $\succeq^{sing}$.

For the converse, assume that the preference $\succeq^{sing}$ is strict, and that $|I_l \cap T| < l/2$ for some $l = 1, 2, \ldots, m$. Let $\epsilon > 0$ be a small constant, and suppose that the preference $\succeq$ is given by an additive utility function u such that:

- $u(x_i) = 1 + (l - i)\epsilon$ for $1 \leq i \leq l$;
- $u(x_i) = (m - i)\epsilon$ for $l < i \leq m$.

Since any preference that can be represented by an additive utility function is responsive, $\succeq$ is responsive. Moreover, we have $u(S \backslash T) > l/2$, whereas $u(T) < l/2$ when ϵ is small enough. It follows that $\succeq$ is a responsive preference on $\mathcal{S}$ consistent with $\succeq^{sing}$ such that $S \backslash T \succ T$. Hence T is not necessarily agreeable with respect to $\succeq^{sing}$. □

Finally, any logarithm written without a base in this chapter is assumed to have base 2.

3.3 Worst-Case Bounds

In this section, we establish upper bounds on the size of the smallest agreeable set, both when the algorithm has access to the agents' full preferences and when the algorithm only has access to the agents' preferences on single items. In addition, we present algorithms that compute agreeable sets whose size matches the worst-case bounds under both assumptions.

3.3.1 General Worst-Case Bound

We commence our study of agreeable sets by deriving a tight worst-case bound on the number of items that may need to be included in such a set, for any number of items and any number of agents with arbitrary preferences on the items. Even with a single agent, there already exists a preference for which we need to include at least half of the items, e.g., a preference represented by an additive utility function that gives the same positive utility to every item. In light of this, it may seem that there is little hope of obtaining a small agreeable set when there are several agents, possibly with wildly varying preferences. Nevertheless, we show that the number of extra items that we need to include to accommodate the additional agents is surprisingly small even in the worst case—this number is only half an item per additional agent.

Theorem 3.3.1. *For any number of agents and items, there exists a subset $T \subseteq S$ such that*

$$|T| \leq \min\left(\left\lfloor \frac{m+n}{2} \right\rfloor, m\right)$$

and T is agreeable to all agents. Moreover, there exist preferences for which this bound is tight.

Theorem 3.3.1 can be seen as a discrete version of consensus halving, where the goal is to partition a divisible item such as cake or land into two parts that all agents think are worth exactly the same. Interestingly, a consensus halving partition can be found for any number of agents [3, 138]. It follows that we can find a part of the item that is at most half of the item but that all agents think is worth at least half of the item (in particular, we choose the smaller of the two parts in the consensus halving partition). When items are indivisible, however, it may no longer be possible to choose a set containing at most half of the items such that all agents believe this set is worth at least as much as its complement. Indeed, if there is only one item and all agents value this item positively, the item needs to be included in the set. Theorem 3.3.1 gives us a precise bound on how many additional items need to be included in the worst case.

In the case of two agents, we give a direct proof of Theorem 3.3.1 in Appendix A.1. For the general case, however, our proof of the theorem will rely on the following combinatorial result, which is best known as Kneser's conjecture.

Lemma 3.3.2 (Kneser's conjecture)**.** *Let G be the undirected graph with all k-element subsets of the set $\{1, 2, \ldots, n\}$ as vertices such that there exists an edge between two vertices if and only if the corresponding sets are disjoint. The chromatic number*[3] *of G is given by*

$$\chi(G) = \begin{cases} n - 2k + 2 & \text{if } n \geq 2k; \\ 1 & \text{otherwise.} \end{cases}$$

The statement of the lemma is due to Kneser [86], who proposed it as a conjecture in the problem column of a German mathematics magazine in 1955. In spite of the simple statement, it was not until 1978 that the conjecture was first resolved by Lovász [103] using topological methods. The proof was later simplified by Bárány [14] and Greene [74], before Matoušek [110] gave the first purely combinatorial proof in 2004. Lovász's proof of the conjecture, which makes use of the Borsuk-Ulam theorem, marked the first time that methods from algebraic topology were used to solve a

[3]The chromatic number of a graph is defined as the smallest number of colors needed to color the vertices of the graph so that no two adjacent vertices share the same color.

combinatorial problem, and gave rise to the field of topological combinatorics.

With Lemma 3.3.2 in hand, we are ready to establish our theorem.

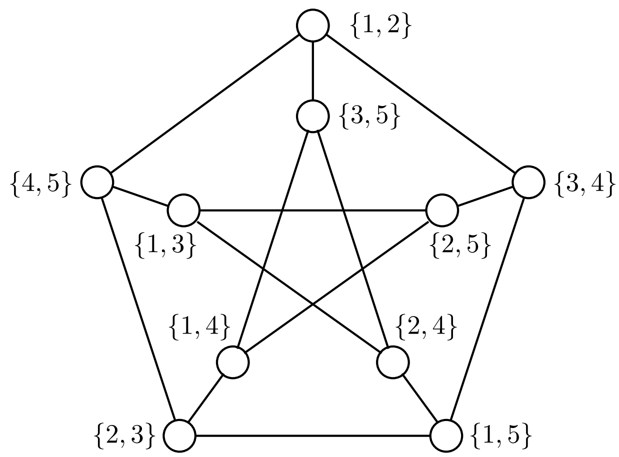

Figure 3.1: The graph G in the proof of Theorem 3.3.1 when $n = 2$ and $m = 5$, also known as the Petersen graph. A vertex with a label $\{i, j\}$ corresponds to the set $\{x_i, x_j\}$.

Proof of Theorem 3.3.1. Let $k = \lfloor \frac{m+n}{2} \rfloor$. If $k \geq m$, the set S of all items has size $m = \min(k, m)$ and is agreeable to all agents since preferences are monotonic. Assume from now on that $k < m$, and consider the undirected graph G with all $(m-k)$-element subsets of $\{x_1, x_2, \ldots, x_m\}$ as vertices and with edges connecting vertices whose corresponding sets are disjoint. For example, if $n = 2$ and $m = 5$ (so $k = 3$ and $m - k = 2$), the graph G corresponds to the well-known Petersen graph and is shown in Figure 3.1.

If all agents weakly prefer $S \backslash T$ to T for some $(m - k)$-element subset $T \subseteq S$, then $S \backslash T$ is our desired agreeable subset of size k. Suppose for contradiction that for any $(m - k)$-element subset $T \subseteq S$, there exists an agent who strictly prefers T to $S \backslash T$. We color the vertices of G with n colors in the following way: For each vertex v of G corresponding to a set T, color it with the color corresponding to an agent who strictly prefers T to $S \backslash T$. If there is more than one such agent, choose one arbitrarily.

Since $k = \lfloor \frac{m+n}{2} \rfloor \geq \frac{m}{2}$, we have $m \geq 2(m - k)$. By Lemma 3.3.2, the chromatic number of G is

$$m - 2(m - k) + 2 = 2k - m + 2 \geq 2\left(\frac{m + n - 1}{2}\right) - m + 2 = n + 1.$$

Since we colored G with n colors, there exist two adjacent vertices sharing the same color. Let T_1 and T_2 be the sets corresponding to these vertices. This means that $T_1 \succ_i S \backslash T_1$ and $T_2 \succ_i S \backslash T_2$ for some agent i. Since T_1 and T_2 are disjoint, we have $T_1 \subseteq S \backslash T_2$ and $T_2 \subseteq S \backslash T_1$. Monotonicity now implies that

$$S\backslash T_1 \succeq_i T_2 \succ_i S\backslash T_2 \succeq_i T_1 \succ_i S\backslash T_1,$$

a contradiction. This means that we can always find a subset of size k that is agreeable to all agents.

Finally, we show that there exist monotonic preferences for which the bound $\min(k, m)$ is tight. In fact, we can even choose preferences that are represented by additive utility functions. We consider two cases.

- $n \geq m$. Then $\min(k, m) = m$. For $i = 1, 2, \ldots, n$, let the preference of agent i be given by an additive utility function u such that $u(x_{\min(i,m)}) = 1$ and $u(x_j) = 0$ for all $j \neq \min(i, m)$. Then any subset that is agreeable to agent i must contain item $x_{\min(i,m)}$. Hence a subset that is agreeable to all agents must contain all m items.

- $n < m$. Then $\min(k, m) = k$. For $i = 1, 2, \ldots, n-1$, let the preference of agent i be given by an additive utility function u such that $u(x_i) = 1$ and $u(x_j) = 0$ for all $j \neq i$. Let the preference of agent n be given by an additive utility function u such that $u(x_j) = 1$ for $j \in \{n, n+1, \ldots, m\}$ and $u(x_j) = 0$ otherwise.

  For $i = 1, 2, \ldots, n-1$, any subset that is agreeable to agent i must contain item x_i. Also, any subset that is agreeable to agent n must contain at least half of the items $x_n, x_{n+1}, \ldots, x_m$. Hence a subset that is agreeable to all agents must have size at least $n - 1 + \left\lceil \frac{m-n+1}{2} \right\rceil = \left\lceil \frac{m+n-1}{2} \right\rceil = \left\lfloor \frac{m+n}{2} \right\rfloor = k$, as desired.

This completes the proof. □

3.3.2 Matching the Worst-Case Bound

Theorem 3.3.1 gives us a tight worst-case bound on the size of the smallest agreeable set for any number of agents and items. However, its proof does not yield a method for obtaining a set of that size. Since the number of sets that we have to consider is exponential in the number of items, brute-force search is infeasible even for moderate numbers of items. Our goal in this section is to show that when there are two or three agents with responsive preferences, it is in fact possible to compute an agreeable set whose size matches the worst-case bound in polynomial time. This implies that we can compute such a set even when the number of items is large.

An important issue when we discuss algorithms is how we represent the agents' preferences. Since preferences on subsets, unlike preferences on single items, might not have a succinct representation, it is not possible to design algorithms that run in time polynomial in the number of items if the algorithm is required to read the entire preference. To circumvent this problem, we assume in this section that preferences are responsive; this allows us to extend preferences on single items to partial preferences on subsets. Our algorithm for two agents will only make use of the preferences on single items and compute a necessarily agreeable subset.[4] On the other hand, our algorithm for three agents will also query the agents' preferences on subsets through a preference oracle in addition to utilizing the preferences on single items.

We first handle the case of two agents.

Theorem 3.3.3. *Assume that there are two agents with preferences $\succeq_1^{sing}$ and $\succeq_2^{sing}$ on the single items in S. There exists a subset $T \subseteq S$ such that $|T| \leq \left\lfloor \frac{m+2}{2} \right\rfloor$ and T is necessarily agreeable with respect to both $\succeq_1^{sing}$ and $\succeq_2^{sing}$. Also, there exists a polynomial-time algorithm that computes such a subset T.*

Moreover, there exist preferences on the single items in S for which the bound $\left\lfloor \frac{m+2}{2} \right\rfloor$ is tight.

Proof. Assume first that $m = 2k+1$ is odd, and suppose without loss of generality that $x_1 \succeq_1^{sing} x_2 \succeq_1^{sing} \cdots \succeq_1^{sing} x_{2k+1}$. We choose our set T of $\left\lfloor \frac{m+2}{2} \right\rfloor = k+1$ items as follows:

1. Choose x_1.
2. Between each of the k pairs of items $(x_2, x_3), (x_4, x_5), \ldots, (x_{2k}, x_{2k+1})$, choose the item that is preferred according to $\succeq_2^{sing}$. If $\succeq_2^{sing}$ is indifferent between any pair of items, choose an arbitrary item from that pair.

For any $j = 1, 2, \ldots, m$, our set T contains at least $j/2$ of the j items $x_1, x_2, \ldots, x_j$; by Proposition 3.2.5, T is necessarily agreeable with respect to $\succeq_1^{sing}$. Moreover, since we choose the item that is preferred according to $\succeq_2^{sing}$ from each of the sets $\{x_2, x_3\}$, $\{x_4, x_5\}, \ldots, \{x_{2k}, x_{2k+1}\}$ along with x_1, Proposition 3.2.5 implies that T is also necessarily agreeable with respect to $\succeq_2^{sing}$. Hence T is necessarily agreeable with respect to both $\succeq_1^{sing}$ and $\succeq_2^{sing}$.

Assume now that $m = 2k$ is even. Let $S' = S \backslash \{x_1\}$. We apply the algorithm from the case of m odd to choose a set $T \subseteq S'$ of size k that is necessarily agreeable with respect to both $\succeq_1^{sing}$ and $\succeq_2^{sing}$ when the universe considered is S'. It follows that $T \cup \{x_1\}$ is a subset of size $\left\lfloor \frac{m+2}{2} \right\rfloor = k+1$ that is necessarily agreeable with respect to both $\succeq_1^{sing}$ and $\succeq_2^{sing}$ when the universe considered is S.

Next, we show that there exist preferences on single items for which the bound $\left\lfloor \frac{m+2}{2} \right\rfloor$ is tight. If $m = 2k+1$ is odd and the preference $\succeq_1^{sing}$ is strict, then by Proposition 3.2.5, any subset that is necessarily agreeable with respect to $\succeq_1^{sing}$ alone must already contain at least $\left\lfloor \frac{m+2}{2} \right\rfloor = k+1$ items.

Finally, suppose that $m = 2k$ is even, and let $\succeq_1^{sing}$ and $\succeq_2^{sing}$ be such that $x_1 \succ_1^{sing} x_2 \succ_1^{sing} \cdots \succ_1^{sing} x_{2k}$ and $x_{2k} \succ_2^{sing} x_{2k-1} \succ_2^{sing} \cdots \succ_2^{sing} x_1$. By Proposition 3.2.5, any subset $T \subseteq S$ that is necessarily agreeable with respect to $\succeq_1^{sing}$ alone must contain at least k items, one of which is x_1. If T contains exactly k items, then it contains exactly $k-1$ items among $x_2, x_3, \ldots, x_{2k}$.

[4] If we do not assume responsiveness, there still exists a polynomial-time algorithm for two agents that discovers the agents' preferences on subsets through a preference oracle; this algorithm is described in Appendix A.1.

Proposition 3.2.5 implies that such a set T is not necessarily agreeable with respect to $\succeq_2^{sing}$. Hence any subset $T \subseteq S$ that is necessarily agreeable with respect to both $\succeq_1^{sing}$ and $\succeq_2^{sing}$ must contain at least $\left\lfloor \frac{m+2}{2} \right\rfloor = k+1$ items, as desired. □

At a high level, the algorithm in Theorem 3.3.3 bears a resemblance to the "Trump rule", which was proposed by Pruhs and Woeginger [121] for fair division of indivisible items between two agents. Like our algorithm, the Trump rule takes as input the preferences on single items of the two agents. Using our terminology, the rule is guaranteed to produce an allocation with the property that each agent views her bundle as necessarily agreeable, whenever such an allocation exists. The difference between the Trump rule and our algorithm is that the Trump rule produces a partition of the items into two subsets with each agent taking one subset, whereas our algorithm produces a single subset that both agents share.

Observe that in the case of two agents, the upper bound for the size of the smallest necessarily agreeable set (Theorem 3.3.3) coincides with the bound for the size of the smallest agreeable set (Theorem 3.3.1). This is somewhat surprising because the definition of a necessarily agreeable set only involves preferences on single items, and yet the worst-case bound remains unchanged even if we have access to the full preferences. The following example shows that the same statement ceases to hold when there are three agents.

Example 3.3.4. *Let $m = 6$, and assume that the preferences on single items of the three agents are as follows:*

1. $x_1 \succ_1^{sing} x_4 \succ_1^{sing} x_5 \succ_1^{sing} x_6 \succ_1^{sing} x_2 \succ_1^{sing} x_3$;
2. $x_2 \succ_2^{sing} x_5 \succ_2^{sing} x_6 \succ_2^{sing} x_4 \succ_2^{sing} x_3 \succ_2^{sing} x_1$;
3. $x_3 \succ_3^{sing} x_6 \succ_3^{sing} x_4 \succ_3^{sing} x_5 \succ_3^{sing} x_1 \succ_3^{sing} x_2$.

In Example 3.3.4, any subset that is necessarily agreeable to all three agents must contain x_1, x_2, x_3, since each of them is ranked first by some agent. Moreover, choosing only one of x_4, x_5, x_6 does not yield a necessarily agreeable set for the agent who ranks that item fourth. Hence a necessarily agreeable set must contain at least five items. On the other hand, if we have access to the agents' full preferences, Theorem 3.3.1 implies that we can find a set of size $\left\lfloor \frac{6+3}{2} \right\rfloor = 4$ that is agreeable to all agents.

Therefore, to compute an agreeable set whose size matches the worst-case bound when there are three agents, it is not sufficient to consider only preferences on single items. Nevertheless, if the algorithm has access to the agents' full preferences, it is possible to find such a subset in polynomial time. To access the preferences, the algorithm is allowed to make a polynomial number of queries to a *preference oracle*. In each query, the algorithm can specify an agent and two subsets of items to the preference oracle, and the oracle reveals the preference of that agent between the two subsets.

Theorem 3.3.5. *Assume that there are three agents with responsive preferences $\succeq_1$, $\succeq_2$, and $\succeq_3$ on S. There exists a polynomial-time algorithm that computes a subset $T \subseteq S$ such that $|T| \leq \left\lfloor \frac{m+3}{2} \right\rfloor$ and T is agreeable to all three agents.*

Proof. Assume first that $m = 2k$ is even. Our goal is to find a subset of size $\left\lfloor \frac{m+3}{2} \right\rfloor = k+1$ that is agreeable to all three agents. Suppose without loss of generality that x_{2k-1} is the most preferred item according to $\succeq_1$, x_{2k} is the most preferred item other than x_{2k-1} according to $\succeq_2$, and among the remaining $2k-2$ items, the preference $\succeq_1$ ranks them as $x_1 \succeq_1 x_2 \succeq_1 \cdots \succeq_1 x_{2k-2}$.

Let $A = \{x_1, x_2, \ldots, x_{2k-2}\}$, and consider the pairs $(x_1, x_2), (x_3, x_4), \ldots, (x_{2k-3}, x_{2k-2})$. Let B be a set of $k-1$ items containing an item from each pair that is *not* preferred to the other item in the pair according to $\succeq_2$. If $\succeq_2$ is indifferent between any pair of items, we choose arbitrarily. Responsiveness implies that $A \backslash B \succeq_2 B$.

As long as $A \backslash B \succeq_2 B$, we remove an element from B that was also originally in B, and insert the other item in its pair into B. We must eventually reach a point where $B \succeq_2 A \backslash B$, at the latest after $k-1$ moves. We consider two cases.

- We have not performed any move. By definition of B, we have that $B \succeq_2 A \backslash B$ and $A \backslash B \succeq_2 B$, and therefore $A \backslash B \sim_2 B$. Since preferences are monotonic, it follows that $(A \backslash B) \cup \{x_{2k}\} \succeq_2 B$ and $B \cup \{x_{2k}\} \succeq_2 A \backslash B$.

- We have performed at least one move. Suppose without loss of generality that in our last move, we inserted x_{2i-1} into and removed x_{2i} from B. Let $C = (A \backslash (B \cup \{x_{2i}\})) \cup \{x_{2i-1}\}$ and $D = (B \backslash \{x_{2i-1}\}) \cup \{x_{2i}\}$, i.e., C and D are the sets $A \backslash B$ and B before the last move, respectively. We have that $C \succ_2 D$ and $B \succeq_2 A \backslash B$, and it follows from monotonicity that $C \cup \{x_{2k}\} \succeq_2 D$ and $B \cup \{x_{2k}\} \succeq_2 A \backslash B$. We claim that at least one of $D \cup \{x_{2k}\} \succeq_2 C$ and $(A \backslash B) \cup \{x_{2k}\} \succeq_2 B$ holds.

  Assume for contradiction that $C \succ_2 D \cup \{x_{2k}\}$ and $B \succ_2 (A \backslash B) \cup \{x_{2k}\}$. Responsiveness implies that

  $$C \succ_2 D \cup \{x_{2k}\} \succeq_2 B \succ_2 (A \backslash B) \cup \{x_{2k}\} \succeq_2 C,$$

  a contradiction. Hence at least one of $D \cup \{x_{2k}\} \succeq_2 C$ and $(A \backslash B) \cup \{x_{2k}\} \succeq_2 B$ holds, as claimed.

In both cases, we can find in polynomial time a subset $E \subseteq A$ of size $k-1$ containing an item from each of the pairs $(x_1, x_2), (x_3, x_4), \ldots, (x_{2k-3}, x_{2k-2})$ such that $E \cup \{x_{2k}\} \succeq_2 A \backslash E$ and $(A \backslash E) \cup \{x_{2k}\} \succeq_2 E$.

We now choose our agreeable set of size $k+1$ as follows:

1. Choose both x_{2k-1} and x_{2k}.

2. Choose either E or $A \backslash E$ according to which set agent 3 prefers. (If agent 3 is indifferent between the two sets, choose one of them arbitrarily.)

We claim that our chosen set T is agreeable to all three agents. We prove the claim separately for each of the agents.

- For any $j = 1, 2, \ldots, m$, the set T contains at least $j/2$ of the j most preferred items according to $\succeq_1$. Since $\succeq_1$ is responsive, Proposition 3.2.5 implies that T is necessarily agreeable to agent 1.
- Since $E \cup \{x_{2k}\} \succeq_2 A\backslash E$ and $A\backslash E \cup \{x_{2k}\} \succeq_2 E$, and T contains either E or $A\backslash E$ along with both x_{2k-1} and x_{2k}, T is agreeable to agent 2.
- Since we choose the set E or $A\backslash E$ that agent 3 prefers and we include both of the remaining items x_{2k-1} and x_{2k}, T is agreeable to agent 3.

Hence T is agreeable to all three agents, as claimed. This concludes the analysis for the case where m is even.

Finally, assume that $m = 2k+1$ is odd. Our goal is to find a subset of size $\left\lfloor \frac{m+3}{2} \right\rfloor = k+2$ that is agreeable to all three agents. Let $S' = S\backslash\{x_1\}$. We apply the algorithm from the case of m even to choose a set $T \subseteq S'$ of size $k+1$ that is agreeable to all three agents when the universe considered is S'. It follows that $T \cup \{x_1\}$ is a subset of size $k+2$ that is agreeable to all three agents when the universe considered is S. □

3.3.3 Computing Small Necessarily Agreeable Sets

In this section, we consider a model in which the algorithm only has access to each agent's ranking over the items. We will therefore be interested in computing a small subset that is *necessarily* agreeable to every agent. While the algorithm has significantly less information at its disposal than before, as we will see, it is still possible to find small subsets that are necessarily agreeable to all agents.

If the algorithm had access to the agents' preferences over all subsets of items, Theorem 3.3.1 implies that it could always find a subset of size $\left\lfloor \frac{m+n}{2} \right\rfloor$ that is agreeable to all agents. For two agents, the algorithm in Theorem 3.3.3 only uses the agents' rankings to compute a subset of this size that is necessarily agreeable to both agents. As Example 3.3.4 shows, however, a necessarily agreeable subset of size $\left\lfloor \frac{m+n}{2} \right\rfloor$ might not exist even when there are three agents. Indeed, it is not clear how much extra "penalty" we have to pay for the information restriction that we are imposing. For example, it could be that with three agents, there already exist preferences on single items for which any necessarily agreeable subset contains at least cm items for some constant $c > 1/2$. We show in the next theorem that this is in fact not the case—there always exists a necessarily agreeable subset of size $m/2 + O(\log m)$ as long as the number of agents is constant. Moreover, such a set can be computed in polynomial time.

Theorem 3.3.6. *For any constant number of agents, there exists a subset of S of size $m/2+O(\log m)$ that is necessarily agreeable to all agents. Moreover, such a subset can be computed in polynomial time.*

To prove this theorem, we will use the following result from discrepancy theory due to Bohus [22].

Lemma 3.3.7 ([22])**.** *Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be permutations of the set $M = \{1, 2, \ldots, m\}$. There exists a function $f : M \to \{-1, 1\}$ such that for any $1 \le p \le q \le m$ and any $1 \le j \le n$,*

$$\left| \sum_{i=p}^{q} f(\sigma_j(i)) \right| \le 8n \log m.$$

Moreover, such a function can be computed in polynomial time.

Proof of Theorem 3.3.6. Suppose that agent j ranks the single items as $x_{\sigma_j(1)} \succeq_j^{sing} x_{\sigma_j(2)} \succeq_j^{sing} \cdots \succeq_j^{sing} x_{\sigma_j(m)}$. By Lemma 3.3.7, we can efficiently compute a function $f : S \to \{-1, 1\}$ such that

$$\left| \sum_{i=1}^{q} f(x_{\sigma_j(i)}) \right| \le 8n \log m$$

for any $q = 1, 2, \ldots, m$ and any $j = 1, 2, \ldots, n$.

We construct our agreeable subset as follows. We include in our subset all items x_i such that $f(x_i) = 1$, as well as the $\lceil 4n \log m \rceil$ most preferred items of each agent that are not yet included. (If some agent has fewer than $\lceil 4n \log m \rceil$ items that are not yet included, we simply include all of that agent's items.) For $i = 1, 2, \ldots, m$, let X_i be an indicator variable such that $X_i = 1$ if item x_i is included in the set and $X_i = -1$ if not. For any agent j and any $i = 1, 2, \ldots, m$, we have

$$X_{\sigma_j(1)} + \cdots + X_{\sigma_j(i)} \ge \min\{i, -8n \log m + 2 \cdot \lceil 4n \log m \rceil\} \ge 0.$$

By Proposition 3.2.5, this implies that the chosen set is necessarily agreeable to all agents. Moreover, our subset includes at most

$$\frac{m}{2} + (n+1) \cdot \lceil 4n \log m \rceil = \frac{m}{2} + O(\log m)$$

items, as desired. □

Next, we address the tightness of the bound in Theorem 3.3.6. Bohus's result is known to be asymptotically tight for constant n: Newman et al. [118] constructed, for every m that is a power of three, an example of three permutations whose discrepancy is $\Omega(\log m)$. While upper bounds on the discrepancy of permutations can be easily turned into upper bounds on the size of necessarily agreeable sets as seen above, lower bounds are somewhat more delicate. Nevertheless, Newman et al.'s examples satisfy stronger conditions than merely having a large discrepancy. One of these conditions, which we state in the following lemma, will be sufficient for proving a lower bound on the size of necessarily agreeable sets. The lemma is a restatement of Corollary 2 in the work of Newman et al. [118].

Lemma 3.3.8 ([118])**.** *Given any positive integer k, let $m = 3^k$ and $M = \{1, 2, \dots, m\}$. There exist three permutations $\sigma_1, \sigma_2, \sigma_3$ of M such that for any function $f : M \to \{-1, 1\}$, if $\Delta := \sum_{i \in M} f(i) \geq 1$, then there exist $1 \leq q \leq m$ and $1 \leq j \leq 3$ such that*

$$\sum_{i=1}^{q} f(\sigma_j(i)) \leq \frac{-k + 2\Delta - 2}{3}.$$

We now show that the bound in Theorem 3.3.6 is tight even when there are three agents. Recall that if there are two agents, it is possible to compute a subset of size $\lfloor \frac{m+2}{2} \rfloor$ that is necessarily agreeable to both agents (Theorem 3.3.3).

Theorem 3.3.9. *Suppose that $m = 3^k$ for some positive integer k. There exist preferences on single items of three agents such that every necessarily agreeable subset of items have size at least $m/2 + \Omega(\log m)$.*

Proof. Let $\sigma_1, \sigma_2, \sigma_3$ be the permutations of S from Lemma 3.3.8, where we use $S = \{x_1, x_2, \dots, x_m\}$ in place of $M = \{1, 2, \dots, m\}$. For each $j = 1, 2, 3$, let the preference on single items of agent j be $x_{\sigma_j(1)} \succeq_j^{sing} x_{\sigma_j(2)} \succeq_j^{sing} \cdots \succeq_j^{sing} x_{\sigma_j(m)}$.

Consider any subset $T \subseteq S$ of size at most $m/2 + k/4$. We will show that T cannot be necessarily agreeable to all three agents, which immediately implies the theorem since $k = \log_3 m$. To see that this is the case, let $f_T : S \to \{-1, 1\}$ denote the indicator function of T, i.e., $f_T(x_i) = 1$ if $x_i \in T$ and $f_T(x_i) = -1$ if $x_i \notin T$. Since T is of size at most $m/2 + k/4$, we have $\Delta_T := \sum_{i=1}^{m} f_T(x_i) = |T| - |S \setminus T| \leq k/2$. If $\Delta_T < 0$, then T is not necessarily agreeable by Proposition 3.2.5, so we may assume that $\Delta_T \geq 0$. Since m is odd and Δ_T is an integer, we also have $\Delta_T \geq 1$. By Lemma 3.3.8, there exists $1 \leq q \leq m$ and $1 \leq j \leq 3$ such that

$$\sum_{i=1}^{q} f_T(\sigma_j(x_i)) \leq \frac{-k + 2\Delta_T - 2}{3} \leq \frac{-k + k - 2}{3} < 0.$$

By Proposition 3.2.5, T is not necessarily agreeable to agent j, as desired. □

Theorems 3.3.6 and 3.3.9 show that the bound $m/2 + O(\log m)$ for the size of the smallest necessarily agreeable set is asymptotically tight. We next present a randomized algorithm that, despite its simplicity, computes a necessarily agreeable subset of size $m/2 + O(\sqrt{m})$ in polynomial time. The algorithm works by first choosing whether to include each item independently with 50% probability, and then including the $O(\sqrt{m})$ most preferred items of each agent that were excluded in the first step.

For the analysis of the algorithm, we will require two probabilistic results: the Chernoff bound and Lévy's inequality. The statements of both results can be found in Section 2.3.

Theorem 3.3.10. *Assume that the number of agents is constant. Let $\epsilon \in (0,1)$, and let $c > 0$ be a constant such that $e^{-2c^2/3} < \epsilon/4n$. Consider the following randomized polynomial-time algorithm:*

1. *For each item, either include it in our set or not with probability $1/2$, independently of the remaining items.*
2. *Include the $\lceil c\sqrt{m}\rceil$ most preferred items of each agent that were excluded in Step 1.*

With probability at least $1-\epsilon$, the algorithm computes a subset of size $m/2+O(\sqrt{m})$ that is necessarily agreeable to all agents.

Proof. Let $X_1, X_2, \ldots, X_m$ be independent random variables such that $X_i = 1$ if item x_i is included in our subset in the first step, and $X_i = 0$ if not. By the definition of the algorithm, each X_i is either 0 or 1 with probability $1/2$, independently of the other X_i's.

For $j = 1, 2, \ldots, m$, suppose that agent j ranks the single items as $x_{\sigma_j(1)} \succeq_j^{sing} x_{\sigma_j(2)} \succeq_j^{sing} \cdots \succeq_j^{sing} x_{\sigma_j(m)}$. Let $Y_i^j := X_{\sigma_j(1)} + \cdots + X_{\sigma_j(i)}$ for $i = 1, 2, \ldots, n$. For any agent j, we have $\mathbb{E}[Y_m^j] = m/2$. The Chernoff bound (Lemma 2.3.1) with $\delta = 2c/\sqrt{m}$ implies that

$$\Pr\left[Y_m^j \geq \frac{m}{2} + c\sqrt{m}\right] \leq e^{-\frac{2c^2}{3}}.$$

Similarly,

$$\Pr\left[Y_m^j \leq \frac{m}{2} - c\sqrt{m}\right] \leq e^{-c^2}.$$

Combining the two inequalities, we have

$$\Pr\left[\left|Y_m^j - \frac{m}{2}\right| \geq c\sqrt{m}\right] \leq e^{-\frac{2c^2}{3}} + e^{-c^2} \leq 2e^{-\frac{2c^2}{3}}.$$

Using Lévy's inequality (Lemma 2.3.2) with the random variables $X_{\sigma_j(i)} - 1/2$ for $i = 1, 2, \ldots, n$, it follows that

$$\Pr\left[\max_{1\leq i\leq m}\left|Y_i^j - \frac{i}{2}\right| \geq c\sqrt{m}\right] \leq 4e^{-\frac{2c^2}{3}}.$$

Using the union bound over all agents j, we have

$$\Pr\left[\max_{1\leq i\leq m}\left|Y_i^j - \frac{i}{2}\right| \geq c\sqrt{m} \text{ for some } j \in \{1, 2, \ldots, m\}\right] \leq 4ne^{-\frac{2c^2}{3}} < \epsilon.$$

Hence, with probability at least $1-\epsilon$,

$$\left|Y_i^j - \frac{i}{2}\right| < c\sqrt{m}$$

for all $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$, or equivalently, $Y_i^j - i/2 \in (-c\sqrt{m}, c\sqrt{m})$ for all i, j. Now, since in Step 2 we include the $\lceil c\sqrt{m} \rceil$ most preferred items of each agent that were excluded in Step 1, if we update the random variables to reflect these changes, we have $Y_i^j - i/2 \geq 0$ for all i, j. By Proposition 3.2.5, the set that the algorithm returns is necessarily agreeable to all agents. Since we include at most $m/2 + c\sqrt{m}$ items in Step 1 and $\lceil c\sqrt{m} \rceil$ items for each agent in Step 2, the set contains at most

$$\frac{m}{2} + (n+1) \cdot \lceil c\sqrt{m} \rceil = \frac{m}{2} + O(\sqrt{m})$$

items, as desired. □

3.4 Efficient Approximation

While our results in Sections 3.3 provide insights on small agreeable sets and how to compute them efficiently, an important issue is still left unaddressed by these results. In many instances, the minimum size of an agreeable set is much smaller than the worst-case bound over all instances with that number of agents and items. Indeed, an extreme example is when there is a single item that every agent likes more than all of the remaining items combined. In this case, it suffices to select that item alone. This results in a much smaller set than the worst-case bound, which is at least half of the items for any number of agents.

In this section, we investigate the problem of computing an agreeable subset of approximately optimal size for any given instance, as opposed to one whose size matches the worst-case bound over all instances with the same number of agents and items. We show that finding an optimal agreeable set is computationally hard, and therefore focus on finding an approximate solution. We do so using two well-known models for representing preferences, namely the value oracle model and additive valuations. For each of these models, we present polynomial-time algorithms for computing an agreeable set of approximately optimal size. Moreover, we show that the approximation ratios obtained by our algorithms are asymptotically tight for both models.

3.4.1 General Preferences

We begin with a model in which agents can have arbitrary preferences on subsets of items. Recall that our results so far do not yield any guarantee on the approximation ratio beyond the obvious $O(m)$ upper bound for arbitrary preferences over subsets of items. The goal of this section is to explore the approximation ratios that we can achieve in this general setting.

Before we move on to our results, let us be more precise about the model that we work with. First, we work with the agents' utility functions $u_1, u_2, \ldots, u_n$ instead of directly with the preferences themselves. Since the number of subsets of S is exponentially large, the utility functions take exponential space to write down. For this reason, it is undesirable to include them as part of the input. Instead, we work with the *value oracle model* [63], in which the algorithm can query the value of $u_i(T)$ for any subset $T \subseteq S$ and any $i = 1, 2, \ldots, n$. We also note that we do not assume responsiveness of the agents' preferences in this section.

Our first result is a simple polynomial-time approximation algorithm with approximation ratio $O(m/\log m)$. Even though this approximation guarantee is only $\Omega(\log m)$ better than the obvious $O(m)$ bound, we will see later that this is already the best we can hope for in polynomial time.

Theorem 3.4.1. *There exists a polynomial-time $O(m/\log m)$-approximation algorithm for computing a minimum size agreeable set in the value oracle model.*

Proof. We start by partitioning the set S of items into $\lceil \log m \rceil$ parts $S_1, \ldots, S_{\lceil \log m \rceil}$, where each part is of size at most $\lceil m/\log m \rceil$. For each set $A \subseteq \{1, 2, \ldots, \lceil \log m \rceil\}$, we check whether the set $\bigcup_{i \in A} S_i$ is agreeable or not by comparing each agent's value for the set to that for its complement. We then output the smallest agreeable set that we find. Since the number of possible sets A is linear in m, the running time of our algorithm is polynomial in m and n.

To prove the approximation guarantee of the algorithm, let S^* be a smallest agreeable set. Suppose that $|S^*| = k$. By monotonicity, the union of all sets S_i containing elements of S^* is also agreeable, and it is one of the sets that we check. Moreover, this union has size at most $k \cdot \lceil m/\log m \rceil$, implying that our algorithm indeed has approximation ratio $O(m/\log m)$. ☐

Even though our algorithm is very simple, we show next that its approximation guarantee is in fact the best one can hope for, even when there is a single agent.

Theorem 3.4.2. *For every constant $c > 0$, there exists m_0 such that for every $m > m_0$, there is no (possibly randomized and adaptive) algorithm that makes at most $m^{c/8}$ queries to the value oracle and always outputs an agreeable set with expected size at most $m/(c \log m)$ times the optimum, even when there is only one agent.*

In other words, the above theorem implies that there is no polynomial time algorithm with approximation ratio $o(m/\log m)$. We note here that our lower bound is information-theoretic and is not based on any computational complexity assumptions. Moreover, it rules out any algorithm that makes a polynomial number of queries, not only those that run in polynomial time.

Proof of Theorem 3.4.2. Let $g : \mathcal{S} \to \mathbb{R}_{\geq 0}$ be a function such that

$$g(T) = \begin{cases} 1 & \text{if } |T| \geq m/2; \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, for each subset $T^* \subseteq S$, let $f_{T^*} : \mathcal{S} \to \mathbb{R}_{\geq 0}$ denote the function

$$f_{T^*}(T) = \begin{cases} 1 & \text{if } |T| \geq m/2 \text{ or } T^* \subseteq T; \\ 0 & \text{otherwise.} \end{cases}$$

That is, f_{T^*} is g together with a planted solution T^*.

Consider any algorithm $\mathcal{A}$ that makes at most $m^{c/8}$ queries. Assume for the moment that $\mathcal{A}$ is deterministic. Let us examine a run of $\mathcal{A}$ when the agent's utility function is g. Suppose that $\mathcal{A}$'s queries to g are on the sets $T_1, T_2, \ldots, T_{\lfloor m^{c/8} \rfloor} \subseteq S$.

Let T^* be a subset of S of size $\lfloor c \log m/4 \rfloor$ chosen uniformly at random. Consider the queries that $\mathcal{A}$ makes when the agent's utility function is f_{T^*}; suppose that the queries made are on the sets $T'_1, T'_2, \ldots, T'_{\lfloor m^{c/8} \rfloor} \subseteq S$. For each $j = 1, 2, \ldots, \lfloor m^{c/8} \rfloor$, if $T_i = T'_i$ and $g(T_i) = f_{T^*}(T'_i)$ for all $i = 1, 2, \ldots, j-1$, then $\mathcal{A}$ goes through the same computation route for both g and f_{T^*}, and hence $T_j = T'_j$. Moreover, when both runs share the same computational route so far and $T_j = T'_j$, we can bound the probability that $g(T_j) \neq f_{T^*}(T'_j)$ as follows. First, if $|T_j| \geq m/2$, then $g(T_j)$ is always equal to $f_{T^*}(T'_j)$. Otherwise, we have

$$\Pr[g(T_j) \neq f_{T^*}(T'_j)] = \Pr[g(T_j) \neq f_{T^*}(T_j)] = \Pr[T^* \subseteq T_j].$$

If $|T_j| < |T^*|$, this probability is 0. Else, since T_j is independent of T^*, we can bound the probability as

$$\begin{aligned} \Pr[T^* \subseteq T_j] &= \frac{\binom{|T_j|}{\lfloor c \log m/4 \rfloor}}{\binom{m}{\lfloor c \log m/4 \rfloor}} \\ &= \left(\frac{|T_j|}{m}\right)\left(\frac{|T_j|-1}{m-1}\right)\cdots\left(\frac{|T_j| - \lfloor c \log m/4 \rfloor + 1}{m - \lfloor c \log m/4 \rfloor + 1}\right) \\ &\leq \left(\frac{|T_j|}{m}\right)^{\lfloor c \log m/4 \rfloor} \\ &\leq 2^{-\lfloor c \log m/4 \rfloor} \\ &\leq 2m^{-c/4}, \end{aligned}$$

where the last inequality holds for large enough m.

By the union bound, the probability that the two sequences of queries are not identical is at most $(2m^{-c/4}) \cdot m^{c/8} = 2m^{-c/8}$, which is less than $1/2$ when m is sufficiently large. Furthermore, observe that when the two sequences are identical, $\mathcal{A}$ must output an agreeable subset with respect to the utility function g; any such set is of size at least $m/2$. Thus, the expected size of the output of $\mathcal{A}$ when given the utility function f_{T^*} is more than $m/2 \cdot (1/2) = m/4$. However, the optimal agreeable set for f_{T^*} has size only $\lfloor c \log m/4 \rfloor$. As a result, the expected size of the output of $\mathcal{A}$ is more than $m/(c \log m)$ times the optimum, as desired.

Finally, note that if $\mathcal{A}$ is randomized, we can use the above argument on each choice of randomness and average over all the choices, which gives a similar conclusion. □

We remark that the same result holds even if we require the utility function of the agent to be subadditive or submodular.[5] To obtain the proof for a subadditive utility function, for any $T \neq \emptyset$ such that $g(T) = 0$, we set instead $g(T) = 1/2$; we perform an analogous modification to f_{T^*}. Subadditivity holds for g since

$$g(A \cup B) \leq 1 = 1/2 + 1/2 \leq g(A) + g(B)$$

for any $A, B \neq \emptyset$, and similarly for f_{T^*}. The rest of the proof then proceeds as before.

On the other hand, more work is required to adapt the proof to submodular functions. In particular, we let $k = \lfloor c \log m/4 \rfloor$ and define g as follows:

$$g(T) = \begin{cases} 1 & \text{if } |T| \geq m/2, \\ 1 - \frac{1}{2^{|T|-k}(k+1)} & \text{if } k \leq |T| < m/2, \\ \frac{|T|}{k+1} & \text{otherwise.} \end{cases}$$

Likewise, for any T such that originally $f_{T^*}(T) = 0$, we modify the value of $f_{T^*}(T)$ to be the same as $g(T)$. One can check that g and f_{T^*} are submodular, and the proof again proceeds as before.

3.4.2 Additive Utilities

In this section, we assume that the agents' preferences are represented by additive utility functions. Each agent i has some nonnegative utility $u_i(x_j)$ for item x_j, and $u_i(T) = \sum_{x \in T} u_i(x)$ for any subset of items $T \subseteq S$.

Clearly, the problem of deciding whether there exists an agreeable set of a certain size is in NP. The following theorem shows that it is NP-complete, even when there are two agents. Note that if there is only one agent, it is easy to find an optimal agreeable set by repeatedly choosing an item that yields the highest utility to the agent among the remaining items until the set of chosen items is agreeable.

Theorem 3.4.3. *For two agents with additive utility functions, it is NP-hard to decide whether there is an agreeable set of size exactly $m/2$.*

Proof. We will reduce from the following problem called Balanced 2-Partition: Given a multiset A of non-negative integers, decide whether there exists a subset $B \subseteq A$ such that $|B| = |A \setminus B| = |A|/2$ and $\sum_{a \in B} a = \sum_{a \in A \setminus B} a = \sum_{a \in A} a/2$.

[5] A function $f : \mathcal{S} \to \mathbb{R}_{\geq 0}$ is said to be *submodular* if $f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$ for any $A \subseteq B \subseteq S$ and $x \in S \backslash B$. Any submodular function is also subadditive.

Like the well-known 2-Partition problem where the cardinality constraint is not included, Balanced 2-Partition is NP-hard. For completeness, we give a proof of NP-hardness of Balanced 2-Partition in Appendix A.2.

The reduction from Balanced 2-Partition proceeds as follows. Let $a_1, \ldots, a_{|A|}$ be the elements of A. The set S contains $|A|$ items $x_1, \ldots, x_{|A|}$, each associated with an element of A. The utility functions are then defined by $u_1(x_i) = a_i$ and $u_2(x_i) = M - a_i$, where $M = \sum_{a \in A} a$. We show next that this reduction is indeed a valid reduction.

(YES Case) Suppose that there exists $B \subseteq A$ such that $|B| = |A|/2$ and $\sum_{a \in B} a = \sum_{a \in A} a/2$. Let T be the set of all items corresponding to the elements of B. It is obvious that T has size $|A|/2 = m/2$ and that T is agreeable.

(NO Case) We prove the contrapositive; suppose that there is an agreeable subset $T \subseteq S$ of size $m/2$. Let $B \subseteq A$ be the set of elements corresponding to the items in T. Since T is agreeable, $\sum_{x \in T} u_i(x) \geq \sum_{x \in S \setminus T} u_i(x)$ for $i = 1, 2$. When $i = 1$, this implies that $\sum_{a \in B} a \geq \sum_{a \in A} a/2$. When $i = 2$, using the fact that $|T| = m/2$, we have $\sum_{a \in B} a \leq \sum_{a \in A} a/2$. It follows that $\sum_{a \in B} a = \sum_{a \in A} a/2$. Since $|B| = m/2 = |A|/2$, this concludes the proof. □

Theorem 3.4.3 shows that the problem is weakly NP-hard even when there are two agents. Nevertheless, when the number of agents is constant, the following theorem shows that there exists a pseudo-polynomial time dynamic programming algorithm for computing an optimal agreeable set. In particular, the problem is not strongly NP-hard for any constant number of agents.

Theorem 3.4.4. *For any constant number of agents with additive utility functions, there exists a pseudo-polynomial time algorithm that computes an agreeable set of minimum size.*

Proof. The algorithm uses dynamic programming. Assume that the utilities of agent i for the items are nonnegative integers with sum σ_i. We construct a table Σ of size $(m+1)(\sigma_1+1)\ldots(\sigma_n+1)$, where for each $0 \leq m' \leq m$ and each tuple $(y_1, \ldots, y_n)$ with $0 \leq y_i \leq \sigma_i$, the entry $\Sigma(m', y_1, \ldots, y_n)$ of the table corresponds to the minimum number of items from among the items $x_1, x_2, \ldots, x_{m'}$ that we need to include so that agent i has utility exactly y_i for all i (if this is achievable). Initially we have $\Sigma(0, 0, \ldots, 0) = 0$ and $\Sigma(m', y_1, \ldots, y_n) = \infty$ otherwise. We then iterate through the values of m' in increasing order. For each $m' \geq 1$, we update the entries of the table as follows:

- If $u_i(x_{m'}) \leq y_i$ for all i and

  $$1 + \Sigma\left(m'-1, y_1 - u_1(x_{m'}) \ldots, y_n - u_n(x_{m'})\right) < \Sigma(m'-1, y_1, \ldots, y_n),$$

  let $\Sigma(m', y_1, \ldots, y_n) = 1 + \Sigma\left(m'-1, y_1 - u_1(x_{m'}) \ldots, y_n - u_n(x_{m'})\right)$.

- Else, let $\Sigma(m', y_1, \ldots, y_n) = \Sigma(m'-1, y_1, \ldots, y_n)$.

Finally, we look up the entries $\Sigma(m, y_1, \ldots, y_n)$ such that $y_i \geq \sigma_i/2$ for all i and return the minimum value over all such entries. The algorithm runs in time $O(m\sigma_1 \ldots \sigma_n)$. Note that if we also want to return an agreeable set (rather than just the size), we can also keep track of the sets of items along with the values in our table. □

While there is a pseudo-polynomial time algorithm for the problem when the number of agents is constant, we show next that if the number of agents is not constant, the problem becomes strongly NP-hard. In other words, there is no pseudo-polynomial time algorithm for this variant unless P=NP.

Theorem 3.4.5. *When the number of agents is not constant, it is strongly NP-hard to decide whether there is an agreeable set of size exactly* $(m+1)/2$.

Proof. We reduce from 3SAT. Given a 3SAT formula ϕ with m' clauses $C_1, C_2, \ldots, C_{m'}$ on n' variables $y_1, y_2, \ldots, y_{n'}$, let there be $n = m' + n'$ agents, where we abuse notation and call the agents $C_1, C_2, \ldots, C_{m'}, y_1, y_2, \ldots, y_{n'}$, and $m = 2n' + 1$ items, where $2n'$ items correspond to all the literals $y_1, \neg y_1, y_2, \neg y_2, \ldots, y_{n'}, \neg y_{n'}$ and the remaining item is called a. We assume without loss of generality that each clause of ϕ has at least two variables—it is obvious that every 3SAT formula can be transformed into this form in polynomial time. The utility functions of the agents are defined by

$$u_{C_i}(b) = \begin{cases} 1 & \text{if } b = a \text{ or the literal } b \text{ is present in } C_i; \\ 0 & \text{otherwise;} \end{cases}$$

and

$$u_{y_i}(b) = \begin{cases} 1 & \text{if } b = a, b = y_i, \text{ or } b = \neg y_i; \\ 0 & \text{otherwise.} \end{cases}$$

We show next that this is a valid reduction. First, note that all of the integer parameters are polynomial in the size of the input. Hence, we are left to show that YES and NO instances of 3SAT map to YES and NO instances of our problem respectively.

(YES Case) Suppose that there exists an assignment that satisfies ϕ. For each y_i, let b_i be the literal of y_i that is true according to this assignment. Let $T = \{a, b_1, b_2, \ldots, b_{n'}\}$. Since each clause C_j is satisfied by the assignment, we have $\sum_{i=1}^{n'} u_{C_j}(b_i) \geq 1$. It follows that $\sum_{x \in T} u_{C_j}(x) \geq 2$, and therefore $T \succeq_{C_j} S \setminus T$. Moreover, for each variable y_i, we have $\sum_{x \in T} u_{y_i}(x) = 2$, which also implies that $T \succeq_{y_i} S \setminus T$. As a result, T is an agreeable set of size $n' + 1 = (m+1)/2$ as desired.

(NO Case) We prove the contrapositive; suppose that there exists an agreeable set $T \subseteq S$ of size $(m+1)/2 = n' + 1$. We assume without loss of generality that $a \in T$; indeed, since the utility of any agent for a is at least as much as the utility of the agent for any other item, if $a \notin T$ we can replace an arbitrary item in T by a and maintain the agreeability of T.

Since $T \succeq_{y_i} S \setminus T$, at least one literal corresponding to y_i is included in T. Moreover, since the size of T is $n' + 1$ and $a \in T$, exactly one literal of each y_i is in T; let b_i be this literal. Consider the assignment to the variables such that all the b_i's are set to true. Since $T \succeq_{C_j} S \setminus T$ for every C_j and C_j contains at least two literals, at least one literal in C_j is set to true by this assignment. Hence the assignment satisfies the formula ϕ. □

Given that computing an agreeable set of minimum size is NP-hard, it is natural to attempt to find an approximation algorithm for the problem. When the utilities are additive, this turns out to be closely related to approximating the classical problem Set Cover. In Set Cover, we are given a ground set U and a collection $\mathcal{C}$ of subsets of U. The goal is to select a minimum number of subsets whose union is the entire set U.

Set Cover was one of the first problems shown to be NP-hard in Karp's seminal paper [82]. Since then, its approximability has been intensively studied and is now well understood. A simple greedy algorithm yields a $(\ln |U| + 1)$-approximation for the problem [81, 102]. On the other hand, a long line of work in hardness of approximation [4, 61, 104, 116, 123] culminates in Dinur and Steurer's work [55], in which the NP-hardness of approximating Set Cover within a factor of $(1 - \varepsilon) \ln |U|$ was proved for every constant $\varepsilon > 0$.

The first connection we will make between Set Cover and approximating minimum size agreeable set is on the negative side—we will show that any inapproximability result for Set Cover can be translated to that for approximating minimum size agreeable set as well. To do so, we will first state Dinur and Steurer's result more precisely.

Lemma 3.4.6 ([55]). *For every constant $\varepsilon > 0$, there is a polynomial time reduction from any* 3SAT *formula ϕ to a* Set Cover *instance $(U, \mathcal{C})$ and a function $f(U)$ which is polynomial in $|U|$ such that*

- *(Completeness) if ϕ is satisfiable, the optimum of $(U, \mathcal{C})$ is at most $f(U)$;*
- *(Soundness) if ϕ is unsatisfiable, the optimum of $(U, \mathcal{C})$ is at least $((1 - \varepsilon) \ln |U|) f(U)$.*

We are now ready to prove the hardness of approximating minimum size agreeable set.

Theorem 3.4.7. *For any constant $\delta > 0$, it is NP-hard to approximate minimum size agreeable set to within a factor $(1 - \delta) \ln n$ of the optimum.*

Proof. Let $\varepsilon = \delta/2$. Given a 3SAT formula ϕ, we first use Dinur and Steurer's reduction to produce a Set Cover instance $(U, \mathcal{C})$. Let there be $|U|$ agents, each of whom is associated with a distinct element of U; it is convenient to think of the set of agents as simply U. As for the items, let there be one item for each subset $C \in \mathcal{C}$ and additionally let there be one special item called t. In other words, $S = \mathcal{C} \cup \{t\}$.

The utility function of each agent $a \in U$ is then defined by

$$u_a(s) = \begin{cases} |\{C \in \mathcal{C} \mid a \in C\}| - 1 & \text{if } s = t; \\ 1 & \text{if } s \in \mathcal{C} \text{ and } a \in s; \\ 0 & \text{otherwise.} \end{cases}$$

We show next that this reduction indeed gives the desired inapproximability result.

(Completeness) If ϕ is satisfiable, then there are $f(U)$ subsets from $\mathcal{C}$ that together cover U. We can take T to contain all of these subsets and the special item t. Clearly, T has size $f(U) + 1$ and is agreeable.

(Soundness) If ϕ is unsatisfiable, then any set cover of $(S, \mathcal{C})$ contains at least $((1-\varepsilon)\ln|U|)f(U)$ subsets. Consider any agreeable set T. For each $a \in U$, from our definition of $u_a(t)$, the set T must include at least one subset that contains a. In other words, $T \setminus \{t\}$ is a set cover of $(S, \mathcal{C})$. Hence, $|T| \geq ((1-\varepsilon)\ln|U|)f(U)$.

The two parts together imply that it is NP-hard to approximate minimum size agreeable set to within a factor $\frac{((1-\varepsilon)\ln|U|)f(U)}{f(U)+1}$ of the optimum. This ratio is at least $(1-\delta)\ln n$ when $f(U) \geq 2/\delta$, which can be assumed without loss of generality (since otherwise we can solve the Set Cover instance in time $|U|^{O(f(U))} = |U|^{O(1)}$, implying that P = NP). □

Unlike the above inapproximability result, it is unclear how algorithms for Set Cover can be used to approximate minimum size agreeable set. Fortunately, our problem is in fact a special case of a generalization of Set Cover called Covering Integer Program (CIP), which can be written as follows:

$$\begin{aligned} \text{minimize } & c^T x \\ \text{subject to } & Ax \geq 1, \\ & 0 \leq x \leq u, \\ & x \in \mathbb{Z}^m, \end{aligned}$$

where $c, u \in \mathbb{R}^m$ and $A \in \mathbb{R}^{n \times m}$ are given as input.

The problem of finding a minimum size agreeable set can be formulated in this form by setting c, u and A as follows:

$$\begin{aligned} c_s &= 1 & \forall s \in S \\ u_s &= 1 & \forall s \in S \\ A_{i,s} &= \frac{2u_i(s)}{\sum_{s' \in S} u_i(s')} & \forall i = 1, 2, \ldots, n, \forall s \in S \end{aligned}$$

Similarly to Set Cover, the approximability of CIP is well studied. In particular, the problem is known to be approximable to within a factor $O(\log n)$ of the optimum in polynomial time [90]. This immediately implies an $O(\log n)$-approximation algorithm for finding a minimum size agreeable set as well.

Theorem 3.4.8. *For agents with additive utility functions, there exists a polynomial-time $O(\log n)$-approximation algorithm for computing a minimum size agreeable set.*

3.5 Conclusion and Future Work

In this chapter, we introduce the notion of agreeability, which captures an agent's acceptance of the set of items allocated to her group, and present a number of fundamental results on the notion. For any number of agents and items, we derive a tight upper bound on the number of items that may need to be included in an agreeable subset. We also present polynomial-time algorithms for computing an agreeable set whose size matches the upper bound or approximates the optimal size for a given instance using well-known models for representing preferences.

Our work suggests a number of possible future directions. With polynomial-time algorithms for computing an agreeable set whose size matches the upper bound for two and three agents in hand, a natural question is whether we can similarly obtain efficient algorithms when there are more agents. The algorithm for three agents is already quite involved, so one might suspect that the problem is intractable for larger numbers of agents. If that were to be the case, it would be useful to have a confirmation by means of a hardness result, even for some fixed large number of agents. Since the problem is a search problem for which we know that a solution always exists, it cannot be NP-hard, but could potentially be hard with respect to a subclass of TFNP such as PPAD or PLS. One could also investigate the complexity of deciding the existence of agreeable subsets of certain sizes for which there is no guarantee of existence, as we do in Theorems 3.4.3 and 3.4.5.

Another avenue for future work is to extend the notion of agreeability to more general settings. For instance, a motivating example that we give is that the group of agents receive some items as prizes from a team competition that they won against another group. One could consider a generalization where there are more than two competing groups. However, in this case there are several reasonable ways of defining agreeability, since we do not know how the remaining items are distributed among the remaining groups. One possibility is to require that each agent in the group find the set of items to be worth at least $1/k$ of the whole set, where k is the number of groups. An alternative definition is to impose the condition that for each agent in the group, we can partition the remaining items among the other $k-1$ groups so that the agent does not envy any of the other groups. While both definitions reduce to our notion in the case of two groups and additive utilities, the equivalence ceases to hold when there are at least three groups or if utilities are not additive. As such, the results that we can obtain will likely depend on the definition that we use.

Chapter 4

Asymptotic Existence of Fair Divisions for Groups

4.1 Introduction

In this chapter, we study envy-free divisions in the group setting. Given that such divisions do not always exist even in the individual setting,[1] we investigate in Section 4.3 the asymptotic existence and non-existence of envy-free divisions using a probabilistic model, previously used in the setting with one agent per group [54]. We show in Section 4.3.1 that under additive valuations and other mild technical conditions, when all groups contain an equal number of agents, an envy-free division is likely to exist if the number of goods exceeds the total number of agents by a logarithmic factor, no matter whether the agents are distributed into several groups of small size or few groups of large size (Theorem 4.3.1). In particular, any allocation that maximizes social welfare is likely to be envy-free. In addition, when there are two groups with possibly unequal numbers of agents and the distribution on the valuation of each good is symmetric, an envy-free division is likely to exist if the number of goods exceeds the total number of agents by a logarithmic factor as well (Theorem 4.3.2). Although it might not be surprising that a welfare-maximizing allocation is envy-free with high probability when there are sufficiently many goods, the fact that only an extra logarithmic factor is required is perhaps somewhat unexpected. Indeed, as the number of agents in each group increases, it seems as though the independence between the preferences of each agent would make it much harder to satisfy all of them simultaneously, since they all need to be allocated the same goods.

To complement our existence results, we show on the other hand in Section 4.3.2 that we cannot get away with a much lower number of goods and still have an envy-free division with high probability. In particular, if the number of goods is less than the total number of agents by a superconstant factor,

[1] See Section 2.1.

or if the number of goods is less than the total number of agents and the number of groups is large, the probability that an envy-free division exists is low (Corollaries 4.3.5 and 4.3.6). This leaves the gap between asymptotic existence and non-existence of envy-free divisions at a mere logarithmic factor.

While the techniques used to show asymptotic existence of envy-free divisions in Section 4.3 give rise to mechanisms that compute such divisions with high probability, these mechanisms are unfortunately not truthful. In other words, implementing these mechanisms in the presence of strategic agents can lead to undesirable outcomes. In Section 4.4, we tackle the issue of truthfulness and show that a simple truthful mechanism, namely the random assignment mechanism, is *α-approximate envy-free* with high probability for any constant $\alpha \in [0, 1)$ (Theorem 4.4.1). Approximate envy-freeness means that even though an agent may envy another agent in the resulting division, the values of the agent for her own allocation and for the other agent's allocation differ by no more than a multiplicative factor of α. In other words, the agent's envy is relatively small compared to her value for her own allocation. The number of goods required to obtain approximate envy-freeness with high probability increases as we increase α. Our result shows that it is possible to achieve truthfulness and approximate envy-freeness simultaneously in a wide range of random instances, and improves upon the previous result for the setting with one agent per group [6] in several ways.

Our results in Section 4.3 can be viewed as generalizations of previous results by Dickerson et al. [54], who showed asymptotic existence and non-existence under a similar model but in a more limited setting where each group has only one agent. In particular, these authors proved that under certain technical conditions on the probability distributions, an allocation that maximizes social welfare is envy-free with high probability if the number of goods is larger than the number of agents by a logarithmic factor. In fact, their result also holds when the number of agents stay constant, as long as the number of goods goes to infinity. Similarly, we show that a welfare-maximizing allocation is likely to be envy-free if the number of goods exceeds the number of agents by a logarithmic factor. While we require that the number of agent per group goes to infinity, the number of groups can stay small, even constant. On the non-existence front, Dickerson et al. showed that if the utility for each good is independent and identically distributed across agents, then envy-free allocations are unlikely to exist when the number of goods is larger than the number of agents by a linear fraction. On the other hand, our non-existence results apply to the regime where the number of goods is smaller than the number of agents. Note that while this regime is uninteresting in Dickerson et al.'s setting since envy-free allocations cannot exist, in our generalized setting an envy-free allocation can already exist when the number of goods is at least the number of *groups*.

Besides the asymptotic results on envy-free divisions, results of this type have also been shown for other fairness notions, including proportionality and the maximin share criterion. These two notions are weaker than envy-freeness when utilities are additive. Suksompong [149] showed that proportional allocations exist with high probability if the number of goods is a multiple of the number of agents or if the number of goods grows asymptotically faster than the number of agents. Kurokawa et al. [92] showed that if either the number of agents or the number of goods goes to infinity, then an allocation satisfying the maximin share criterion is likely to exist as long as each probability distribution has at least constant variance. Amanatidis et al. [7] analyzed the rate of convergence for the existence of allocations satisfying the maximin share criterion when the utilities are drawn from the uniform distribution over the unit interval.

4.2 Preliminaries

The basic definitions and notation of the group fair division setting are introduced in Section 2.1. We assume in this chapter that agents have additive utility functions. We may suppose without loss of generality that $u_{ij}(g) \in [0, 1]$ for each agent a_{ij} and each good g, since otherwise we can scale down all utilities by their maximum. We also assume for most of the chapter that all groups contain the same number of agents $n' := n_1 = n_2 = \cdots = n_k$. The *social welfare* of an allocation is the sum of the utilities of all agents from the allocation.

Let us now state two assumptions on distributions of utilities; in Section 4.3 we will work with the first and in Section 4.4 with the second.

[A1] For each good $g \in G$, the utilities $u_{ij}(g) \in [0, 1]$ for each agent $a_{ij} \in A$ are drawn independently at random from a distribution $\mathcal{D}_g$. Each distribution $\mathcal{D}_g$ is *non-atomic*, i.e., $\Pr[u_{ij}(g) = x] = 0$ for every $x \in [0, 1]$. Moreover, the variances of the distributions are bounded away from zero, i.e., there exists a constant $\sigma_{min} > 0$ such that the variance of $\mathcal{D}_g$ is at least σ^2_{min} for every good g.

[A2] For each agent $a_{ij} \in A$ and each good $g \in G$, the utility $u_{ij}(g) \in [0, 1]$ is drawn independently at random from a probability distribution $\mathcal{D}_{ij,g}$. The mean of each distribution is bounded away from zero, i.e., there exists a constant $\mu_{min} > 0$ such that $\mathbb{E}[u_{ij}(g)] \geq \mu_{min}$ for every $a_{ij} \in A$ and $g \in G$.

Note that assumption [A2] is weaker than [A1]. Indeed, in [A2] we do not require $\mathcal{D}_{ij,g}$ to be the same for every a_{ij}. In addition, since $u_{ij}(g) \in [0, 1]$ for all $a_{ij} \in A$ and $g \in G$, we have $\mathbb{E}[u_{ij}(g)] \geq \mathbb{E}[u_{ij}(g)^2] \geq \mathbb{E}[u_{ij}(g)^2] - \mathbb{E}[u_{ij}(g)]^2 = \mathrm{Var}(u_{ij}(g))$. Hence, the condition that the means of the distributions are bounded away from zero follows from the analogous condition on the variances.

In Section 4.4, we consider the notion of *approximate envy-freeness*, which means that for each agent, there is no bundle of another group for which the agent's utility is a certain (multiplicative) factor larger than the utility of the agent for the allocation of her own group. The notion is defined formally below.

Definition 4.2.1. *We write $G_p \succsim_{ij}^{\alpha} G_q$ for $\alpha \in [0,1]$ if and only if $u_{ij}(G_p) \geq \alpha u_{ij}(G_q)$. Agent a_{ij} considers an allocation $(G_1, G_2, \ldots, G_k)$ of goods to the k groups* α-approximate envy-free *if $G_i \succsim_i^{\alpha} G_{i'}$ for every group $A_{i'} \neq A_i$. We say that an allocation is* α-approximate envy-free *if it is α-approximate envy-free for every agent.*

Finally, we give the definition of a truthful mechanism, which we will use in Section 4.4.

Definition 4.2.2. *A* mechanism *is a function that takes as input the utility of agent a_{ij} for good g for all $a_{ij} \in A$ and $g \in G$, and outputs a (possibly random) allocation of goods to the groups. A mechanism is said to be* truthful *if every agent always obtains the highest possible (expected) utility by submitting her true utilities to the mechanism, regardless of the utilities that the remaining agents submit.*

4.3 Asymptotic Existence and Non-Existence of Fair Divisions

In this section, we study the existence and non-existence of fair divisions. First, we show that when $m \in \Omega(n \log n)$, where $\Omega(\cdot)$ hides a sufficiently large constant, there exists an envy-free division with high probability (Theorem 4.3.1). In particular, we prove that a welfare-maximizing allocation is likely to be envy-free. This gives rise to a simple algorithm that finds such a fair division with high probability. We also extend our existence result to the case where there are two groups but the groups need not have the same number of agents; we show a similar result in this case, provided that each distribution $\mathcal{D}_g$ satisfies an additional symmetry condition (Theorem 4.3.2).

Moreover, on the non-existence front, we prove that when m is smaller than n, the probability that a fair division exists is at most $1/k^{n-m}$ (Theorem 4.3.4). This has as consequences that if the number of goods is less than the total number of agents by a superconstant factor, or if the number of goods is less than the total number of agents and the number of groups is large, then the probability that an envy-free division exists is low (Corollaries 4.3.5 and 4.3.6).

4.3.1 Existence

We begin with our main existence result.

Theorem 4.3.1. *Assume that [A1] holds. For any fixed $\sigma_{min} > 0$, there exists a constant $C > 0$ such that, for any sufficiently large n', if $m > Cn \log n$, then there exists an envy-free allocation with high probability.*

In fact, we not only prove that an envy-free allocation exists but also give a simple greedy algorithm that finds one such allocation with high probability. The algorithm is simple: we greedily allocate each good to the group that maximizes the total utility of the good with respect to the agents in that group. This yields an allocation that maximizes the social welfare. The allocation is therefore Pareto optimal, i.e., there exists no other allocation in which every agent is weakly better off and at least one agent is strictly better off. The pseudocode of the algorithm is shown below.

Algorithm 1 Greedy Allocation Algorithm for Multiple Groups

```
procedure GREEDY–ALLOCATION–MULTIPLE
    let G_1 = G_2 = ··· = G_k = ∅.
    for each good g ∈ G do
        choose i* from arg max_{i=1,2,...,k} Σ_{j=1}^{n'} u_ij(g)
        let G_{i*} ← G_{i*} ∪ {g}
    end for
end procedure
```

The analysis of the algorithm contains similarities to that of the corresponding result in the setting with one agent per group [54]. However, significantly more technical care will be required to handle our setting in which each group contains multiple agents. This is reflected by our use of the Berry-Esseen theorem (Lemma 2.3.3). Here we provide a proof sketch that contains all the high-level ideas but leaves out some tedious details, especially calculations; the full proof can be found in Appendix B.1.

Proof sketch of Theorem 4.3.1. We will first bound $\Pr[u_{ij}(G_{i'}) > u_{ij}(G_i)]$ for each agent a_{ij} and each group $A_{i'} \neq A_i$; we then use the union bound at the end to conclude Theorem 4.3.1. To bound $\Pr[u_{ij}(G_{i'}) > u_{ij}(G_i)]$, we define a random variable $B_{ij,g}$ to be $u_{ij}(g)$ if good g is allocated to group A_i and zero otherwise. Similarly, define $C^{i'}_{ij,g}$ to be $u_{ij}(g)$ if the good is allocated to group $A_{i'}$ and zero otherwise.

Intuitively, with respect to agent a_{ij}, $B_{ij,g}$ is the utility contribution of good g to the group A_i. On the other hand, $C^{i'}_{ij,g}$ is the utility that is "lost" to group i'. In other words, $u_{ij}(G_{i'}) > u_{ij}(G_i)$ if and only if $S_B < S_C$, where $S_B = \sum_{g\in G} B_{ij,g}$ and $S_C = \sum_{g\in G} C^{i'}_{ij,g}$. We will use the Chernoff bound to estimate the probability of this event. To do so, we first need to bound $\mathbb{E}[B_{ij,g}]$ and $\mathbb{E}[C^{i'}_{ij,g}]$.

From the symmetry between different groups, the probability that good g is allocated to each group is $1/k$. Thus, we have $\mathbb{E}[B_{ij,g}] = \frac{1}{k}\,\mathbb{E}\left[u_{ij}(g) \mid \text{good } g \text{ is allocated to } A_i\right]$ and $\mathbb{E}[C^{i'}_{ij,g}] = \frac{1}{k}\,\mathbb{E}\left[u_{ij}(g) \mid \text{good } g \text{ is allocated to } A_{i'}\right]$. It is now fairly easy to see that $\mathbb{E}[C^{i'}_{ij,g}] \leq \mu_g/k$, where μ_g is the mean of $\mathcal{D}_g$; the reason is that the expected value of $u_{ij}(g)$ when g is not allocated to A_i is clearly at most μ_g. For convenience, we will assume in this proof sketch that $\mathbb{E}[C^{i'}_{ij,g}]$ is roughly μ_g/k.

Now, we will bound the expected value of $B_{ij,g}$. For each $p = 1, 2, \ldots, k$, let X_p denote the sum of the utilities of good g with respect to all agents in A_p. Due to the symmetry among agents within the same group, we have

$$\begin{aligned}\mathbb{E}[B_{ij,g}] &= \frac{1}{n'k}\,\mathbb{E}\left[X_i \mid X_i = \max\{X_1, X_2, \ldots, X_k\}\right] \\ &= \frac{1}{n'k}\,\mathbb{E}[\max\{X_1, X_2, \ldots, X_k\}].\end{aligned}$$

The latter equality comes from the symmetry between different groups.

Now, we use the Berry-Esseen theorem (Lemma 2.3.3), which tells us that each of $X_1, X_2, \ldots, X_k$ is close to $\mathcal{N}(\mu_g n', \Omega(\sigma_{min}^2 n'))$. With simple calculations, one can see that the expectation of the maximum of k identically independent random variables sampled from $\mathcal{N}(\mu_g n', \Omega(\sigma_{min}^2 n'))$ is $\mu_g n' + \Omega(\sigma_{min}\sqrt{n'})$. Roughly speaking, we also have

$$\mathbb{E}\left[B_{ij,g}\right] = \frac{\mu_g}{k} + \Omega\left(\frac{\sigma_{min}}{k\sqrt{n'}}\right).$$

Having bounded the expectations of $B_{ij,g}$ and $C_{ij,g}^{i'}$, we are ready to apply the Chernoff bound. Let $\delta = \Theta\left(\frac{\sigma_{min}}{\mu_g\sqrt{n'}}\right)$ where $\Theta(\cdot)$ hides some sufficiently small constant. When n' is sufficiently large, we can see that $(1+\delta)\,\mathbb{E}[C_{ij,g}^{i'}] < (1-\delta)\,\mathbb{E}[B_{ij,g}]$, which implies that $(1+\delta)\,\mathbb{E}[S_C] < (1-\delta)\,\mathbb{E}[S_B]$. Using the Chernoff bound (Lemma 2.3.1) on S_B and S_C, we have

$$\Pr[S_B \le (1-\delta)\,\mathbb{E}[S_B]] \le \exp\left(\frac{-\delta^2\,\mathbb{E}[S_B]}{2}\right),$$

and,

$$\Pr[S_C \ge (1+\delta)\,\mathbb{E}[S_C]] \le \exp\left(\frac{-\delta^2\,\mathbb{E}[S_C]}{3}\right).$$

Thus, we have

$$\begin{aligned}\Pr[S_B < S_C] &\le \exp\left(\frac{-\delta^2\,\mathbb{E}[S_B]}{2}\right) + \exp\left(\frac{-\delta^2\,\mathbb{E}[S_C]}{3}\right) \\ &\le 2\exp\left(-\Omega\left(\frac{\sigma_{min}^2 m}{n\mu_g}\right)\right) \\ (\text{Since } \mu_g \le 1) &\le 2\exp\left(-\Omega\left(\frac{\sigma_{min}^2 m}{n}\right)\right).\end{aligned}$$

Recall that $\Pr[u_{ij}(G_{i'}) > u_{ij}(G_i)] = \Pr[S_B < S_C]$. Using the union bound for all a_{ij} and all $A_{i'} \neq A_i$, the probability that the allocation output by the algorithm is not envy-free is at most

$$2n(k-1)\exp\left(-\Omega\left(\frac{\sigma_{min}^2 m}{n}\right)\right),$$

which is at most $1/m$ when $m \ge Cn\log n$ for some sufficiently large C. This completes the proof sketch of the theorem. □

Unfortunately, the algorithm in Theorem 4.3.1 cannot be extended to give a proof for the case where the groups do not have the same number of agents. However, in a more restricted setting where there are only two groups with potentially different numbers of agents and an additional symmetry condition on the distributions $\mathcal{D}_g$ is enforced, a result similar to that in Theorem 4.3.1 can be shown, as stated in the theorem below.

Theorem 4.3.2. *Assume that [A1] holds. Suppose that there are only two groups but not necessarily with the same number of agents; let n_1, n_2 denote the numbers of agents of the first and second group respectively (so $n = n_1 + n_2$). Assume also that the distributions $\mathcal{D}_g$ are symmetric (around 1/2),*[2] *i.e.,*

$$\Pr_{X \sim \mathcal{D}_g}\left[X \leq \frac{1}{2} - x\right] = \Pr_{X \sim \mathcal{D}_g}\left[X \geq \frac{1}{2} + x\right]$$

for all $x \in [0, 1/2]$. For any fixed $\sigma_{min} > 0$, there exists a constant $C > 0$ such that for any sufficiently large n_1 and n_2, if $m > Cn \log n$, then there exists an envy-free allocation with high probability.

The algorithm is similar to that in Theorem 4.3.1; the only difference is that, instead of allocating each good to the group with the highest *total* utility over its agents, we allocate the good to the group with the highest *average* utility, as seen in the pseudocode of Algorithm 2.

Algorithm 2 Greedy Allocation Algorithm for Two Possibly Unequal-Sized Groups

```
procedure GREEDY–ALLOCATION–TWO
    let G_1 = G_2 = ∅.
    for each good g ∈ G do
        choose i* from argmax_{i=1,2} (Σ_{j=1}^{n_i} u_ij(g)) / n_i
        let G_{i*} ← G_{i*} ∪ {g}
    end for
end procedure
```

The proof is essentially the same as that of Theorem 4.3.1 after the random variables defined are changed corresponding to the modification in the algorithm. For instance, $B_{ij,g}$ is now defined as

$$B_{ij,g} = u_{ij}(g) \cdot \mathbf{1}\left[i = \arg\max_{q=1,2} \frac{\sum_{p=1}^{n_q} u_{qp}(g)}{n_q}\right]$$

where $\mathbf{1}[E]$ denotes an indicator variable for event E.

Due to the similarities between the two proofs, we will not repeat the whole proof. Instead, we would like to point out that all the arguments from Theorem 4.3.1 work here save for only one

[2]There is nothing special about the number 1/2; a similar result holds if the distributions are supported on a subset of an interval $[a, b]$ and are symmetric around $(a + b)/2$, for some $0 < a < b$.

additional fact that we need to prove:

Lemma 4.3.3. *Let X_1 and X_2 denote $\sum_{p=1}^{n_1} u_{qp}(g)/n_1$ and $\sum_{p=1}^{n_2} u_{qp}(g)/n_2$ respectively. Then,*

$$\Pr[X_1 \geq X_2] = \frac{1}{2}.$$

Proof. To show this, observe first that, since $\mathcal{D}_g$ is symmetric over 1/2, the distributions of X_1 and X_2 are also symmetric over 1/2. Let f_1 and f_2 be the probability density functions of X_1 and X_2 respectively, we have

$$\begin{aligned}\Pr[X_1 \geq X_2] &= \int_0^1 \int_0^x f_1(x)f_2(y)dydx \\ &= \int_0^1 \int_0^x f_1(1-x)f_2(1-y)dydx \\ &= \int_0^1 \int_x^1 f_1(x)f_2(y)dydx \\ &= \Pr[X_2 \geq X_1].\end{aligned}$$

Hence, $\Pr[X_1 \geq X_2] = \Pr[X_2 \geq X_1] = 1/2$, as desired. □

4.3.2 Non-Existence

Next, we state and prove an upper bound for the probability that an envy-free allocation exists when the number of agents exceeds the number of goods. Such an allocation obviously does not exist under this condition if every group contains only one agent. In fact, the theorem holds even without the assumption that the variances of the distributions $\mathcal{D}_g$ are at least $\sigma_{min}^2 > 0$.

Theorem 4.3.4. *Assume that [A1] holds. If $m < n$, then there exists an envy-free allocation with probability at most $1/k^{n-m}$.*

Proof. Suppose that $m \leq n-1$, and fix an allocation $G_1, \dots, G_k$. We will bound the probability that this allocation is envy-free. Consider any agent A_{ij}. The probability that the allocation is envy-free for this particular agent is the probability that the total utility of the agent for the bundle G_i is no less than that for any other bundle $G_{i'}$. This can be written as follows:

$$\Pr_{u_{ij}(g) \in \mathcal{D}_g \text{ for all } g \in G}\left[\sum_{g \in G_i} u_{ij}(g) = \max_{i'=1,2,\dots,k} \sum_{g \in G_{i'}} u_{ij}(g)\right].$$

For each $q = 1, 2, \ldots, k$, define p_q as

$$p_q = \Pr_{\substack{x_g \in \mathcal{D}_g \\ \text{for all } g \in G}} \left[\sum_{g \in G_q} x_g = \max_{q'=1,2,\ldots,k} \sum_{g \in G_{q'}} x_g \right].$$

Notice that the probability that the allocation is envy-free for agent a_{ij} is p_i.

Since the utilities of different agents are chosen independently from one another, the probability that this allocation is envy-free for every agent is simply the product of the probability that the allocation is envy-free for each agent, i.e.,

$$\prod_{a_{ij} \in A}^{n} p_i = \prod_{q=1}^{k} p_q^{n'}.$$

Using the inequality of arithmetic and geometric means, we arrive at the following bound:

$$\prod_{q=1}^{k} p_q^{n'} \leq \left(\frac{1}{k} \sum_{q=1}^{k} p_q \right)^{n'k}.$$

Recall our assumption that the distributions $\mathcal{D}_g$ are non-atomic. Hence we may assume that the events $\sum_{g \in G_q} x_g = \max_{q'=1,2,\ldots,k} \sum_{g \in G_{q'}} x_g$ are disjoint for different j. This implies that $\sum_{q=1}^{k} p_q = 1$. Thus, the probability that this fixed allocation is envy-free is at most

$$\left(\frac{1}{k} \sum_{q=1}^{k} p_q \right)^{n'k} = \left(\frac{1}{k} \right)^{n'k} = \frac{1}{k^n}.$$

Finally, since each allocation is envy-free with probability at most $1/k^n$ and there are k^m possible allocations, by union bound the probability that there exists an envy-free allocation is at most $1/k^{n-m}$. This completes the proof of the theorem. □

The following corollaries can be immediately derived from Theorem 4.3.4. They say that an envy-free allocation is unlikely to exist when the number of goods is less than the number of agents by a superconstant factor, or when the number of goods is less than the number of agents and the number of groups is large.

Corollary 4.3.5. *Assume that [A1] holds. When $m = n - \omega(1)$, the probability that there exists an envy-free allocation converges to zero as $n \to \infty$.*

Corollary 4.3.6. *Assume that [A1] holds. When $m < n$, the probability that there exists an envy-free allocation converges to zero as $k \to \infty$.*

4.4 Truthful Mechanism for Approximate Envy-Freeness

While the algorithms in Section 4.3 translate to mechanisms that yield with high probability envy-free divisions that are compatible with social welfare assuming that agents are truth-telling, the resulting mechanisms suffer from the setback that they are easily manipulable. Indeed, since they aim to maximize (total or average) welfare, strategic agents will declare their values for the goods to be high, regardless of what the actual values are. This presents a significant disadvantage: implementing these mechanisms in most practical situations, where we do not know the true valuations of the agents and have no reason to assume that they will reveal their valuations in a honest manner, can lead to potentially undesirable outcomes.

In this section, we work with the weaker notion of approximate envy-freeness and show that a simple truthful mechanism yields an approximately envy-free allocation with high probability. In particular, we prove that the random allocation mechanism, which allocates each good to an agent chosen uniformly and independently at random, is likely to produce such an allocation. In the setting where each group consists of only one agent, Amanatidis et al. [6] showed that when the distribution is as above and the number of goods m is large enough compared to n, the random allocation mechanism yields an approximately envy-free allocation with high probability. Our statement is an analogous statement for the case where each group can have multiple agents.

Theorem 4.4.1. *Assume that [A2] holds. For every $\alpha \in [0, 1)$, there exists a constant C depending only on α and μ_{min} such that if $m > Ck \log n$, then the random allocation, where each good $g \in G$ is allocated independently and uniformly at random to a group, is α-approximate envy-free with high probability.*

Before we prove Theorem 4.4.1, we note some ways in which our result is stronger than that of Amanatidis et al.'s apart from the fact that multiple agents per group are allowed in our setting. First, Amanatidis et al. required $\mathcal{D}_{ij,g}$ to be the same for all g, which we do not assume here. Next, they only showed that the random allocation is likely to be *approximately proportional*, a weaker notion that is implied by approximate envy-freeness. Moreover, in their result, m needs to be as high as $\Omega(n^2)$, whereas in our case it suffices for m to be in the range $\Omega(k \log n)$. Finally, we also derive a stronger probabilistic bound; they showed a "success probability" of the algorithm of $1 - O(n^2/m)$, while our success probability is $1 - \exp(-\Omega(m/k))$.

Proof of Theorem 4.4.1. For each agent $a_{ij} \in A$, each good $g \in G$ and each $q \in \{1, 2, \ldots, k\}$, let $B^q_{ij,g}$ be a random variable representing the contribution of good g's utility with respect to agent a_{ij} to group A_q, i.e., $B^q_{ij,g}$ is $u_{ij}(g)$ if good g is allocated to group A_q and is zero otherwise.

Define $S^q_{ij} := \sum_{g \in G} A^q_{ij,g}$. Observe that each agent a_{ij} considers the allocation to be α-approximate envy-free if and only if $S^i_{ij} \geq \alpha S^q_{ij}$ for every q. Let $\delta = \frac{1-\alpha}{1+\alpha}$; from this choice of δ and since $\mathbb{E}[S^q_{ij}]$ is equal for every q, we can conclude that $S^i_{ij} \geq \alpha S^q_{ij}$ is implied by $S^i_{ij} \geq (1 - \delta)\, \mathbb{E}[S^i_{ij}]$

and $S_{ij}^q \leq (1+\delta)\,\mathbb{E}[S_{ij}^q]$. In other words, we can bound the probability that the random allocation is not α-approximate envy-free as follows.

$$\begin{aligned}
&\Pr[\exists a_{ij} \in A, q \in \{1,2,\ldots,k\} : S_{ij}^i < \alpha S_{ij}^q] \\
&\quad \leq \sum_{a_{ij}\in A, q\in\{1,2,\ldots,k\}} \Pr[S_{ij}^i < \alpha S_{ij}^q] \\
&\quad \leq \sum_{a_{ij}\in A, q\in\{1,2,\ldots,k\}} \Pr[S_{ij}^i < (1-\delta)\,\mathbb{E}[S_{ij}^i] \text{ or } S_{ij}^q > (1+\delta)\,\mathbb{E}[S_{ij}^i]] \\
&\quad \leq \sum_{a_{ij}\in A, q\in\{1,2,\ldots,k\}} (\Pr[S_{ij}^i < (1-\delta)\,\mathbb{E}[S_{ij}^i]] + \Pr[S_{ij}^q > (1+\delta)\,\mathbb{E}[S_{ij}^q]]).
\end{aligned}$$

Since $S_{ij}^q = \sum_{g\in G} B_{ij,g}^q$ and the random variables $B_{ij,g}^q$ are independent and lie in $[0,1]$, we can use the Chernoff bound (Lemma 2.3.1) to upper bound the last terms. Hence, the probability that the allocation is not α-approximate envy-free is at most

$$\sum_{a_{ij}\in A, q\in\{1,2,\ldots,k\}} \exp\left(\frac{-\delta^2\,\mathbb{E}[S_{ij}^i]]}{2}\right) + \exp\left(\frac{-\delta^2\,\mathbb{E}[S_{ij}^q]]}{3}\right).$$

Finally, observe that

$$\mathbb{E}[S_{ij}^q] = \sum_{g\in G} \mathbb{E}[B_{ij,g}^q] = \sum_{g\in G} \frac{1}{k}\,\mathbb{E}[u_{ij}(g)] \geq \frac{m\mu_{min}}{k}.$$

This means that the desired probability is bounded above by

$$\begin{aligned}
&\sum_{a_{ij}\in A, q\in\{1,2,\ldots,k\}} \exp\left(\frac{-\delta^2 m\mu_{min}}{2k}\right) + \exp\left(\frac{-\delta^2 m\mu_{min}}{3k}\right) \\
&\quad \leq 2nk \exp\left(\frac{-\delta^2 m\mu_{min}}{3k}\right) \\
&\quad \leq \exp\left(-\frac{\delta^2 m\mu_{min}}{3k} + 3\log n\right).
\end{aligned}$$

When $m > \left(\frac{10}{\mu_{min}\delta^2}\right) k \log n$, the above expression is at most $\exp(-\Omega(m/k))$, concluding our proof. □

4.5 Conclusion and Future Work

In this chapter, we study the group fair division setting and establish nearly tight bounds on the number of agents and goods under which a fair division is likely or unlikely to exist. Furthermore, we consider the issue of truthfulness and show that a simple truthful mechanism produces an allocation that is approximately envy-free with high probability.

While the assumptions of additivity and independence are somewhat restrictive and might not apply fully to settings in the real world, our results give indications as to what we can expect if the assumptions are relaxed, such as if a certain degree of dependence is introduced. An interesting future direction is to generalize the results to settings with more general valuations. In particular, if the utility functions are low-degree polynomials, then one could try applying the invariance principle [117], which is a generalization of the Berry-Esseen theorem that we use.

We end the chapter with some further questions that remain after this work. A natural question is whether we can generalize our existence and non-existence results (Theorems 4.3.1 and 4.3.4) to the setting where the groups do not contain the same number of agents. This non-symmetry between the groups seems to complicate the approaches that we use in this chapter. For example, it breaks the greedy algorithm used in Theorem 4.3.1. Nevertheless, it might still be possible to prove existence of an envy-free division using other algorithms or without relying on a specific algorithm.

Another direction for future research is to invent procedures for computing envy-free divisions, whenever such divisions exist, for the general setting where each group contains multiple agents and agents have arbitrary monotonic valuations. Even procedures that only depend on rankings of single goods [28] do not appear to extend easily to this setting. Indeed, if a group contains two agents whose preferences are opposite of each other, it is not immediately clear what we should allocate to the group. It would be useful to have a procedure that produces a desirable outcome, even for a small number of agents in each group.

Lastly, one could explore the limitations that arise when we impose the condition of truthfulness, an important property when we implement the mechanisms in practice. For instance, truthful allocation mechanisms have recently been characterized in the case of two agents [5], and it has been shown that there is a separation between truthful and non-truthful mechanisms for approximating maximin shares [6]. In our setting, a negative result on the existence of a truthful mechanism that yields an envy-free division with high probability would provide such a separation as well, while a positive result in this direction would have even more consequences for practical applications.

Chapter 5

Approximate Maximin Shares for Groups of Agents

5.1 Introduction

The previous chapter shows an approach that uses asymptotic analysis to circumvent the potential non-existence of fair divisions, which leads to statements on the probability that a fair allocation exists as the size of the instance grows. In this chapter, we present a different approach to obtain positive results using worst-case analysis and approximation. In particular, we study the existence of allocations satisfying the maximin share criteria introduced in Section 2.1.

In Section 5.3, we consider the setting where there are two groups of agents. For this setting, we completely determine the cardinality of agents in the groups for which it is possible to approximate the maximin share within a positive factor that depends only on the number of agents and not on the number of goods. In particular, an approximation is possible when one of the groups contain a single agent, when both groups contain two agents, or when the groups contain three and two agents respectively. In all other cases, no approximation is possible in a strong sense: There exists an instance with only four goods in which some agent with positive maximin share necessarily gets zero utility. These results, along with bounds for the approximation factors, are summarized in Table 5.1.

In Section 5.4, we generalize our results to the setting with several groups of agents. On the positive side, we show that a positive approximation is possible if only one group contains more than a single agent (Theorem 5.4.1). On the other hand, we show on the negative side that when all groups contain at least two agents and one group contains at least five agents, it is possible that some agent with positive maximin share will be forced to obtain zero utility (Theorem 5.4.2), which means that there is no hope of obtaining an approximation in this case.

Number of agents	Approximation ratio
$(n_1, n_2) = (1, 1)$	$\alpha = 1$ (Cut-and-choose protocol; see, e.g., [26])
$(n_1, n_2) = (2, 1)$	$2/3 \leq \alpha \leq 3/4$ (Theorem 5.3.3)
$n_2 = 1$	$2/(n_1 + 1) \leq \alpha \leq 1/\left\lfloor \left\lfloor \sqrt{2n_1} \right\rfloor /2 \right\rfloor$ (Corollary 5.3.5)
$(n_1, n_2) = (2, 2)$	$1/8 \leq \alpha \leq 1/2$ (Theorem 5.3.6)
$(n_1, n_2) = (3, 2)$	$1/16 \leq \alpha \leq 1/2$ (Theorem 5.3.7)
$n_1 \geq 4, n_2 \geq 2$	$\alpha = 0$ (Proposition 5.3.1)
$n_1, n_2 \geq 3$	$\alpha = 0$ (Proposition 5.3.2)

Table 5.1: Values of the best possible approximation ratio, denoted by α, for the maximin share when there are two groups with $n_1 \geq n_2$ agents. The approximation ratios hold regardless of the number of goods.

5.2 Preliminaries

The basic definitions and notation of the group fair division setting are introduced in Section 2.1. We assume in this chapter that agents have additive utility functions. Let us state two lemmas that we will use in the analysis of the modified round-robin algorithm in Section 5.3.2.

Lemma 5.2.1 ([7])**.** *Suppose that each group contains one agent. Consider a round-robin algorithm in which the agents take turns taking their favorite good from the remaining goods; if there are goods remaining after the last agent takes a good, we circle back to the first agent. In the resulting allocation, the envy that an agent has toward any other agent is at most the maximum utility of the former agent for any single good. Moreover, if an agent is ahead of another agent in the round-robin ordering, then the former agent has no envy toward the latter agent.*

Lemma 5.2.2 ([7, 26])**.** *Given an arbitrary instance in which each group contains one agent, if we allocate an arbitrary good to an agent as her only good, then the maximin share of any remaining agent with respect to the remaining goods does not decrease.*

Both of these lemmas admit rather straightforward proofs which can also be found in the cited references.

5.3 Two Groups of Agents

In this section, we consider the setting where there are two groups of agents and characterize the cardinality of the groups for which a positive approximation of the maximin share is possible regardless of the number of goods. In particular, suppose that the two groups contain n_1 and n_2 agents, where we assume without loss of generality that $n_1 \geq n_2$. Then a positive approximation is possible when $n_2 = 1$ as well as when $(n_1, n_2) = (2, 2)$ or $(3, 2)$. The results are summarized in Table 5.1.

5.3.1 Large Number of Agents: No Possible Approximation

We begin by showing that when the numbers of agents in the groups are large enough, no approximation of the maximin share is possible. Observe that if we prove that a maximin share approximation is not possible for groups with n_1 and n_2 agents, then it is also not possible for groups with $n_1' \geq n_1$ and $n_2' \geq n_2$ agents, since we would still need to fulfill the approximation for the first n_1 and n_2 agents in the respective groups.

Proposition 5.3.1. *If $n_1 \geq 4$ and $n_2 \geq 2$, then there exists an instance in which some agent with nonzero maximin share necessarily receives zero utility.*

Proof. Assume that $n_1 = 4$ and $n_2 = 2$, and suppose that there are four goods. The utilities of the agents in the first group are $\mathbf{u}_{11} = (0,1,0,1)$, $\mathbf{u}_{12} = (0,1,1,0)$, $\mathbf{u}_{13} = (1,0,0,1)$, and $\mathbf{u}_{14} = (1,0,1,0)$, while the utilities of those in the second group are $\mathbf{u}_{21} = (1,1,0,0)$ and $\mathbf{u}_{22} = (0,0,1,1)$.

In this example, every agent has a maximin share of 1. To guarantee nonzero utility for the agents in the second group, we must allocate at least one of the first two goods and at least one of the last two goods to the group. But this implies that some agent in the first group receives zero utility. □

Proposition 5.3.2. *If $n_1, n_2 \geq 3$, then there exists an instance in which some agent with nonzero maximin share necessarily receives zero utility.*

Proof. Assume that $n_1 = n_2 = 3$, and suppose that there are three goods. The utilities of the agents in the both groups are $\mathbf{u}_{i1} = (1,1,0)$, $\mathbf{u}_{i2} = (1,0,1)$, and $\mathbf{u}_{i3} = (0,1,1)$ for $i = 1,2$.

In this example, every agent has a maximin share of 1. In any allocation, one of the groups gets at most one good, and some agent in that group receives zero utility. □

5.3.2 Approximation via Modified Round-Robin Algorithm

When both groups contain a single agent, it is known that a simple "cut-and-choose" protocol similar to a famous cake-cutting protocol yields the full maximin share for both agents (see, e.g., [26]). It turns out that as soon as at least one group contains more than one agent, the full maximin share can no longer be guaranteed. We next consider the simplest such case where the groups contain one and two agents, respectively. The maximin share approximation algorithm for this case is similar to the modified round-robin algorithm that yields a 1/2-approximation for an arbitrary number of agents [7], but we will need to make some adjustments to handle more than one agent being in the same group.

Theorem 5.3.3. *Let $(n_1, n_2) = (2,1)$, and suppose that α is the best possible approximation ratio for the maximin share. Then $2/3 \leq \alpha \leq 3/4$.*

Proof. We first show the upper bound. Suppose that there are four goods. The utilities of the agents in the first group for the goods are $\mathbf{u}_{11} = (3, 1, 2, 2)$ and $\mathbf{u}_{12} = (2, 3, 2, 1)$, while the utilities of the agent in the second group are $\mathbf{u}_{21} = (3, 2, 2, 1)$.

In this example, every agent has a maximin share of 4. We will show that any allocation gives some agent a utility of at most 3. Note that an allocation that would give every agent a utility of at least 4 must allocate two goods to both groups. If the fourth and one of the second and third goods are allocated to the second group, the agent gets a utility of 3. Otherwise, one can check that one of the agents in the first group gets a utility of 3.

Next, we exhibit an algorithm that guarantees each agent a 2/3 fraction of her maximin share. Since we do not engage in interpersonal comparisons of utilities, we may assume without loss of generality that every agent has utility 1 for the whole bundle of goods. Since the maximin share of an agent is always at most 1/2, it suffices to allocate a bundle worth at least 1/3 to her.

If some good is worth at least 1/3 to a_{21}, let her take that good. By Lemma 5.2.2, the maximin shares of a_{11} and a_{12} do not decrease. Since they receive all of the remaining goods, they obtain their maximin share. Hence we may now assume that no good is worth at least 1/3 to a_{21}. We let each of a_{11} and a_{12}, in arbitrary order, take a good worth at least 1/3 if there is any. There are three cases.

- Both of them take a good. Then each of them gets a utility of at least 1/3. Since every good is worth less than 1/3 to a_{21}, she also gets a utility of at least $1 - 1/3 - 1/3 = 1/3$ from the remaining goods.

- One of them takes a good and the other does not. Assume without loss of generality that a_{11} is the agent who takes a good. We run the round-robin algorithm on a_{12} and a_{21} using the remaining goods, starting with a_{21}. Since every good is worth less than 1/3 to a_{21}, the value of the whole bundle of goods minus the good that a_{11} takes is at least 2/3. By Lemma 5.2.1, a_{21} gets a utility of at least $1/2 \times 2/3 = 1/3$. Similarly, the envy of a_{12} toward a_{21} is at most the maximum utility of a_{12} for a good allocated during the round-robin algorithm, which is at most 1/3. This implies that a_{21}'s bundle is worth at most 2/3 to a_{12}, and hence a_{12}'s bundle in the final allocation (i.e., her bundle from the round-robin algorithm combined with the good that a_{11} takes) is worth at least 1/3 to her.

- Neither of them takes a good. We run the round-robin algorithm on all three agents, starting with a_{21}. By Lemma 5.2.1, a_{21} gets a utility of at least 1/3. The envy of a_{11} toward a_{21} is at most the maximum utility of a_{11} for a good, which is at most 1/3. Hence a_{21}'s bundle is worth at most 2/3 to a_{11}, which means that a_{11}'s bundle in the final allocation (i.e., her bundle combined with a_{12}'s bundle) is worth at least 1/3 to her. An analogous argument holds for a_{12}.

This covers all three possible cases. □

Next, we generalize to the setting where the first group contains an arbitrary number of agents while the second group contains a single agent. In this case, an algorithm similar to that in Theorem 5.3.3 can be used to obtain a constant factor approximation when the number of agents is constant. In addition, we show that the approximation ratio necessarily degrades as the number of agents grows.

Algorithm 3 Algorithm for approximate maximin share when the groups contain $n_1 \geq 2$ and $n_2 = 1$ agents (Theorem 5.3.4).

procedure APPROXIMATE–MAXIMIN–SHARE–1
  if Agent a_{21} values some good g at least $\frac{1}{n_1+1}$ **then**
    Allocate g to a_{21} and the remaining goods to the first group.
  else
    Let each agent in the first group, in arbitrary order, take a good worth at least $\frac{1}{n_1+1}$ to her if there is any.
    Allocate the remaining goods to the agents who have not taken a good using the round-robin algorithm, starting with a_{21}.
  end if
end procedure

Theorem 5.3.4. *Let $n_1 \geq 2$ and $n_2 = 1$, and suppose that α is the best approximation ratio for the maximin share. Then $\frac{2}{n_1+1} \leq \alpha \leq \frac{1}{\lfloor f(n_1)/2 \rfloor}$, where $f(n_1)$ is the largest integer such that $\binom{f(n_1)}{2} \leq n_1$.*

Proof. We first show the upper bound. Let $l = f(n_1)$, and suppose that there are l goods. Let $\binom{l}{2}$ of the agents in the first group positively value a distinct set of two goods. In particular, each of them has utility 1 for both goods in their set and 0 for the remaining goods. Let the agent in the second group have utility 1 for all goods.

In this example, each agent in the first group has a maximin share of 1, while the agent a_{21} has a maximin share of $\lfloor l/2 \rfloor$. To guarantee nonzero utility for the l agents in the first group, we must allocate all but at most one good to the group, leaving at most one good for the second group. So the agent in the second group obtains at most a $1/\lfloor l/2 \rfloor$ fraction of her maximin share.

An algorithm that guarantees a $\frac{2}{n_1+1}$-approximation of the maximin share (Algorithm 3) is similar to that for the case $n_1 = 2$ (Theorem 5.3.3). Again, we normalize the utility of each agent for the whole set of goods to 1. First, let a_{21} take a good worth at least $\frac{1}{n_1+1}$ to her if there is any. If she takes a good, we allocate the remaining goods to the first group and are done by Lemma 5.2.2. Else, we let each of the agents in the first group, in arbitrary order, take a good worth at least $\frac{1}{n_1+1}$ to her if there is any. After that, we run the round-robin algorithm on the agents who have not taken a good, starting with a_{21}.

Suppose that r agents in the first group take a good. Each of them obtains a utility of at least $\frac{1}{n_1+1}$. The remaining goods, which are allocated by the round-robin algorithm, are worth a total of at least $\frac{n_1+1-r}{n_1+1}$ to a_{21}. Since there are n_1+1-r agents who participate in the round-robin algorithm, and a_{21} is the first to choose, she obtains utility at least $\frac{1}{n_1+1-r} \cdot \frac{n_1+1-r}{n_1+1} = \frac{1}{n_1+1}$. Finally, by Lemma 5.2.1, each agent in the first group who does not take a good in the first stage has envy at most $\frac{1}{n_1+1} \leq \frac{1}{3}$ toward a_{21}. Hence for such an agent, a_{21}'s bundle is worth at most 2/3, and so the bundle allocated to the first group is worth at least $\frac{1}{3} \geq \frac{1}{n_1+1}$. □

Algorithm 3 can be implemented in time polynomial in the number of agents and goods. Also, since $\binom{\lfloor\sqrt{2n_1}\rfloor}{2} \leq \frac{(\sqrt{2n_1})^2}{2} = n_1$, we have the following corollary.

Corollary 5.3.5. *Let $n_1 \geq 2$ and $n_2 = 1$, and suppose that α is the best approximation ratio for the maximin share. Then* $\frac{2}{n_1+1} \leq \alpha \leq \frac{1}{\lceil\lfloor\sqrt{2n_1}\rfloor/2\rceil}$.

5.3.3 Approximation via Maximin Partitions

We now consider the two remaining cases, $(n_1, n_2) = (2, 2)$ and $(3, 2)$. We show that in both cases, a positive approximation is also possible. However, the algorithms for these two cases will rely on a different idea than the previous algorithms. These positive results provide a clear distinction between the settings where it is possible to approximate the maximin share and those where it is not. For the former settings, the maximin share can be approximated within a positive factor independent of the number of goods. On the other hand, for the latter settings, there exist instances in which some agent with positive maximin share necessarily gets zero utility even when there are only four goods (Propositions 5.3.1 and 5.3.2), and therefore no approximation is possible even if we allow dependence on the number of goods.

Algorithm 4 Algorithm for approximate maximin share when the groups contain $n_1 = n_2 = 2$ agents (Theorem 5.3.6).

procedure APPROXIMATE–MAXIMIN–SHARE–2
  for each agent a_{ij} **do**
    Compute her maximin partition $(G_{ij}, G \backslash G_{ij})$.
  end for
  Partition G into 16 subsets $(H_1, H_2, \ldots, H_{16})$ according to whether each good belongs to G_{ij} or $G\backslash G_{ij}$ for each $1 \leq i, j \leq 2$.
  if Some subset H_p is *important* (i.e., of value at least 1/8 of the agent's maximin share) to both a_{11} and a_{12} **then**
    Allocate H_p to the first group and the remaining goods to the second group.
  else
    Suppose that H_p, H_q are important to a_{11} and H_r, H_s to a_{12}.
    Find a pair from $(H_p, H_r), (H_p, H_s), (H_q, H_r), (H_q, H_s)$ that does not coincide with the important subsets for an agent in the second group.
    Allocate that pair of subsets to the first group and the remaining goods to the second group.
  end if
end procedure

Theorem 5.3.6. *Let $(n_1, n_2) = (2, 2)$, and suppose that α is the best possible approximation ratio for the maximin share. Then $1/8 \leq \alpha \leq 1/2$.*

Proof. We first show the upper bound. Suppose that there are four goods. The utilities of the agents in the first group for the goods are $\mathbf{u}_{11} = (0, 2, 1, 1)$ and $\mathbf{u}_{12} = (2, 0, 1, 1)$, while the utilities of those in the second group are $\mathbf{u}_{21} = (1, 1, 0, 0)$ and $\mathbf{u}_{22} = (0, 0, 1, 1)$.

In this example, both agents in the first group have a maximin share of 2, and both agents in the second group have a maximin share of 1. To ensure nonzero utility for the agents in the second group, we must allocate at least one of the first two goods and at least one of the last two goods to the group. However, this implies that some agent in the first group gets a utility of at most 1.

Next, we exhibit an algorithm that yields a 1/8-approximation of the maximin share (Algorithm 4). For each agent a_{ij}, let $(G_{ij}, G\backslash G_{ij})$ be one of her maximin partitions. By definition, we have that both $u_{ij}(G_{ij})$ and $u_{ij}(G\backslash G_{ij})$ are at least the maximin share of a_{ij}. Let $(H_1, H_2, \dots, H_{16})$ be a partition of G into 16 subsets according to whether each good belongs to G_{ij} or $G\backslash G_{ij}$ for each $1 \leq i, j \leq 2$; in other words, for every i, j, the goods in each set H_k either all belong to G_{ij} or all belong to $G\backslash G_{ij}$. By the pigeonhole principle, for each agent a_{ij}, among the eight sets H_k whose union is G_{ij}, the set that she values most gives her a utility of at least 1/8 of her maximin share; call this set H_p. Likewise, we can find a set $H_q \subseteq G\backslash G_{ij}$ that the agent values at least 1/8 of her maximin share. We call these subsets *important* to a_{ij}. It suffices for every agent to obtain a subset that is important to her.

If some subset H_p is important to both a_{11} and a_{12}, we allocate that subset to the first group and the remaining goods to the second group. Since each agent in the second group has at least two important subsets, and only one is taken away from them, this yields the desired guarantee. Else, two subsets H_p, H_q are important to a_{11} and two other subsets H_r, H_s are important to a_{12}. We will assign one of the pairs $(H_p, H_r), (H_p, H_s), (H_q, H_r), (H_q, H_s)$ to the first group. If a pair does not work, that means that some agent in the second group has exactly that pair as her important subsets. But there are four pairs and only two agents in the second group, hence some pair must work. □

We briefly discuss the running time of Algorithm 4. The algorithm can be implemented efficiently except for one step: computing a maximin partition of each agent. This step is NP-hard even when the two agents have identical utility functions by a straightforward reduction from the partition problem. Nevertheless, Woeginger [160] showed that a PTAS for the problem exists.[1] Using the PTAS, we can compute an approximate maximin partition instead of an exact one and obtain a $(1/8 - \epsilon)$-approximate algorithm for the maximin share in time polynomial in the number of goods for any constant $\epsilon > 0$.

A similar idea can be used to show that a positive approximation of the maximin share is possible when $(n_1, n_2) = (3, 2)$.

Theorem 5.3.7. *Let $(n_1, n_2) = (3, 2)$, and suppose that α is the best possible approximation ratio for the maximin share. Then $1/16 \leq \alpha \leq 1/2$.*

[1] Woeginger also showed that an FPTAS for this problem does not exist unless P = NP.

Proof. The upper bound follows from Theorem 5.3.6 and the observation preceding Proposition 5.3.1.

For the lower bound, compute the maximin partition for each agent, and partition G into 32 subsets according to which part of the partition of each agent a good belongs to. For each agent, at least 2 of the subsets are *important*, i.e., of value at least 1/16 of her maximin share. If some subset is important to both a_{21} and a_{22}, allocate that subset to them and the remaining goods to the first group. Otherwise, we can allocate some pair of important subsets to the second group using a similar argument as in Theorem 5.3.6. □

5.3.4 Experimental Results

To complement our theoretical results, we ran computer experiments to see the extent to which it is possible to approximate the maximin share in random instances. For $(n_1, n_2) = (2, 2)$ and $(3, 2)$, we generated 100000 random instances where there are four goods and the utility of each agent for each good is drawn independently and uniformly at random from the interval $[0, 1]$. The results are shown in Table 5.2a. An approximation ratio of 0.9 can be guaranteed in over 90% when there are two agents in each group, and in over 80% when there are three agents in one group and two in the other. In other words, an allocation that "almost" satisfies the maximin criterion can be found in a large majority of the instances. However, the proportion drops significantly to around 70% and 50% if we demand that the partition yield the full maximin share to the agents, indicating that this is a much more stringent requirement. We also ran the experiment on instances where the utilities are drawn from an exponential distribution and from a log-normal distribution. As shown in Tables 5.2b and 5.2c, the number of instances for which the (approximate) maximin criterion is satisfied is lower for both distributions than for the uniform distribution. This is to be expected since the utilities are less spread out, meaning that conflicts are more likely to occur. The heavy drop as we increase the requirement from $\alpha \geq 0.9$ to $\alpha \geq 1$ is present for these distributions as well.

We also remark here that the case with four goods seems to be the hardest case for maximin approximation. Indeed, with two goods an allocation yielding the full maximin share always exists, with three goods the maximin share is low since any partition leaves at most one good to one group, and with more than four goods there are more allocations and therefore more possibilities to exploit the differences between the utilities of the agents.

5.4 Several Groups of Agents

In this section, we consider a more general setting where there are several groups of agents. We show that when only one group contains more than a single agent, a positive approximation is possible independent of the number of goods. On the other hand, when all groups contain at least two agents and one group contains at least five agents, no approximation is possible in a strong sense.

	$\alpha \geq 0.5$	$\alpha \geq 0.6$	$\alpha \geq 0.7$	$\alpha \geq 0.8$	$\alpha \geq 0.9$	$\alpha \geq 1$
$(n_1, n_2) = (2, 2)$	100000	100000	99937	98803	92015	69248
$(n_1, n_2) = (3, 2)$	100000	99997	99672	96174	81709	49386

(a) The uniform distribution over the interval $[0, 1]$

	$\alpha \geq 0.5$	$\alpha \geq 0.6$	$\alpha \geq 0.7$	$\alpha \geq 0.8$	$\alpha \geq 0.9$	$\alpha \geq 1$
$(n_1, n_2) = (2, 2)$	100000	99982	99280	94464	80683	55833
$(n_1, n_2) = (3, 2)$	100000	99827	97295	86293	63914	36626

(b) The exponential distribution with mean 1

	$\alpha \geq 0.5$	$\alpha \geq 0.6$	$\alpha \geq 0.7$	$\alpha \geq 0.8$	$\alpha \geq 0.9$	$\alpha \geq 1$
$(n_1, n_2) = (2, 2)$	100000	99990	99220	92658	74966	55768
$(n_1, n_2) = (3, 2)$	100000	99895	97159	82918	57068	36802

(c) The log-normal distribution with parameters $\mu = 0$ and $\sigma = 1$ for the associated normal distribution

Table 5.2: Experimental results showing the number of instances, out of 100000, for which the respective maximin approximation ratio is achievable by some allocation of goods when utilities are drawn independently from the specified probability distribution.

5.4.1 Positive Results

We first show a positive result when all groups but one contain a single agent. This is a generalization of the corresponding result for two groups (Theorem 5.3.4). The algorithm also uses the round-robin algorithm as a subroutine, but more care must be taken to account for the extra groups.

Theorem 5.4.1. *Let $n_1 \geq 2$ and $n_2 = n_3 = \cdots = n_k = 1$. Then it is possible to give every agent at least $\frac{2}{n_1+2k-3}$ of her maximin share.*

Proof. Let $\alpha := \frac{1}{n_1+2k-3}$. If some agent in a singleton group values a good at least α times her value for the whole set of goods, put that good as the only good in her allocation. Since her maximin share is at most $1/k \leq 1/2$ times her value for the whole set of goods, this agent obtains the desired guarantee. We will give the remaining agents their guarantees with respect to the reduced set of goods and agents. By Lemma 5.2.2, the maximin share of an agent can only increase as we remove an agent and a good, and the approximation ratio $\frac{2}{n_1+2k-3}$ also increases as k decreases. This implies that guarantees for the reduced instance also translate to ones for the original instance. We recompute each agent's value for the whole set of goods as well as the number of groups and repeat this step until no agent in a singleton group values a good at least α times her value for the whole set of goods.

Next, we normalize the utility of each agent for the whole set of goods to 1 as in Theorem 5.3.3. We let each of the agents in the first group, in arbitrary order, take a good worth at least α to her if there is any. The approximation guarantee is satisfied for any agent who takes a good. Suppose that after this step, there are $n_0 \leq n_1$ agents in the first group and $k_0 \leq k-1$ agents in the remaining groups who have not taken a good. We run the round-robin algorithm on these $n_0 + k_0$ agents, starting with the k_0 agents who do not belong to the first group.

Consider one of the n_0 agents in the first group. Since no good allocated by the round-robin algorithm is worth at least α to her, she has envy at most α toward each of the k_0 agents. Assume for contradiction that the bundle allocated to the first group is worth less than α to her. Then she values the bundle of each of the k_0 agents at most 2α. Hence her utility for the whole set of goods is less than $\alpha + 2k_0\alpha = \frac{2k_0+1}{n_1+2k-3} \leq \frac{2k-1}{2k-1} = 1$, a contradiction.

Consider now one of the k_0 agents in the remaining group. She has utility at least $1-(n_1-n_0)\alpha$ for the set of goods allocated by the round-robin algorithm. With respect to the bundles allocated by the round-robin algorithm, she has no envy toward herself or any of the n_0 agents in the first group, and she has envy at most α toward the remaining $k_0 - 1$ agents. Summing up the corresponding inequalities, averaging, and using the fact that her utility for all bundles combined is at least $1 - (n_1 - n_0)\alpha$, we find that her utility for her own bundle is at least $\frac{1-(n_1-n_0)\alpha-(k_0-1)\alpha}{n_0+k_0}$. It suffices to show that this is at least α. The inequality is equivalent to $\alpha(2k_0 + n_1 - 1) \leq 1$, which holds since $k_0 \leq k - 1$. □

5.4.2 Negative Results

We next show that when all groups contain at least two agents and one group contains at least five agents, no approximation is possible.

Theorem 5.4.2. *Let $n_1 \geq 4$ if k is even and $n_1 \geq 5$ if k is odd, and $n_2 = n_3 = \cdots = n_k = 2$. Then there exists an instance in which some agent with nonzero maximin share necessarily receives zero utility.*

Proof. Let $n_1 = 4$ if k is even and 5 if k is odd, and suppose that there are $2k$ goods. In each of the groups $2, 3, \ldots, k$, one agent has utility 1 for the first k goods and 0 for the last k, while the other agent has utility 0 for the first k goods and 1 for the last k. Hence all of these agents have a maximin share of 1. To ensure that they all get nonzero utility, each group must receive one of the first k and one of the last k goods. This only leaves one good from the first k and one from the last k to the first group.

First, consider the case k even. Let the utilities of the agents in the first group be given by

- $\mathbf{u}_{11} = (\overbrace{1,1,\ldots,1}^{k/2},\overbrace{0,0,\ldots,0}^{k/2},\overbrace{1,1,\ldots,1}^{k/2},\overbrace{0,0,\ldots,0}^{k/2})$;
- $\mathbf{u}_{12} = (\overbrace{1,1,\ldots,1}^{k/2},\overbrace{0,0,\ldots,0}^{k/2},\overbrace{0,0,\ldots,0}^{k/2},\overbrace{1,1,\ldots,1}^{k/2})$;
- $\mathbf{u}_{13} = (\overbrace{0,0,\ldots,0}^{k/2},\overbrace{1,1,\ldots,1}^{k/2},\overbrace{1,1,\ldots,1}^{k/2},\overbrace{0,0,\ldots,0}^{k/2})$;
- $\mathbf{u}_{14} = (\overbrace{0,0,\ldots,0}^{k/2},\overbrace{1,1,\ldots,1}^{k/2},\overbrace{0,0,\ldots,0}^{k/2},\overbrace{1,1,\ldots,1}^{k/2})$.

All four agents have a maximin share of 1, but for any combination of a good from the first k goods and one from the last k, some agent obtains a utility of 0.

Next, consider the case k odd. Let the utilities of the agents in the first group be given by

- $\mathbf{u}_{11} = (\overbrace{1,1,\ldots,1}^{(k-1)/2},\overbrace{0,0,\ldots,0}^{(k+1)/2},\overbrace{1,1,\ldots,1}^{(k+1)/2},\overbrace{0,0,\ldots,0}^{(k-1)/2});$
- $\mathbf{u}_{12} = (\overbrace{1,1,\ldots,1}^{(k+1)/2},\overbrace{0,0,\ldots,0}^{(k-1)/2},\overbrace{0,0,\ldots,0}^{(k+1)/2},\overbrace{1,1,\ldots,1}^{(k-1)/2});$
- $\mathbf{u}_{13} = (\overbrace{0,0,\ldots,0}^{(k+1)/2},\overbrace{1,1,\ldots,1}^{(k-1)/2},\overbrace{0,0,\ldots,0}^{(k-1)/2},\overbrace{1,1,\ldots,1}^{(k+1)/2});$
- $\mathbf{u}_{14} = (\overbrace{0,0,\ldots,0}^{(k-1)/2},\overbrace{1,1,\ldots,1}^{(k+1)/2},\overbrace{1,1,\ldots,1}^{(k-1)/2},\overbrace{0,0,\ldots,0}^{(k+1)/2});$
- $\mathbf{u}_{15} = (\overbrace{1,1,\ldots,1}^{(k-1)/2},0,\overbrace{1,1,\ldots,1}^{k-1},0,\overbrace{1,1,\ldots,1}^{(k-1)/2}).$

All five agents have a maximin share of 1, but as in the previous case, any combination of a good from the first k goods and one from the last k yields no utility to some agent. □

5.5 Conclusion and Future Work

In this chapter, we study the problem of approximating the maximin share when we allocate goods to groups of agents. When there are two groups, we characterize the cardinality of the groups for which we can obtain a positive approximation of the maximin share. We also show positive and negative results for approximation when there are several groups.

We conclude the chapter by listing some future directions. For two groups, closing the gap between the lower and upper bounds of the approximation ratios (Table 5.1) is a significant problem from a theoretical point of view but perhaps even more so from a practical one. In particular, it would be especially interesting to determine the asymptotic behavior of the best approximation ratio when one group contains a single agent and the number of agent in the other group grows. For the case of several groups, one can ask whether it is in general possible to obtain a positive approximation when some groups contain a single agent while others contain two agents; the techniques that we present in this chapter do not seem to extend easily to this case. Another question is to determine whether the dependence on the number of groups in the approximation ratio (Theorem 5.4.1) is necessary. One could also address the issue of truthfulness or add constraints on the allocation, for example by requiring that the allocation form a contiguous block on a line, as has been done for the traditional fair division setting [6, 24, 151].

In light of the fact that the positive results in this chapter only hold for groups with a small number of agents, a natural question is whether we can relax the fairness notion in order to allow for more positive results. For example, one could consider only requiring that a certain fraction of agents in each group, instead of all of them, think that the allocation is fair. Alternatively, if we use envy-freeness as the fairness notion, then a possible relaxation is to require envy-freeness only up to some number of goods, where the number of goods could depend on the number of agents in each group. We will explore both of these directions in the next chapter.

Chapter 6

Democratic Fair Allocation of Indivisible Goods

6.1 Introduction

In the last chapter, we have seen fairness guarantees for groups of agents based on the maximin share. While such guarantees are possible for groups with certain numbers of agents, there does not always exist an allocation that gives all agents a positive fraction of their maximin share. The impossibility occurs even for two groups with three agents each (Proposition 5.3.2). This shows that the "unanimous fairness" notion that we have worked with so far might be too strong to be practical in a number of situations.

What do groups do when they cannot attain unanimity? In democratic societies, they use some kind of voting. The premise of voting is that it is impossible to satisfy everyone, so we should try to satisfy as many members as possible. Based on this observation, we say that a division is *h-democratic fair*, for some fairness notion and for some $h \in [0, 1]$, if at least a fraction h of the agents in each group believe it is fair. We would like h, the fraction of *h*appy agents, to be as large as possible. We thus pose the following question:

> *Given a fairness notion, what is the largest h such that an h-democratic fair allocation of indivisible goods can always be found?*

Our goal in this chapter is to answer the above question for different fairness notions.

Initially, in Section 6.3 we consider two groups of agents with binary valuations. We study a relaxation of envy-freeness called *envy-freeness up to c goods* (EFc).[1] One might expect to have a trade-off curve where a larger c corresponds to a larger h. However, we find that the actual trade-off

[1] See Definition 2.1.1.

$h \downarrow$ | $q \rightarrow$	(0, 1/2]	(1/2, 1]
(0, 1/3]	Yes (Cor. 6.4.2)	**Bin**: Yes (Thm. 6.4.1), **Add**: ?
(1/3, 1/2]		**Bin**: Yes (↑), **Add**: No (↓)
(1/2, 2/3]	?	**Bin**: ?, **Add**: No (Prop. 6.4.3)
(2/3, 1]	No (Prop. 6.3.1)	

h	EFc for any constant $c \geq 1$
(0, 1/2]	Yes (Thm. 6.4.1)
(1/2, 1]	No (Prop. 6.3.5)

Table 6.1: Summary of results for two groups of agents with **Bin**ary and **Add**itive valuations. For each range of $h, q \in (0, 1]$, the table shows whether there always exists an allocation that gives at least a fraction h of the agents in each group at least a fraction q of their maximin share. For EFc, the results hold for monotonic valuations too. The arrows refer to the directions pointed to in the table. The table is clearer when viewed in color.

curve is degenerate: for every constant c, it is not possible to guarantee more than 1/2-democratic EFc (Proposition 6.3.5); as we show later, the bound 1/2 is tight. The same holds for MMS fairness.

Next, in Section 6.4 we consider two groups whose agents have arbitrary monotonic valuations. We present an efficient protocol that guarantees EF1 to at least 1/2 of the agents in each group (Theorem 6.4.1); as mentioned in the previous paragraph, the factor 1/2 is tight even for agents with binary valuations. When all agents have additive valuations, this protocol guarantees 1/2 of the MMS to 1/2 of the agents (Corollary 6.4.2). This is tight: one cannot guarantee more than 1/2 of the MMS to more than 1/3 of the agents (Proposition 6.4.3). If we are instead interested in relaxing envy-freeness, it is possible to guarantee unanimous EF$(n-1)$ when agents have additive valuations, where n denotes the total number of agents in the two groups (Theorem 6.4.4).

Finally, in Section 6.5 we present two generalizations of our results to $k \geq 3$ groups. The first generalization has stronger fairness guarantees: when all valuations are binary, it guarantees to $1/k$ of the agents in all groups both EF1 and MMS; the factor $1/k$ is tight for EF1 (Theorem 6.5.1). When valuations are additive, it guarantees an additive approximation to EF and MMS. However, the running time of the protocol might be exponential. The second generalization relies on a polynomial-time protocol but provides weaker guarantees: when all valuations are binary, it guarantees MMS to $1/k$ of the agents (Theorem 6.5.6), and when valuations are additive, it guarantees an additive approximation to MMS (Theorem 6.5.5).

Some of our results and open questions are summarized in Table 6.1.

6.2 Preliminaries

The basic definitions and notation of the group fair division setting are introduced in Section 2.1. Unlike in the last two chapters, we do not always assume in this chapter that agents have additive valuations. Since we will be primarily concerned with EF1 and MMS-fairness, we state and prove a proposition that shows an interesting link between the two notions.

Proposition 6.2.1. *If an allocation is EF1 for an agent with an additive utility function, then*

(a) it is also $1/k$-MMS-fair for that agent—the factor $1/k$ is tight;

(b) if the agent's utility function is binary, then the allocation is also MMS-fair for that agent.

Proof. Denote by u the utility function of the agent and assume without loss of generality that the agent is in group A_1. We prove the two parts in order.

(a) EF1 implies that in each bundle $G_i \neq G_1$, there exists a subset C_i with $|C_i| \leq 1$ such that $u(G_1) \geq u(G_i \backslash C_i)$. Summing over all bundles gives $k \cdot u(G_1) \geq u(G \backslash (C_2 \cup C_3 \cup \cdots \cup C_k))$. Now, in any partition of G into k bundles, there is at least one bundle that does not contain any good in $C_2 \cup C_3 \cup \cdots \cup C_k$. This bundle is contained in $G \backslash (C_2 \cup C_3 \cup \cdots \cup C_k)$. Therefore, the MMS is at most $u(G \backslash (C_2 \cup C_3 \cup \cdots \cup C_k))$, which is at most $k \cdot u(G_1)$. Therefore, $u(G_1)$ is at least $1/k$ of the MMS.

To show that the factor $1/k$ is tight, assume that there are $2k-1$ goods with $u(g_1) = u(g_2) = \cdots = u(g_k) = 1$ and $u(g_{k+1}) = u(g_{k+2}) = \cdots = u(g_{2k-1}) = k$. If the agent's group A_1 gets g_1 and group i gets $\{g_i, g_{k+i-1}\}$ for $i = 2, 3, \ldots, k$, the agent gets utility 1 and finds the allocation EF1. However, the MMS is k, as can be seen from the partition $(\{g_1, g_2, \ldots, g_k\}, \{g_{k+1}\}, \{g_{k+2}\}, \ldots, \{g_{2k-1}\})$.

(b) Suppose that the agent's group gets l of the agent's desired goods. EF1 implies that each of the other $k-1$ groups gets at most $l+1$ of the agent's desired goods. Hence the agent has at most $l + (k-1)(l+1) = kl + k - 1$ desired goods. Therefore the agent's MMS is at most l, so the allocation is MMS-fair for her.

This concludes the proof. □

6.3 Two Groups with Binary Valuations

This section considers the special case in which there are two groups, the agents have additive valuations, and each agent either *desires* a good (in which case her utility for the good is 1) or does not desire it (in which case her utility is 0). Even in this special case, some fairness guarantees are unattainable. The instance used in the following proposition is the same as the one in Proposition 5.3.2.

Proposition 6.3.1. *For any $h > 2/3$, there is an instance with two groups of agents with binary valuations in which no allocation is h-democratic positive-MMS-fair.*

Proof. There are three goods. In each group there are three agents, each of whom has utility 0 for a unique good and utility 1 for each of the other two goods. Each agent has a positive MMS (1), but no allocation gives all agents a positive utility. □

We next leverage a combinatorial construction of Erdős to show the limitations of 1-out-of-c MMS-fairness.

Lemma 6.3.2. *For any integer $c \geq 2$, there is an instance with two groups each consisting of $c^2 2^{c+1}$ agents with binary valuations, such that each agent desires c goods but no allocation gives all agents a positive utility.*

Proof. Erdős [58] proved that for any positive integer c, there exists a collection $\mathcal{C}$ of $c^2 2^{c+1}$ subsets of size c of a base set G that does not have "property B". This means that no matter how we partition G into two subsets G_1 and G_2, some subset in $\mathcal{C}$ has an empty intersection with G_1 or G_2.

Take the elements of G to be our goods. Each group consists of $c^2 2^{c+1}$ agents, each of whom desires all goods in a unique subset of goods in $\mathcal{C}$. Then every agent desires c goods, but no allocation gives all agents a positive utility. □

Proposition 6.3.3. *For any integer $c \geq 2$ and any $h > 1 - \frac{1}{c^2 2^{c+1}}$, there is an instance with two groups of agents with binary valuations in which no allocation is h-democratic 1-out-of-c MMS-fair.*

Proof. Consider the instance from Lemma 6.3.2. The 1-out-of-c MMS of an agent who desires c goods is positive (1), but no allocation gives all agents a positive utility. □

As Lemma 6.3.2 shows, in certain instances it might be impossible to give every agent a positive utility. Interestingly, deciding whether an instance admits an allocation that leaves no agent with zero utility is an NP-complete problem.

Proposition 6.3.4. *Deciding whether an instance with two groups of agents with binary valuations admits an allocation that gives every agent a positive utility is NP-complete.*

Proof. For any allocation, we can clearly verify in polynomial time whether it yields a positive utility to every agent. To show that the problem is NP-hard, we reduce from Monotone SAT, a variant of the classical satisfiability problem where each clause contains either only positive literals or only negative literals. Monotone SAT is known to be NP-hard [70, p. 259].

Given a Monotone SAT formula ϕ with variables $x_1, x_2, \ldots, x_m$, let there be m items corresponding to the m variables. For each clause that contains only positive literals, we construct an agent in the first group who values exactly the items contained in this clause. Similarly, for each clause that contains only negative literals, we construct an agent in the second group who values exactly the items contained in this clause. Any assignment that satisfies ϕ gives rise to an allocation where the items corresponding to true variables in the assignment are allocated to the first group and those corresponding to false variables in the assignment are allocated to the second group; this allocation gives every agent nonzero utility. Likewise, any allocation that gives every agent nonzero utility yields a satisfying assignment of ϕ. Hence the reduction is valid. □

If we change the fairness requirement from MMS to EFc, then we can satisfy no more than half of the agents in each group.

Proposition 6.3.5. *For any constant integer $c \geq 1$ and any $h > 1/2$, there is an instance with two groups of agents with binary valuations in which no allocation is h-democratic EFc.*

Proof. Consider an instance with $m = 4l$ goods and $\binom{4l}{2l}$ agents in each group, for some $l \geq 1$. Each agent desires a unique subset of $2l$ goods. An allocation is EFc for an agent if and only if her group receives at least $l - \lfloor c/2 \rfloor$ of her $2l$ desired goods.

The symmetry between the groups implies that the best fairness guarantee can be attained by giving exactly $2l$ goods to each group; the symmetry between the goods implies that it does not matter which $2l$ goods are given to which group. In each such allocation, the number of a group's members who receive exactly j desired goods is $\binom{2l}{j} \cdot \binom{4l-2l}{2l-j} = \binom{2l}{j}^2$. Therefore, the number of a group's members who receive at least $l - \lfloor c/2 \rfloor$ desired goods is: $\sum_{j=l-\lfloor c/2 \rfloor}^{2l} \binom{2l}{j}^2 = \sum_{j=l-\lfloor c/2 \rfloor}^{l-1} \binom{2l}{j}^2 + \frac{1}{2}\binom{2l}{l}^2 + \frac{1}{2}\binom{4l}{2l}$, where the equality follows from expanding the central binomial coefficient $\binom{4l}{2l}$. The fraction of a group's members who think the division is EFc is attained by dividing this expression by $\binom{4l}{2l}$. This fraction is:

$$\frac{1}{2} + \frac{1}{2} \cdot \frac{\binom{2l}{l}^2}{\binom{4l}{2l}} + \frac{\sum_{j=l-\lfloor c/2 \rfloor}^{l-1} \binom{2l}{j}^2}{\binom{4l}{2l}} \leq \frac{1}{2}\left(1 + (c+1)\frac{\binom{2l}{l}^2}{\binom{4l}{2l}}\right).$$

Using Stirling's approximation, we find that $\binom{2l}{l} \sim \frac{4^l}{\sqrt{\pi l}}$, so

$$\frac{\binom{2l}{l}^2}{\binom{4l}{2l}} \sim \frac{\frac{4^{2l}}{\pi l}}{\frac{4^{2l}}{\sqrt{2\pi l}}} = \sqrt{\frac{2}{\pi l}}.$$

As $l \to \infty$, the fraction of agents in each group who think that the allocation is EFc approaches $1/2$, as claimed. □

6.4 Two Groups with General Valuations

In this section, we assume that there are two groups and each agent has an arbitrary monotonic utility function.

We start with a positive result: it is always possible to efficiently allocate goods so that at least half of the agents in each group believe the division is EF1. The protocol mirrors the well-known "cut-and-choose" protocol for dividing a cake between two agents.

Theorem 6.4.1. *For two groups of agents with monotonic valuations, $1/2$-democratic EF1 is attainable.*

Proof. We arrange the goods in a line and process them from left to right. Starting from an empty block, we add one good at a time until the current block is EF1 for at least half of the agents in at least one group. We allocate the current block to one such group, and the remaining goods to the other group.

Since the whole set of goods is EF1 for both groups, the protocol terminates. Assume without loss of generality that the left block G_1 is allocated to the first group A_1, and the right block G_2 to the second group A_2. By the description of the protocol, the allocation is EF1 for at least half of the agents in A_1, so it remains to show that the same holds for A_2. Let g be the last good added to the left block. More than half of the agents in A_2 think that $G_1 \backslash \{g\}$ is not EF1, so for these agents, $G_1 \backslash \{g\}$ is worth less than $G_2 \cup \{g\} \backslash \{g'\}$ for any $g' \in G_2 \cup \{g\}$. Taking $g' = g$, we find that these agents value $G_1 \backslash \{g\}$ less than G_2. But this implies that the agents find G_2 to be EF1, completing the proof. □

Theorem 6.4.1 shows that if the goods lie on a line, we can find a $1/2$-democratic EF1 allocation that moreover gives each group a contiguous block on the line. This may be important, for example, if the goods are houses on a street and each group wants to have all its houses in a contiguous block [24, 151].

If agents have additive valuations, Proposition 6.2.1 implies:

Corollary 6.4.2. *For two groups with additive agents, $1/2$-democratic $1/2$-MMS-fairness is attainable.*

For EF1, the factor $1/2$ in Theorem 6.4.1 is tight even for binary valuations, as shown in Proposition 6.3.5. For $1/2$-MMS-fairness, the factor $1/2$ in Corollary 6.4.2 is "almost" tight, as the following proposition shows.

Proposition 6.4.3. *For any $h > 1/3$ and $q > 1/2$, there is an instance with two groups of agents with additive valuations in which no allocation is h-democratic q-MMS-fair.*

Proof. Consider an instance with $m = 3$ goods and $n_1 = n_2 = 3$ agents in each group, with utility vectors: $\mathbf{u}_{i1} = (2, 1, 1)$, $\mathbf{u}_{i2} = (1, 2, 1)$, and $\mathbf{u}_{i3} = (1, 1, 2)$ for $i = 1, 2$. The MMS of every agent is 2. In any allocation, one group receives at most one good, so at most one of its three agents receives utility more than 1. In that group, at most $1/3$ of the agents receive more than $1/2$ of their MMS. □

A corollary of Proposition 6.4.3 is that for every $h \in (1/3, 1/2]$, the maximum fraction q such that there always exists an h-democratic q-MMS-fair allocation is $q = 1/2$.

Proposition 6.3.5 shows that the factor $1/2$ in Theorem 6.4.1 cannot be improved even for agents with binary valuations and even if we relax EF1 to EFc for any constant c. Nevertheless, if we let the relaxation of envy-freeness depend on the number of agents, it is possible to obtain a unanimous fairness guarantee.

Theorem 6.4.4. *For any two groups of agents with additive valuations, there exists an allocation that is EF*$(n-1)$ *for all agents, where* $n = n_1 + n_2$ *is the total number of agents in both groups.*

Proof. Choose an arbitrary agent in one of the groups. We will partition the goods into two parts and let the agent choose the part that she prefers. The resulting allocation is envy-free and therefore EF$(n-1)$ for this agent. Hence it suffices to show that there exists a partition in which each bundle is EF$(n-1)$ (with respect to the other bundle) for all of the remaining $n-1$ agents.

To this end, assume that there is a divisible good, which we refer to as a cake, represented by the half-open interval $(0, m]$. The value density functions of the agents over the cake are piecewise constant: for every $l \in \{1, 2, \ldots, m\}$, the value density v_{ij} in the half-open interval $(l-1, l]$ equals $u_{ij}(g_l)$.

It is known that there exists a partition of the cake into two parts, using at most $n-1$ cuts, in which every agent has equal value for both parts [3]. Starting with two empty bundles, for each $l \in \{1, 2, \ldots, m\}$, we add good g_l to the bundle corresponding to the part that contains at least half of the interval $(l-1, l]$. (If both parts contain exactly half of the interval, we add g_l to an arbitrary bundle.)

We claim that every agent finds either bundle to be EF$(n-1)$. Fix an agent a_{ij} and a bundle G'. From our partitioning choice, we have that $u_{ij}(G \backslash G') - u_{ij}(G') \leq u_{ij}(G'')$ for some set $G'' \subseteq G \backslash G'$ of size at most $n-1$. This implies that the agent finds G' to be EF$(n-1)$ with respect to $G \setminus G'$, as claimed. □

6.5 Three or More Groups

In this section we study the most general setting where we allocate goods among any number of groups. When there are two groups, the protocol in Theorem 6.4.1 is computationally efficient and yields an allocation that is both approximately envy-free and approximately MMS-fair. We present two ways of generalizing the result to multiple groups: one keeps the approximate envy-freeness guarantee but loses computational efficiency, while the other keeps only the approximate MMS-fairness guarantee but also retains computational efficiency.

6.5.1 Approximate Envy-Freeness

The main theorem of this subsection is the following.

Theorem 6.5.1. *When all agents have binary valuations, there exists an allocation that is* $1/k$*-democratic EF1 and* $1/k$*-democratic MMS-fair. The factor* $1/k$ *is tight for EF1.*

To establish this theorem, we prove two lemmas that may be of independent interest—one on the allocation of divisible goods ("cake cutting") and the other on group allocation for agents with additive valuations.

The result on cake cutting generalizes the theorems of Stromquist [142] and Su [144], who proved the existence of contiguous envy-free cake allocations for individual agents. Since these results are well-known, we present the model and proof quite briefly, focusing on the changes required to generalize from individuals to groups.

We consider a "cake" modeled as the interval $[0, 1]$. Each agent a_{ij} has a value density function $v_{ij} : [0, 1] \to \mathbb{R}_{\geq 0}$. The value of an agent for a piece X is $V_{ij}(X) = \int_{x \in X} v_{ij}(x)dx$. Denoting by X_i the allocation to group i, an allocation is *envy-free* for an agent a_{ij} if $V_{ij}(X_i) \geq V_{ij}(X_{i'})$ for every group i'. A *contiguous allocation* is an allocation of the cake in which each group gets a contiguous interval.

Lemma 6.5.2. *There always exists a contiguous cake allocation that is $1/k$-democratic envy-free. The factor $1/k$ is tight.*

Proof. The space of all contiguous partitions corresponds to the standard simplex in $\mathbb{R}^k$. Triangulate that simplex and assign each vertex of the triangulation to one of the groups. In each vertex, ask the group owning that vertex to select one of the k pieces using *plurality voting* among its members, breaking ties arbitrarily. Label that vertex with the group's selection. The resulting labeling satisfies the conditions of *Sperner's lemma* (see Su [144]). Therefore, the triangulation has a *Sperner subsimplex*—a subsimplex all of whose labels are different. We can repeat this process with finer and finer triangulations. This gives an infinite sequence of smaller and smaller Sperner subsimplices. This sequence has a subsequence that converges to a single point. By the continuity of preferences, this limit point corresponds to a partition in which each group selects a different piece. Since the selection is by plurality, at least $1/k$ of the agents in each group prefer their group's piece over all other pieces.

The tightness of the $1/k$ factor follows from Lemma 6 of Segal Halevi and Nitzan [133]. It shows an example with k groups and n' agents in each group with the property that in order to give a positive value to q out of n' agents in each group, we need to cut the cake into at least $k(kq - n')/(k - 1)$ intervals. In a contiguous partition there are exactly k intervals. Therefore, the fraction of agents in each group that can be guaranteed a positive value is $q/n' \leq 1/k + 1/n' - 1/kn'$. Since n' can be arbitrarily large, the largest fraction that can be guaranteed is $1/k$. □

The next lemma presents a reduction from approximate envy-free allocation of indivisible goods to envy-free cake cutting. We call this approximation "EF-minus-2". An allocation is *EF-minus-2* for agent a_{ij} if for every group i', $u_{ij}(G_i) > u_{ij}(G_{i'}) - 2u_{ij,\max}$. The reduction generalizes Theorem 5 of Suksompong [151]; a similar reduction was used in Theorem 3 of Barrera et al. [17].

Lemma 6.5.3. *When agents have additive valuations, there always exists a contiguous allocation of indivisible goods that is $1/k$-democratic EF-minus-2.*

Proof. We create an instance of the cake cutting problem in the following way.

- The cake is the half-open interval $(0, m]$.
- The value density functions are piecewise constant: for every $l \in \{1, \ldots, m\}$, the value density v_{ij} in the half-open interval $(l-1, l]$ equals $u_{ij}(g_l)$.

By Lemma 6.5.2, there exists a contiguous cake allocation that is envy-free for at least $1/k$ of the agents in each group. From this allocation we construct an allocation of goods as follows.

- If point g of the cake is in the interior of a piece, then good g is given to the group owning that piece.
- If point g of the cake is at the boundary between two pieces, then good g is given to the group owning the piece to its left.

A group gets good g only if it owns a positive fraction of the interval $(g-1, g]$. Hence, in the allocation, each group loses strictly less than the value of a good and gains strictly less than the value of a good (relative to its value in the cake division). This means that every agent who believes that the cake allocation is envy-free also believes that the allocation of the goods is EF-minus-2. □

We are now ready to prove Theorem 6.5.1.

Proof of Theorem 6.5.1. Suppose an allocation is EF-minus-2 for some agent a_{ij}. This means that the agent's envy towards any other group is less than $2u_{ij,\max} \leq 2$. Since the agent has binary valuations, the envy is at most 1, meaning that the allocation is EF1 for that agent. Hence any $1/k$-democratic EF-minus-2 allocation, which is guaranteed to exist by Lemma 6.5.3, is also $1/k$-democratic EF1. By Proposition 6.2.1 it is also $1/k$-democratic MMS-fair.

We next show that the factor $1/k$ is tight. Assume that there are $m = km'$ goods for some large positive integer m'. Each group consists of 2^m agents, each of whom values a distinct combination of the goods. Consider first an allocation that gives exactly m' goods to each group, and fix a group. We claim that the fraction of the agents in the group whose utilities for some two bundles differ by at most 1 converges to 0 for large m'. Indeed, this follows from the central limit theorem: Fix two bundles and consider a random agent from the group; let X be the random variable denoting the (possibly negative) difference between the agent's utilities for the two bundles. Then X is a sum of m' independent and identically distributed random variables with mean 0. The central limit theorem implies that for any fixed $\epsilon > 0$, there exists a constant c such that $\Pr[|X| \leq 1] \leq \Pr[|X| \leq c\sqrt{m'}] \leq \epsilon$ for any sufficiently large m'. Taking the union bound over all pairs of bundles, we find that the fraction of agents in the group who value some two bundles within 1 of each other approaches 0 as

m' goes to infinity. This means that all but a negligible fraction of the agents find only one bundle to be EF1. By symmetry, $1/k$ of these agents find the bundle allocated to the group to be EF1. It follows that the fraction of agents in the group for whom the allocation is EF1 converges to $1/k$.

It remains to consider the case where the allocation does not give the same number of goods to all groups. In this case, let $\mathcal{G}$ denote the set of bundles with the smallest number of goods, which must be strictly smaller than m'. If we move goods from bundles with more than m' goods to bundles in $\mathcal{G}$ in such a way that the number of goods in each bundle in $\mathcal{G}$ increases by exactly one, the fraction of agents in an arbitrary group that receives a bundle in $\mathcal{G}$ who finds the allocation to be EF1 can only increase. We can repeat this process, at each step possibly adding bundles to $\mathcal{G}$, until all bundles contain the same number of goods, which is the case we have already handled. Since the fraction of agents for whom the allocation is EF1 is bounded above by $1/k$ for large m' in the latter allocation, and this fraction only increases during our process of moving goods, the same is true for the original allocation. □

The cake cutting protocol of Lemma 6.5.2 might take infinitely many steps to converge. In fact, there is no finite protocol for contiguous envy-free cake cutting even for individuals [143]. However, the division guaranteed by Lemma 6.5.3 and Theorem 6.5.1 can be found in finite time (exponential in the input size) by checking all possible allocations. An interesting open question is whether a faster algorithm exists.

6.5.2 Approximate Maximin Share

In this subsection, we show that if we weaken our fairness requirement to approximate MMS, it is possible to compute a fair allocation in time polynomial in the input size.

Lemma 6.5.4. *When agents have additive valuations, there always exists an allocation such that at least $1/k$ of the agents a_{ij} in each group A_i receive utility at least $\frac{1}{k} \cdot u_{ij}(G) - \frac{k-1}{k} \cdot u_{ij,max}$, and such an allocation can be computed efficiently.*

This lemma generalizes the corresponding result for the setting with one agent per group by Suksompong [151]. The factor $(k-1)/k$ is tight even for individual agents.

Proof. We arrange the goods in a line and process them from left to right. Starting from an empty block, we add one good at a time until the current block yields utility at least $\frac{1}{k} \cdot u_{ij}(G) - \frac{k-1}{k} \cdot u_{ij,\max}$ for at least $1/k$ of the agents in at least one group. We allocate the current block to one such group and repeat the process with the remaining $k-1$ groups. It is clear that this algorithm can be implemented efficiently. Any group that receives a block from the algorithm meets the requirement, so it suffices to show that the algorithm allocates a block to every group. We claim that if l groups are yet to receive a block, at least l/k of the agents a_{ij} in each of these groups have utility at least $\frac{l}{k} \cdot u_{ij}(G) - \frac{k-l}{k} \cdot u_{ij,\max}$ for the remaining goods. This would imply the desired result because for the last group, at least $1/k$ of the agents have utility $\frac{1}{k} \cdot u_{ij}(G) - \frac{k-1}{k} \cdot u_{ij,\max}$, which is exactly our requirement.

To show the claim, we proceed by backward induction on l. The claim trivially holds when $l = k$. Suppose that the statement holds when there are $l + 1$ groups left, and consider a group j that is not the next one to receive a block. At least $(l + 1)/k$ of the agents a_{ij} in the group have utility at least $\frac{l+1}{k} \cdot u_{ij}(G) - \frac{k-l-1}{k} \cdot u_{ij,\max}$ for the remaining goods. Since the group does not receive the next block, less than $1/k$ of the agents in the group have utility at least $\frac{1}{k} \cdot u_{ij}(G) - \frac{k-1}{k} \cdot u_{ij,\max}$ for the block excluding the last good. Hence, less than $1/k$ of the agents have utility at least $\frac{1}{k} \cdot u_{ij}(G) + \frac{1}{k} \cdot u_{ij,\max}$ for the whole block. This means that at least l/k of the agents have utility at least $\left(\frac{l+1}{k} \cdot u_{ij}(G) - \frac{k-l-1}{k} \cdot u_{ij,\max}\right) - \left(\frac{1}{k} \cdot u_{ij}(G) + \frac{1}{k} \cdot u_{ij,\max}\right) = \frac{l}{k} \cdot u_{ij}(G) - \frac{k-l}{k} \cdot u_{ij,\max}$, completing the induction. □

It is clear by definition that the MMS of any agent a_{ij} is at most $\frac{1}{k} \cdot u_{ij}(G)$. Lemma 6.5.4 therefore implies the following theorem.

Theorem 6.5.5. *When agents have additive valuations, there always exists an allocation such that at least $1/k$ of the agents a_{ij} in each group A_i receive utility at least $MMS_{ij}(G) - \frac{k-1}{k} \cdot u_{ij,max}$, and such an allocation can be computed efficiently.*

For binary valuations, if we change the stopping condition in Lemma 6.5.4 to be when the current block yields the MMS for at least $1/k$ of the agents in some group, we get the following guarantee.

Theorem 6.5.6. *When agents have binary valuations, there always exists a $1/k$-democratic MMS-fair allocation, and such an allocation can be computed efficiently.*

6.6 Conclusion and Future Work

In this chapter, we initiate the study of democratic fairness in the allocation of indivisible goods among groups. For two groups, we have a comprehensive understanding of possible democratic fairness guarantees. We have a complete characterization of possible envy-freeness approximations, and upper and lower bounds for maximin-share-fairness approximations. Some remaining gaps are shown in Table 6.1; closing them raises interesting combinatorial challenges.

For $k \geq 3$ groups, the challenges are much greater. Currently all of our fairness guarantees are to no more than $1/k$ of the agents in each group. From a practical perspective, it may be important in some settings to give fairness guarantees to at least half of the agents in all groups. Finding protocols that provide such guarantees is an avenue for future work. From an algorithmic perspective, it is interesting whether there exists a polynomial-time algorithm that guarantees EF1 to any positive fraction of the agents.

Part II

Decision Making

Chapter 7

On the Structure of Stable Tournament Solutions

7.1 Introduction

In the second part of the dissertation, we turn our attention to decision making problems, where our goal is to choose the "best" alternatives from a given set of alternatives. We begin in this chapter by assuming the existence of a dominance relation between the alternatives—in other words, the alternatives form a *tournament*.[1] The question of selecting the best alternatives in this setting has been studied in detail in the literature on tournament solutions. The lack of transitivity is typically attributed to the independence of pairwise comparisons as they arise in sports competitions, multi-criteria decision analysis, and preference aggregation.[2] In particular, the pairwise majority relation of a profile of transitive individual preference relations often forms the basis of the study of tournament solutions. This is justified by a theorem due to McGarvey [112], which shows that every tournament can be induced by some underlying preference profile. Many tournament solutions therefore correspond to well-known social choice functions such as Copeland's rule, Slater's rule, the Banks set, and the bipartisan set.

Over the years, many desirable properties of tournament solutions have been proposed. Some of these properties, so-called *choice consistency conditions*, make no reference to the actual tournament but only relate choices from different subtournaments to each other. An important choice consistency condition, which goes under various names, requires that the choice set is invariant under the removal of unchosen alternatives. In conjunction with a dual condition on expanded feasible sets, this property is known as *stability* [41]. Stability implies that choices are made in a robust and coherent

[1] See Section 2.2.

[2] Due to their generality, tournament solutions have also found applications in unrelated areas such as biology [2, 96, 129, 139].

way. Furthermore, stable choice functions can be rationalized by a preference relation on *sets* of alternatives.

Examples of stable tournament solutions are the *top cycle*, *the minimal covering set*, and the *bipartisan set*. The latter is elegantly defined via the support of the unique mixed maximin strategies of the zero-sum game given by the tournament's skew-adjacency matrix. Curiously, for some tournament solutions, including the *tournament equilibrium set* and the *minimal extending set*, proving or disproving stability turned out to be exceedingly difficult. As a matter of fact, whether the tournament equilibrium set satisfies stability was open for more than two decades before the existence of counterexamples with about 10^{136} alternatives was shown using the probabilistic method.

Brandt [31] systematically constructed stable tournament solutions by applying a well-defined operation to existing (non-stable) tournament solutions. Brandt's study was restricted to a particular class of generating tournament solutions, namely tournament solutions that can be defined via qualified subsets (such as the *uncovered set* and the *Banks set*). For any such generator, Brandt gave sufficient conditions for the resulting tournament solution to be stable. Later, Brandt et al. [42] showed that for one particular generator, the Banks set, the sufficient conditions for stability are also necessary.

In this chapter, we show that *every* stable choice function is generated by a unique underlying simple choice function, which never excludes more than one alternative (Theorem 7.3.1). We go on to prove a general characterization of stable tournament solutions that is not restricted to generators defined via qualified subsets (Theorem 7.3.4). As a corollary, we obtain that the sufficient conditions for generators defined via qualified subsets are also necessary (Corollary 7.3.6). Finally, we prove a strong connection between stability and a new property of tournament solutions called *local reversal symmetry* (Theorem 7.6.3). Local reversal symmetry requires that an alternative is chosen if and only if it is unchosen when all its incident edges are inverted. This result allows us to settle two important problems in the theory of tournament solutions. We provide the first concrete tournament—consisting of 24 alternatives—in which the tournament equilibrium set violates stability (Theorem 7.6.4). Secondly, we prove that there is no more discriminating stable tournament solution than the bipartisan set (Corollary 7.6.7). We also axiomatically characterize the bipartisan set by using only properties that have been previously proposed in the literature (Corollary 7.6.10). We believe that these results serve as a strong argument in favor of the bipartisan set if choice consistency is desired.

7.2 Stable Sets and Stable Choice Functions

Let U be a universal set of alternatives. Any finite non-empty subset of U will be called a *feasible set*. Before we analyze tournament solutions in Section 7.4, we first consider a more general model of choice which does not impose any structure on feasible sets. A *choice function* is a function that maps every feasible set A to a non-empty subset of A called the *choice set* of A. For two choice functions S and S', we write $S' \subseteq S$, and say that S' is a *refinement* of S and S a *coarsening* of S', if $S'(A) \subseteq S(A)$ for all feasible sets A. A choice function S is called *trivial* if $S(A) = A$ for all feasible sets A.

Brandt [31] proposed a general method for refining a choice function S by defining minimal sets that satisfy internal and external stability criteria with respect to S, similar to von-Neumann–Morgenstern stable sets in cooperative game theory.[3]

A subset of alternatives $X \subseteq A$ is called S-*stable* within feasible set A for choice function S if it consists precisely of those alternatives that are chosen in the presence of all alternatives in X. Formally, X is S-stable in A if

$$X = \{a \in A\colon a \in S(X \cup \{a\})\}.$$

Equivalently, X is S-stable if and only if

$$S(X) = X, \text{ and} \qquad \text{(internal stability)}$$
$$a \notin S(X \cup \{a\}) \text{ for all } a \in A \setminus X. \qquad \text{(external stability)}$$

The intuition underlying this formulation is that there should be no reason to restrict the choice set by excluding some alternative from it (internal stability) and there should be an argument against each proposal to include an outside alternative into the choice set (external stability).

An S-stable set is *inclusion-minimal* (or simply *minimal*) if it does not contain another S-stable set. $\widehat{S}(A)$ is defined as the union of all minimal S-stable sets in A. $\widehat{S}$ defines a choice function whenever every feasible set admits at least one S-stable set. In general, however, neither the existence of S-stable sets nor the uniqueness of minimal S-stable sets is guaranteed. We say that $\widehat{S}$ is *well-defined* if every choice set admits exactly one minimal S-stable set. We can now define the central concept of this chapter.

Definition 7.2.1. *A choice function S is* stable *if $\widehat{S}$ is well-defined and $S = \widehat{S}$.*

Stability is connected to rationalizability and non-manipulability. In fact, every stable choice function can be rationalized via a preference relation on *sets* of alternatives [41] and, in the context of social choice, stability and monotonicity imply strategyproofness with respect to Kelly's preference extension [31].

The following example illustrates the preceding definitions. Consider universe $U = \{a, b, c\}$ and choice function S given by the table below (choices from singleton sets are trivial and therefore

[3] This is a generalization of earlier work by Dutta [57], who defined the minimal covering set as the unique minimal set that is internally and externally stable with respect to the uncovered set (see Section 7.4).

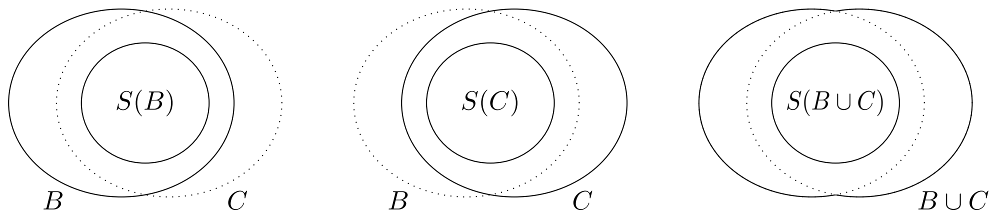

Figure 7.1: Visualization of stability. A stable choice function S chooses a set from both B (left) and C (middle) if and only if it chooses the same set from $B \cup C$ (right). The direction from the left and middle diagrams to the right diagram corresponds to $\widehat{\gamma}$ while the converse direction corresponds to $\widehat{\alpha}$.

omitted).

X	$S(X)$	$\widehat{S}(X)$
$\{a,b\}$	$\{a\}$	$\{a\}$
$\{b,c\}$	$\{b\}$	$\{b\}$
$\{a,c\}$	$\{a\}$	$\{a\}$
$\{a,b,c\}$	$\{a,b,c\}$	$\{a\}$

The feasible set $\{a,b,c\}$ admits exactly two S-stable sets, $\{a,b,c\}$ itself and $\{a\}$. The latter holds because $S(\{a\}) = \{a\}$ (internal stability) and $S(\{a,b\}) = S(\{a,c\}) = \{a\}$ (external stability). All other feasible sets X admit unique S-stable sets, which coincide with $S(X)$. Hence, $\widehat{S}$ is well-defined and given by the entries in the rightmost column of the table. Since $S \neq \widehat{S}$, S fails to be stable. $\widehat{S}$, on the other hand, satisfies stability.

Choice functions are usually evaluated by checking whether they satisfy choice consistency conditions that relate choices from different feasible sets to each other. The following two properties, $\widehat{\alpha}$ and $\widehat{\gamma}$, are set-based variants of Sen's α and γ [136]. $\widehat{\alpha}$ is a rather prominent choice-theoretic condition, also known as Chernoff's *postulate* 5^* [50], the *strong superset property* [23], *outcast* [1], and the *attention filter axiom* [109].[4]

Definition 7.2.2. *A choice function S satisfies $\widehat{\alpha}$ if for all feasible sets B and C,*

$$S(B) \subseteq C \subseteq B \textit{ implies } S(C) = S(B). \qquad (\widehat{\alpha})$$

A choice function S satisfies $\widehat{\gamma}$ if for all feasible sets B and C,

$$S(B) = S(C) \textit{ implies } S(B) = S(B \cup C). \qquad (\widehat{\gamma})$$

It has been shown that stability is equivalent to the conjunction of $\widehat{\alpha}$ and $\widehat{\gamma}$.

[4]We refer to [114] for a more thorough discussion of the origins of this condition.

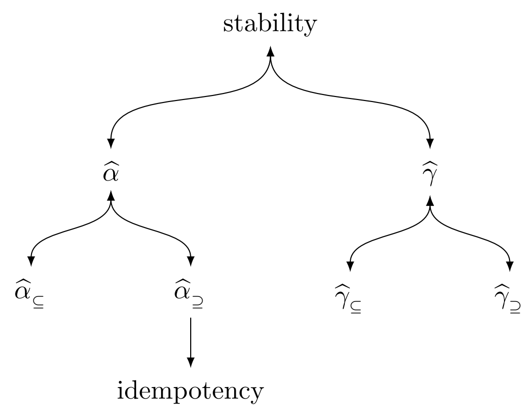

Figure 7.2: Logical relationships between choice-theoretic properties

Theorem 7.2.3 ([41])**.** *A choice function is stable if and only if it satisfies $\widehat{\alpha}$ and $\widehat{\gamma}$.*

Hence, a choice function S is stable if and only if for all feasible sets B, C, and X with $X \subseteq B \cap C$,

$$X = S(B) \text{ and } X = S(C) \quad \text{if and only if} \quad X = S(B \cup C).$$

Stability, $\widehat{\alpha}$, and $\widehat{\gamma}$ are illustrated in Figure 7.1.

For a finer analysis, we split $\widehat{\alpha}$ and $\widehat{\gamma}$ into two conditions [41, Remark 1].

Definition 7.2.4. *A choice function S satisfies*

- $\widehat{\alpha}_{\subseteq}$ *if for all B, C, it holds that $S(B) \subseteq C \subseteq B$ implies $S(C) \subseteq S(B)$,*[5]
- $\widehat{\alpha}_{\supseteq}$ *if for all B, C, it holds that $S(B) \subseteq C \subseteq B$ implies $S(C) \supseteq S(B)$,*
- $\widehat{\gamma}_{\subseteq}$ *if for all B, C, it holds that $S(B) = S(C)$ implies $S(B) \subseteq S(B \cup C)$, and*
- $\widehat{\gamma}_{\supseteq}$ *if for all B, C, it holds that $S(B) = S(C)$ implies $S(B) \supseteq S(B \cup C)$.*

Obviously, for any choice function S we have

$$\begin{aligned} S \text{ satisfies } \widehat{\alpha} \quad &\text{if and only if} \quad S \text{ satisfies } \widehat{\alpha}_{\subseteq} \text{ and } \widehat{\alpha}_{\supseteq}, \text{ and} \\ S \text{ satisfies } \widehat{\gamma} \quad &\text{if and only if} \quad S \text{ satisfies } \widehat{\gamma}_{\subseteq} \text{ and } \widehat{\gamma}_{\supseteq}. \end{aligned}$$

A choice function is *idempotent* if the choice set is invariant under repeated application of the choice function, i.e., $S(S(A)) = S(A)$ for all feasible sets A. It is easily seen that $\widehat{\alpha}_{\supseteq}$ is stronger than idempotency since $S(S(A)) \supseteq S(A)$ implies $S(S(A)) = S(A)$.

Figure 7.2 shows the logical relationships between stability and its weakenings.

[5] $\widehat{\alpha}_{\subseteq}$ has also been called the *Aïzerman property* or the *weak superset property* (e.g., [31, 98]).

7.3 Generators of Stable Choice Functions

We say that a choice function S' *generates* a stable choice function S if $S = \widehat{S'}$. Understanding stable choice functions can be reduced to understanding their generators. It turns out that important generators of stable choice functions are *simple* choice functions, i.e., choice functions S' with $|S'(A)| \geq |A| - 1$ for all A. In fact, every stable choice function S is generated by a unique simple choice function. To this end, we define the *root* of a choice function S as

$$\lceil S \rceil(A) = \begin{cases} S(A) & \text{if } |S(A)| = |A| - 1, \\ A & \text{otherwise.} \end{cases}$$

Not only does $\lceil S \rceil$ generate S, but any choice function sandwiched between S and $\lceil S \rceil$ is a generator of S.

Theorem 7.3.1. *Let S and S' be choice functions such that S is stable and $S \subseteq S' \subseteq \lceil S \rceil$. Then, $\widehat{S'}$ is well-defined and $\widehat{S'} = S$. In particular, S is generated by the simple choice function $\lceil S \rceil$.*

Proof. We first show that any S-stable set is also S'-stable. Suppose that a set $X \subseteq A$ is S-stable in A. Then $S(X) = X$, and $a \notin S(X \cup \{a\})$ for all $a \in A \backslash X$. Since S satisfies $\widehat{\alpha}$, we have $S(X \cup \{a\}) = X$ and therefore $\lceil S \rceil(X \cup \{a\}) = X$ for all $a \in A \backslash X$. Using the inclusion relationship $S \subseteq S' \subseteq \lceil S \rceil$, we find that $S'(X) = X$ and $S'(X \cup \{a\}) = X$ for all $a \in A \backslash X$. Hence, X is S'-stable in A.

Next, we show that every S'-stable set contains an S-stable set. Suppose that a set $X \subseteq A$ is S'-stable in A. Then $S'(X) = X$ and $a \notin S'(X \cup \{a\})$ for all $a \in A \backslash X$. Using the relation $S \subseteq S'$, we find that $a \notin S(X \cup \{a\})$ for all $a \in A \backslash X$. We will show that $S(X) \subseteq X$ is S-stable in A. Since S satisfies $\widehat{\alpha}$, we have $S(S(X)) = S(X)$ and $S(X \cup \{a\}) = S(X)$ for all $a \in A \backslash X$. It remains to show that $b \notin S(S(X) \cup \{b\})$ for all $b \in A \backslash S(X)$. If $b \in A \backslash X$, we already have that $S(X \cup \{b\}) = S(X)$ and therefore $S(S(X) \cup \{b\}) = S(X)$ by $\widehat{\alpha}$. Otherwise, if $b \in X \backslash S(X)$, $\widehat{\alpha}$ again implies that $S(S(X) \cup \{b\}) = S(X)$.

Since S is stable, for any feasible set A there exists a unique minimal S-stable set in A, which is given by $S(A) = \widehat{S}(A)$. From what we have shown, this set is also S'-stable, and moreover any S'-stable set contains an S-stable set which in turn contains $S(A)$. Hence $S(A)$ is also the unique minimal S'-stable set in A. This implies that $\widehat{S'}$ is well-defined and $\widehat{S'} = \widehat{S} = S$. □

Theorem 7.3.1 entails that in order to understand stable choice functions, we only need to understand the circumstances under which a single alternative is discarded.[6]

[6]Together with Theorem 7.3.4, Theorem 7.3.1 also implies that, for any stable tournament solution S, $\lceil S \rceil$ is a coarsest generator of S. When only considering generators that satisfy $\widehat{\alpha}_{\subseteq}$, $\lceil S \rceil$ is also *the* coarsest generator of S. In addition, since simple choice functions trivially satisfy $\widehat{\alpha}_{\subseteq}$, the two theorems imply that $\lceil S \rceil$ is the unique simple choice function generating S.

An important question is which simple choice functions are roots of stable choice functions. It follows from the definition of root functions that any root of a stable choice function needs to satisfy $\widehat{\alpha}$. This condition is, however, not sufficient as it is easy to construct a simple choice function S that satisfies $\widehat{\alpha}$ such that $\widehat{S}$ violates $\widehat{\alpha}$. Nevertheless, the theorem implies that the number of stable choice functions can be bounded by counting the number of simple choice functions that satisfy $\widehat{\alpha}$. The number of simple choice functions for a universe of size $n \geq 2$ is only $\prod_{i=2}^{n}(i+1)^{\binom{n}{i}}$, compared to $\prod_{i=2}^{n}(2^i-1)^{\binom{n}{i}}$ for arbitrary choice functions.

In order to give a complete characterization of choice functions that generate stable choice functions, we need to introduce a new property. A choice function S satisfies local $\widehat{\alpha}$ if minimal S-stable sets are invariant under removing outside alternatives.[7]

Definition 7.3.2. *A choice function S satisfies* local $\widehat{\alpha}$ *if for any sets $X \subseteq Y \subseteq Z$ such that X is minimally S-stable in Z, we have that X is also minimally S-stable in Y.*

Recall that a choice function S satisfies $\widehat{\alpha}_{\subseteq}$ if for any sets A, B such that $S(A) \subseteq B \subseteq A$, we have $S(B) \subseteq S(A)$. In particular, every simple choice function satisfies $\widehat{\alpha}_{\subseteq}$. We will provide a characterization of choice functions S satisfying $\widehat{\alpha}_{\subseteq}$ such that $\widehat{S}$ is stable. First we need the following (known) lemma.

Lemma 7.3.3 ([41])**.** *Let S be a choice function such that $\widehat{S}$ is well-defined. Then $\widehat{S}$ satisfies $\widehat{\gamma}$.*

Theorem 7.3.4. *Let S be a choice function satisfying $\widehat{\alpha}_{\subseteq}$. Then $\widehat{S}$ is stable if and only if $\widehat{S}$ is well-defined and S satisfies local $\widehat{\alpha}$.*

Proof. For the direction from right to left, suppose that $\widehat{S}$ is well-defined and S satisfies local $\widehat{\alpha}$. Then Lemma 7.3.3 implies that $\widehat{S}$ satisfies $\widehat{\gamma}$. Moreover, it follows directly from local $\widehat{\alpha}$ and the fact that $\widehat{S}$ is well-defined that $\widehat{S}$ satisfies $\widehat{\alpha}$. Hence, $\widehat{S}$ is stable.

For the converse direction, suppose that $\widehat{S}$ is stable. We first show that $\widehat{S}$ is well-defined.

Every feasible set A contains at least one S-stable set because otherwise $\widehat{S}$ is not a choice function. Next, suppose for contradiction that there exists a feasible set that contains two distinct minimal S-stable sets. Consider such a feasible set A of minimum size, and pick any two distinct minimal S-stable sets in A, which we denote by B and C. If $|B \backslash C| = |C \backslash B| = 1$, then $\widehat{\alpha}_{\subseteq}$ implies $S(B \cup C) = B = C$, a contradiction. Otherwise, assume without loss of generality that $|C \backslash B| \geq 2$, and pick $x, y \in C \backslash B$ with $x \neq y$. Then $A \backslash \{x\}$ contains a unique minimal S-stable set. As B is also S-stable in $A \backslash \{x\}$, it follows that $\widehat{S}(A \backslash \{x\}) \subseteq B$. Since $\widehat{S}$ satisfies $\widehat{\alpha}$, we have $\widehat{S}(A \backslash \{x\}) = \widehat{S}(A \backslash \{x, y\})$. Similarly, we have $\widehat{S}(A \backslash \{y\}) = \widehat{S}(A \backslash \{x, y\})$. But then $\widehat{\gamma}$ implies that $\widehat{S}(A) = \widehat{S}(A \backslash \{x, y\}) \subseteq A$, which contradicts the fact that C is minimal S-stable in A.

[7] It can be checked that we obtain an equivalent condition even if we require that *all* outside alternatives have to be removed. When defining local $\widehat{\alpha}$ in this way, it can be interpreted as some form of transitivity of stability: stable sets of minimally stable sets are also stable within the original feasible set (cf. [31, Lemma 3]).

We now show that S satisfies local $\widehat{\alpha}$. Since $\widehat{S}$ is well-defined, minimal S-stable sets are unique and given by $\widehat{S}$. Since $\widehat{S}$ satisfies $\widehat{\alpha}$, minimal S-stable sets are invariant under deleting outside alternatives. Hence, S satisfies local $\widehat{\alpha}$, as desired. $\square$

Remark 7.3.5. *Theorem 7.3.4 does not hold without the condition that S satisfies $\widehat{\alpha}_{\subseteq}$. To this end, let $U = \{a,b,c\}$, $S(\{a,b,c\}) = \{b\}$, and $S(X) = X$ for all other feasible sets X. Then both $\{a,b\}$ and $\{b,c\}$ are minimally S-stable in $\{a,b,c\}$, implying that $\widehat{S}$ is not well-defined. On the other hand, $\widehat{S}$ is trivial and therefore also stable. This example also shows that a generator of a stable choice function need not be sandwiched between the choice function and its root.*

Combining Theorem 7.3.4 with Theorem 7.2.3, we obtain the following characterization.

Corollary 7.3.6. *Let S be a choice function satisfying $\widehat{\alpha}_{\subseteq}$. Then,*

$$\begin{aligned} \widehat{S} \text{ is stable } &\text{if and only if } \widehat{\widehat{S}} \text{ is well-defined and } \widehat{\widehat{S}} = \widehat{S} \\ &\text{if and only if } \widehat{S} \text{ satisfies } \widehat{\alpha} \text{ and } \widehat{\gamma} \\ &\text{if and only if } \widehat{S} \text{ is well-defined and } S \text{ satisfies local } \widehat{\alpha}. \end{aligned}$$

Since simple choice functions trivially satisfy $\widehat{\alpha}_{\subseteq}$, this corollary completely characterizes which simple choice functions generate stable choice functions.

7.4 Tournament Solutions

We now turn to tournament solutions, an important special case of choice functions whose output depends on a binary relation. See Section 2.2 for an overview of tournaments and tournament solutions.

In this section, we define two tournament solutions that are central to this chapter. The first one, the bipartisan set, generalizes the notion of a Condorcet winner to probability distributions over alternatives. The *skew-adjacency matrix* $G(T) = (g_{ab})_{a,b \in A}$ of a tournament T is defined by letting

$$g_{ab} = \begin{cases} 1 & \text{if } a \succ b \\ -1 & \text{if } b \succ a \\ 0 & \text{if } a = b. \end{cases}$$

The skew-adjacency matrix can be interpreted as a symmetric zero-sum game in which there are two players, one choosing rows and the other choosing columns, and in which the matrix entries are the payoffs of the row player. Laffond et al. [93] and Fisher and Ryan [68] have shown independently that every such game admits a unique mixed Nash equilibrium, which moreover is symmetric. Let $p_T \in \Delta(A)$ denote the mixed strategy played by both players in equilibrium. Then, p_T is the unique probability distribution such that

$$\sum_{a,b\in A} p_T(a)q(b)g_{ab} \geq 0 \quad \text{for all } q \in \Delta(A).$$

In other words, there is no other probability distribution that is more likely to yield a better alternative than p_T. Laffond et al. [93] defined the bipartisan set $BP(T)$ of T as the support of p_T.[8]

Definition 7.4.1. *The* bipartisan set *(BP) of a given tournament $T = (A, \succ)$ is defined as*

$$BP(T) = \{a \in A \mid p_T(a) > 0\}.$$

BP satisfies stability, monotonicity, regularity, and composition-consistency. Moreover, $BP \subseteq UC$, and BP can be computed in polynomial time by solving a linear feasibility problem.

The next tournament solution, the tournament equilibrium set, was defined by Schwartz [131]. Given a tournament $T = (A, \succ)$ and a tournament solution S, a nonempty subset of alternatives $X \subseteq A$ is called *S-retentive* if $S(\overline{D}(x)) \subseteq X$ for all $x \in X$ such that $\overline{D}(x) \neq \emptyset$.

Definition 7.4.2. *The* tournament equilibrium set *(TEQ) of a given tournament $T = (A, \succ)$ is defined recursively as the union of all inclusion-minimal TEQ-retentive sets in T.*

This is a proper recursive definition because the cardinality of the set of dominators of an alternative in a particular set is always smaller than the cardinality of the set itself. BP and TEQ coincide on all tournaments of order 5 or less [39].[9]

Schwartz [131] showed that $TEQ \subseteq BA$ and conjectured that every tournament contains a *unique* inclusion-minimal TEQ-retentive set, which was later shown to be equivalent to TEQ satisfying any one of a number of desirable properties for tournament solutions including stability and monotonicity [31, 32, 33, 41, 78, 79, 94]. This conjecture was disproved by Brandt et al. [37], who have non-constructively shown the existence of a counterexample with about 10^{136} alternatives using the probabilistic method. Since it was shown that TEQ satisfies the above mentioned desirable properties for all tournaments that are smaller than the smallest counterexample to Schwartz's conjecture, the search for smaller counterexamples remains an important problem. In fact, the counterexample found by Brandt et al. [37] is so large that it has no practical consequences whatsoever for TEQ. Apart from concrete counterexamples, there is ongoing interest in why and under which circumstances TEQ and a related tournament solution called the *minimal extending set* $ME = \widehat{BA}$ violate stability [42, 113, 161].

[8] The probability distribution p_T was independently analyzed by Kreweras [91], Fishburn [66], Felsenthal and Machover [64], and others. An axiomatic characterization in the context of social choice was recently given by Brandl et al. [29].

[9] It is open whether there are tournaments in which BP and TEQ are disjoint.

Computing the tournament equilibrium set of a given tournament was shown to be NP-hard and consequently there does not exist an efficient algorithm for this problem unless P equals NP [40].

7.5 Stable Tournament Solutions and Their Generators

Tournament solutions comprise an important subclass of choice functions. In this section, we examine the consequences of the findings from Sections 7.2 and 7.3, in particular Theorems 7.2.3, 7.3.1, and 7.3.4, for tournament solutions.

Stability is a rather demanding property which is satisfied by only a few tournament solutions. Three well-known tournament solutions that satisfy stability are the top cycle TC, the minimal covering set MC defined by $MC = \widehat{UC}$, and the bipartisan set BP, which is a refinement of MC. By virtue of Theorem 7.3.1, any stable tournament solution is generated by its root $\lceil S \rceil$. For example, $\lceil TC \rceil$ is a tournament solution that excludes an alternative if and only if it is the only alternative not contained in the top cycle (and hence a Condorcet loser). Similarly, one can obtain the roots of other stable tournament solutions such as MC and BP. In some cases, the generator that is typically considered for a stable tournament solution is different from its root; for example, MC is traditionally generated by UC, a refinement of $\lceil MC \rceil$.

Since tournament solutions are invariant under tournament isomorphisms, a simple tournament solution may only exclude an alternative a if any automorphism of T maps a to itself. Note that if a tournament solution S is stable, $\lceil S \rceil$ is different from S unless S is the trivial tournament solution.

It follows from Theorem 7.2.3 that stable tournament solutions satisfy both $\widehat{\alpha}$ and $\widehat{\gamma}$. It can be shown that $\widehat{\alpha}$ and $\widehat{\gamma}$ are independent from each other even in the context of tournament solutions.

Remark 7.5.1. *There are tournament solutions that satisfy only one of $\widehat{\alpha}$ and $\widehat{\gamma}$. Examples are given in Appendix C.1.*

We have shown in Theorem 7.3.1 that stable tournament solutions are generated by unique simple tournament solutions. If we furthermore restrict our attention to *monotonic* stable tournament solutions, the following theorem shows that we only need to consider root solutions that are monotonic.

Theorem 7.5.2. *A stable tournament solution S is monotonic if and only if $\lceil S \rceil$ is monotonic.*

Proof. First, note that monotonicity is equivalent to requiring that unchosen alternatives remain unchosen when they are weakened. Now, for the direction from left to right, suppose that S is monotonic. Let $T = (A, \succ)$, $B = \lceil S \rceil(T)$, and $a \in A \setminus B$. Since $\lceil S \rceil$ is simple, we have $\lceil S \rceil(T) = B \backslash \{a\}$, and therefore $S(T) = B \backslash \{a\}$. Using the fact that S is stable and thus satisfies $\widehat{\alpha}$, we find that $S(T|_{B \backslash \{a\}}) = B \backslash \{a\}$. Let T' be a tournament obtained by weakening a with respect to some alternative in B. Monotonicity of S entails that $a \notin S(T')$. Since $T|_{B \backslash \{a\}} = T'|_{B \backslash \{a\}}$, we have

$S(T'|_{B\setminus\{a\}}) = B\setminus\{a\}$, and $\widehat{\alpha}$ implies that $S(T') = B\setminus\{a\}$ and $\lceil S\rceil(T') = B\setminus\{a\}$ as well. This means that a remains unchosen by $\lceil S\rceil$ in T', as desired.

The converse direction even holds for all generators of S (see [31, Proposition 5]). □

Analogous results do *not* hold for composition-consistency or regularity.

Theorem 7.3.4 characterizes stable choice functions $\widehat{S}$ using well-definedness of $\widehat{S}$ and local $\widehat{\alpha}$ of S. These two properties are independent from each other (and therefore necessary for the characterization) even in the context of tournament solutions.

Remark 7.5.3. *There is a tournament solution S that satisfies local $\widehat{\alpha}$, but $\widehat{S}$ violates $\widehat{\alpha}$. There is a tournament solution S for which $\widehat{S}$ is well-defined, but $\widehat{S}$ is not stable. Examples are given in Appendix C.2.*

Theorem 7.3.4 generalizes previous statements about stable tournament solutions. Brandt studied a particular class of generators defined via qualified subsets and shows the direction from right to left of Theorem 7.3.4 for these generators [31, Theorem 4].[10] Later, Brandt et al. proved Theorem 7.3.4 for one particular generator BA ([42, Corollary 2]).

7.6 Local Reversal Symmetry

In this section, we introduce a new property of tournament solutions called local reversal symmetry (*LRS*).[11] While intuitive by itself, *LRS* is strongly connected to stability and can be leveraged to disprove that *TEQ* is stable and to prove that no refinement of *BP* is stable.

For a tournament T, let T^a be the tournament whose dominance relation is *locally reversed* at alternative a, i.e., $T^a = (A, \succ^a)$ with

$$i \succ^a j \quad \text{if and only if} \quad (i \succ j \text{ and } a \notin \{i,j\}) \text{ or } (j \succ i \text{ and } a \in \{i,j\}).$$

The effect of local reversals is illustrated in Figure 7.3. Note that $T = (T^a)^a$ and $(T^a)^b = \left(T^b\right)^a$ for all alternatives a and b.

Definition 7.6.1. *A tournament solution S satisfies* local reversal symmetry *(LRS) if for all tournaments T and alternatives a,*

$$a \in S(T) \textit{ if and only if } a \notin S(T^a).$$

[10] Brandt's proof relies on a lemma that essentially showed that the generators he considers always satisfy local $\widehat{\alpha}$.

[11] The name of this axiom is inspired by a social choice criterion called *reversal symmetry*. Reversal symmetry prescribes that a uniquely chosen alternative has to be unchosen when the preferences of all voters are reversed [126]. A stronger axiom, called *ballot reversal symmetry*, which demands that the choice set is inverted when all preferences are reversed was recently introduced by Duddy et al. [56].

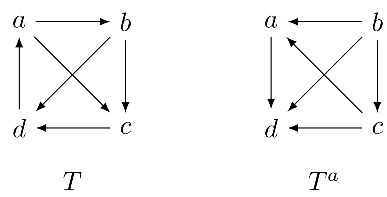

Figure 7.3: Local reversals of tournament T at alternative a result in T^a.

LRS can be naturally split into two properties, LRS_{IN} and LRS_{OUT}. LRS_{IN} corresponds to the direction from right to left in the above equivalence and requires that unchosen alternatives be chosen when all incident edges are reversed. LRS_{OUT} corresponds to the direction from left to right and requires that chosen alternatives be unchosen when all incident edges are reversed.

It follows directly from the definition that LRS_{IN} (resp. LRS_{OUT}) of a tournament solution S carries over to any tournament solution that is a coarsening (resp. refinement) of S.

Lemma 7.6.2. *Let S and S' be two tournament solutions such that $S \subseteq S'$. If S satisfies LRS_{IN}, then so does S'. Conversely, if S' satisfies LRS_{OUT}, then so does S.*

There is an unexpected strong relationship between the purely choice-theoretic condition of stability and LRS.

Theorem 7.6.3. *Every stable tournament solution satisfies LRS_{IN}.*

Proof. Suppose for contradiction that S is stable but violates LRS_{IN}. Then there exists a tournament $T = (A, \succ)$ and an alternative $a \in A$ such that $a \notin S(T)$ and $a \notin S(T^a)$.

Let $T' = (A', \succ')$, where $A' = X \cup Y$ and each of $T'|_X$ and $T'|_Y$ is isomorphic to $T|_{A\setminus\{a\}}$. Also, partition $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$, where X_1 and Y_1 consist of the alternatives that are mapped to $\overline{D}_T(a)$ by the isomorphism. To complete the definition of T', we add the relations $X_1 \succ' Y_2$, $Y_2 \succ' X_2$, $X_2 \succ' Y_1$, and $Y_1 \succ' X_1$. The structure of tournament T' is depicted in Figure 7.4.

We claim that both X and Y are externally S-stable in T'. To this end, we note that for every alternative $x \in X$ (resp. $y \in Y$) the subtournament $T|_{Y\cup\{x\}}$ (resp. $T|_{X\cup\{y\}}$) is isomorphic either to T or to T^a, with x (resp. y) being mapped to a. By assumption, a is neither chosen in T nor in T^a, and therefore X and Y are both externally S-stable in T'.

Now, suppose that $S(X \cup \{y\}) = X' \subseteq X$ for some $y \in Y$. Note that $X' \neq \emptyset$ because tournament solutions always return non-empty sets. Since S satisfies $\widehat{\alpha}$, we have $S(X) = X'$. Hence, $S(X) = X' = S(X \cup \{y\})$ for all $y \in Y$. Since S satisfies $\widehat{\gamma}$, we also have $S(X \cup Y) = X'$. Similarly, we can deduce that $S(X \cup Y) = Y'$ for some $\emptyset \neq Y' \subseteq Y$. This yields the desired contradiction. ☐

7.6.1 Disproving Stability

As discussed in Section 7.4, disproving that a tournament solution satisfies stability can be very difficult. By virtue of Theorem 7.6.3, it now suffices to show that the tournament solution violates

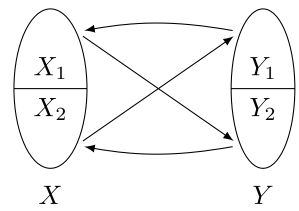

Figure 7.4: Construction of a tournament T' with two S-stable sets X and Y used in the proof of Theorem 7.6.3.

LRS_{IN}. For TEQ, this leads to the first concrete tournament in which TEQ violates stability. With the help of a computer, we found a minimal tournament in which TEQ violates LRS_{IN} using exhaustive search. This tournament is of order 13 and thereby lies exactly at the boundary of the class of tournaments for which exhaustive search is possible. Using the construction explained in the proof of Theorem 7.6.3, we thus obtain a tournament of order 24 in which TEQ violates $\widehat{\gamma}$. This tournament consists of two disjoint isomorphic subtournaments of order 12 both of which are TEQ-retentive.

Theorem 7.6.4. *TEQ violates LRS_{IN}.*

Proof. We define a tournament $T = (\{x_1, x_2, \ldots, x_{13}\}, \succ)$ such that $x_{13} \notin TEQ(T)$ and $x_{13} \notin TEQ(T^{x_{13}})$. The dominator sets in T are defined as follows:

$$\begin{array}{lcllcl}
\overline{D}(x_1) & = & \{x_4, x_5, x_6, x_8, x_9, x_{12}\}, & \overline{D}(x_2) & = & \{x_1, x_6, x_7, x_{10}, x_{12}\}, \\
\overline{D}(x_3) & = & \{x_1, x_2, x_6, x_7, x_9, x_{10}\}, & \overline{D}(x_4) & = & \{x_2, x_3, x_7, x_8, x_{11}\}, \\
\overline{D}(x_5) & = & \{x_2, x_3, x_4, x_8, x_{10}, x_{11}\}, & \overline{D}(x_6) & = & \{x_4, x_5, x_9, x_{11}, x_{12}\}, \\
\overline{D}(x_7) & = & \{x_1, x_5, x_6, x_{11}, x_{12}, x_{13}\}, & \overline{D}(x_8) & = & \{x_2, x_3, x_6, x_7, x_{12}, x_{13}\}, \\
\overline{D}(x_9) & = & \{x_2, x_4, x_5, x_7, x_8, x_{13}\}, & \overline{D}(x_{10}) & = & \{x_1, x_4, x_6, x_7, x_8, x_9, x_{13}\}, \\
\overline{D}(x_{11}) & = & \{x_1, x_2, x_3, x_8, x_9, x_{10}, x_{13}\}, & \overline{D}(x_{12}) & = & \{x_3, x_4, x_5, x_9, x_{10}, x_{11}, x_{13}\}, \\
\overline{D}(x_{13}) & = & \{x_1, x_2, x_3, x_4, x_5, x_6\}. & & &
\end{array}$$

A rather tedious check reveals that

$$\begin{array}{lcllcl}
TEQ(\overline{D}(x_1)) & = & \{x_4, x_8, x_{12}\}, & TEQ(\overline{D}(x_2)) & = & \{x_6, x_{10}, x_{12}\}, \\
TEQ(\overline{D}(x_3)) & = & \{x_6, x_7, x_9\}, & TEQ(\overline{D}(x_4)) & = & \{x_2, x_7, x_{11}\}, \\
TEQ(\overline{D}(x_5)) & = & \{x_2, x_8, x_{10}\}, & TEQ(\overline{D}(x_6)) & = & \{x_4, x_9, x_{11}\}, \\
TEQ(\overline{D}(x_7)) & = & \{x_1, x_5, x_{11}\}, & TEQ(\overline{D}(x_8)) & = & \{x_3, x_6, x_{12}\}, \\
TEQ(\overline{D}(x_9)) & = & \{x_2, x_5, x_7\}, & TEQ(\overline{D}(x_{10})) & = & \{x_4, x_6, x_7\}, \\
TEQ(\overline{D}(x_{11})) & = & \{x_1, x_2, x_8\}, \text{ and} & TEQ(\overline{D}(x_{12})) & = & \{x_3, x_4, x_9\}.
\end{array}$$

It can then be checked that $TEQ(T) = TEQ(T^{x_{13}}) = \{x_1, x_2, \ldots, x_{12}\}$. □

Let n_{TEQ} be the greatest natural number n such that all tournaments of order n or less admit a unique inclusion-minimal *TEQ*-retentive set. Together with earlier results by Brandt et al. [40] and Yang [161], we now have that $14 \leq n_{TEQ} \leq 23$.

The tournament used in the preceding proof does not show that *ME* (or *BA*) violate LRS_{IN}. A computer search for such tournaments was unsuccessful. While it is known that *ME* violates stability, a concrete counterexample thus remains elusive.

7.6.2 Most Discriminating Stable Tournament Solutions

An important property of tournament solutions that is not captured by any of the axioms introduced in Section 2.2 is discriminative power.[12] It is known that *BA* and *MC* (and by the known inclusions also *UC* and *TC*) almost always select all alternatives when tournaments are drawn uniformly at random and the number of alternatives goes to infinity [65, 132].[13] Experimental results suggest that the same is true for *TEQ*. Other tournament solutions, which are known to return small choice sets, fail to satisfy stability and composition-consistency. A challenging question is how discriminating tournament solutions can be while still satisfying desirable axioms.

LRS reveals an illuminating dichotomy in this context for common tournament solutions. We state without proof that discriminating tournament solutions such as Copeland's rule, Slater's rule, and Markov's rule satisfy LRS_{OUT} and violate LRS_{IN}. On the other hand, coarse tournament solutions such as *TC*, *UC*, and *MC* satisfy LRS_{IN} and violate LRS_{OUT}. The bipartisan set hits the sweet spot because it is the only one among the commonly considered tournament solutions that satisfies LRS_{IN} *and* LRS_{OUT} (and hence *LRS*).

Theorem 7.6.5. *BP satisfies LRS.*

Proof. Since *BP* is stable, Theorem 7.6.3 implies that *BP* satisfies LRS_{IN}. Now, assume for contradiction that *BP* violates LRS_{OUT}, i.e., there is a tournament $T = (A, \succ)$ and an alternative a such that $a \in BP(T)$ and $a \in BP(T^a)$. For a probability distribution p and a subset of alternatives $B \subseteq A$, let $p(B) = \sum_{x \in B} p(x)$. Consider the optimal mixed strategy $p_{T|_{A\setminus\{a\}}}$ in tournament $T|_{A\setminus\{a\}}$. It is known (see [98, Proposition 6.4.8]) that $a \in BP(T)$ if and only if $p_{T|_{A\setminus\{a\}}}(D(a)) > p_{T|_{A\setminus\{a\}}}(\overline{D}(a))$. For T^a, we thus have that $p_{T^a|_{A\setminus\{a\}}}(D(a)) > p_{T^a|_{A\setminus\{a\}}}(\overline{D}(a))$. This is a contradiction because $D_T(a) = \overline{D}_{T^a}(a)$ and $\overline{D}_T(a) = D_{T^a}(a)$. □

The relationship between *LRS* and the discriminative power of tournament solutions is no coincidence. To see this, consider all *labeled* tournaments of fixed order and an arbitrary alternative

[12]To see that discriminative power is not captured by the axioms, observe that the trivial tournament solution satisfies stability, monotonicity, regularity, and composition-consistency.

[13]However, these analytic results stand in sharp contrast to empirical observations that Condorcet winners are likely to exist in real-world settings, which implies that tournament solutions are much more discriminative than results for the uniform distribution suggest [43].

a. LRS_{IN} demands that a be chosen in *at least* one of T and T^a while LRS_{OUT} requires that a be chosen in *at most* one of T and T^a. We thus obtain the following consequences.

- A tournament solution satisfying LRS_{IN} chooses on average at least half of the alternatives.
- A tournament solution satisfying LRS_{OUT} chooses on average at most half of the alternatives.
- A tournament solution satisfying LRS chooses on average half of the alternatives.

Hence, the well-known fact that BP chooses on average half of the alternatives [67] follows from Theorem 7.6.5. Also, all coarsenings of BP such as MC, UC, and TC satisfy LRS_{IN} by virtue of Lemma 7.6.2. On the other hand, since these tournament solutions are all different from BP, they choose on average more than half of the alternatives and hence cannot satisfy LRS_{OUT}.

These results already hint at BP being perhaps a "most discriminating" stable tournament solution. In order to make this precise, we formally define the discriminative power of a tournament solution. For two tournament solutions S and S', we say that S is *more discriminating* than S' if there is a natural number n such that the average number of alternatives chosen by S is lower than that of S' over all labeled tournaments of order n. Note that this definition is very weak because we only have an existential, not a universal, quantifier for n. It is therefore even possible that two tournament solutions are more discriminating than each other. However, this only strengthens the following results. Combining Theorems 7.6.3 and 7.6.5 immediately yields the following theorem.

Theorem 7.6.6. *A stable tournament solution satisfies LRS if and only if there is no more discriminating stable tournament solution.*

Proof. First consider the direction from left to right. Let S be a tournament solution that satisfies LRS. Due to the observations made above, S chooses on average half of the alternatives. Since any stable tournament solution satisfies LRS_{IN} by Theorem 7.6.3, it chooses on average at least half of the alternatives and therefore cannot be more discriminating than S.

Now consider the direction from right to left. Let S be a most discriminating stable tournament solution. Again, since any stable tournament solution satisfies LRS_{IN}, S chooses on average at least half of the alternatives. On the other hand, BP is a stable tournament solution that chooses on average exactly half of the alternatives. This means that S must also choose on average half of the alternatives, implying that it also satisfies LRS_{OUT} and hence LRS. □

Corollary 7.6.7. *There is no more discriminating stable tournament solution than BP. In particular, there is no stable refinement of BP.*

Given Corollary 7.6.7, a natural question is whether every stable tournament solution that satisfies mild additional properties such as monotonicity is a coarsening of BP. We give an example in Appendix C.4 which shows that this is not true.

Finally, we provide two axiomatic characterizations of *BP* by leveraging other traditional properties. These characterizations leverage the following lemma, which entails that, in order to show that two stable tournament solutions that satisfy *LRS* are identical, it suffices to show that their roots are contained in each other.

Lemma 7.6.8. *Let S and S' be two stable tournament solutions satisfying LRS. Then $\lceil S \rceil \subseteq \lceil S' \rceil$ if and only if $S = S'$.*

Proof. Suppose that $\lceil S \rceil \subseteq \lceil S' \rceil$, and consider any tournament T. We will show that $S(T) \subseteq S'(T)$. If $S'(T) = T$, this is already the case. Otherwise, we have $S'(S'(T) \cup \{a\}) = S'(T)$ for each $a \notin S'(T)$. By definition of the root function, $\lceil S' \rceil(S'(T) \cup \{a\}) = S'(T)$. Since the root function excludes at most one alternative from any tournament, we also have $\lceil S \rceil(S'(T) \cup \{a\}) = S'(T)$ by our assumption $\lceil S \rceil \subseteq \lceil S' \rceil$. Hence $S(S'(T) \cup \{a\}) = S'(T)$ as well. Using $\widehat{\gamma}$ of S, we find that $S(T) = S'(T)$. So $S(T) \subseteq S'(T)$ for every tournament T. However, since S and S' satisfy *LRS*, and therefore choose on average half of the alternatives, we must have $S = S'$.

Finally, if $S = S'$, then clearly $\lceil S \rceil = \lceil S' \rceil$ and so $\lceil S \rceil \subseteq \lceil S' \rceil$. □

Theorem 7.6.9. *BP is the only tournament solution that satisfies stability, monotonicity, regularity, composition-consistency, and LRS.*

Proof. Let S be a tournament solution satisfying the five aforementioned properties. Since S and *BP* are stable and satisfy *LRS*, by Lemma 7.6.8 it suffices to show that $\lceil S \rceil \subseteq \lceil BP \rceil$. This is equivalent to showing that when $\lceil BP \rceil$ excludes an alternative from a tournament, then $\lceil S \rceil$ excludes the same alternative. In other words, we need to show that when *BP* excludes exactly one alternative a, then S also only excludes a.

Let T be a tournament in which *BP* excludes exactly one alternative a. As defined in Section 7.4, $BP(T)$ corresponds to the support of the unique Nash equilibrium of $G(T)$. Laffond et al. [93] and Fisher and Ryan [68] have shown that this support is always of odd size and that the equilibrium weights associated to the alternatives in $BP(T)$ are odd numbers. Hence, using composition-consistency of *BP* and the fact that the value of a symmetric zero-sum game is zero, T can be transformed into a new (possibly larger) tournament $T_1 = (A, \succ)$ by replacing each alternative except a with a regular tournament of odd order such that $T_1|_{A\setminus\{a\}}$ is regular. Moreover, in T_1, $|\overline{D}(a)| > \frac{|A|}{2}$.

We will now show that $a \notin S(T_1)$. Since S is monotonic, it suffices to prove this when we strengthen a arbitrarily against alternatives in T_1 until $|\overline{D}(a)| = \frac{|A|+1}{2}$.

Let $X = D(a)$ and $Y = \overline{D}(a)$, and let T_2 be a tournament obtained by adding a new alternative b to T_1 so that $X \succ \{b\} \succ Y$ and $a \succ b$. Note that T_2 is again a regular tournament, so $S(T_2) = A \cup \{b\}$. In particular, $b \in S(T_2)$.

Let $T_3 = (T_2)^b$ be the tournament obtained from T_2 by reversing all edges incident to b. By LRS_{OUT}, we have $b \notin S(T_3)$. If it were the case that $a \in S(T_3)$, then it should remain chosen when

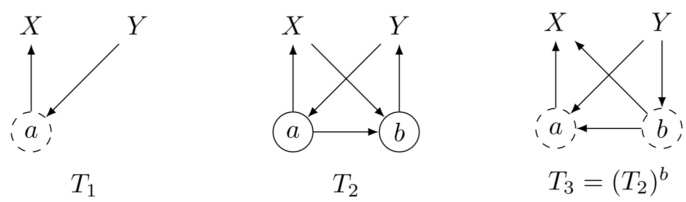

Figure 7.5: Illustration of the proof of Theorem 7.6.9. Circled alternatives are contained in the choice set $S(\cdot)$. Alternatives circled with a dashed line are not contained in the choice set $S(\cdot)$.

we reverse the edge between a and b. However, alternative a in the tournament after reversing the edge is isomorphic to alternative b in T_3, and so we must have $b \in S(T_3)$, a contradiction. Hence $a \notin S(T_3)$. Since S satisfies stability and thus $\widehat{\alpha}$, we also have $a \notin S(T_1)$, as claimed. See Figure 7.5 for an illustration.

Now, $\widehat{\alpha}$ and regularity of S imply that $S(T_1) = S(T_1|_{A\setminus\{a\}}) = A\setminus\{a\}$. Since S satisfies composition-consistency, we also have that S returns all alternatives except a from the original tournament T, completing our proof. □

Based on Theorems 7.6.6 and 7.6.9, we obtain another characterization that does not involve *LRS* and hence only makes use of properties previously considered in the literature.

Corollary 7.6.10. *BP is the unique most discriminating tournament solution that satisfies stability, composition-consistency, monotonicity, and regularity.*

Proof. Suppose that a tournament solution S satisfies stability, composition-consistency, monotonicity, and regularity and is as discriminating as BP. By Theorem 7.6.3, S satisfies LRS_{IN}. Since S chooses on average half of the alternatives, it satisfies LRS_{OUT} and hence LRS as well. Theorem 7.6.9 then implies that $S = BP$. □

The only previous characterization of BP that we are aware of was obtained by Laslier [98, Theorem 6.3.10] and is based on a rather contrived property called *Copeland-dominance*. According to Laslier [98, p. 153], "this axiomatization of the Bipartisan set does not add much to our knowledge of the concept because it is merely a re-statement of previous propositions." Corollary 7.6.10 essentially shows that, for most discriminating stable tournament solutions, Laslier's Copeland-dominance is implied by monotonicity and regularity.

We now address the independence of the axioms used in Theorem 7.6.9.

Remark 7.6.11. *LRS is not implied by the other properties. In fact, the trivial tournament solution satisfies stability, composition-consistency, monotonicity, and regularity.*

Remark 7.6.12. *Monotonicity is not implied by the other properties. In fact, the tournament solution that returns $BP(\overline{T})$, where $\overline{T}$ is the tournament in which all edges in T are reversed, satisfies stability, composition-consistency, regularity, and LRS.*

The question of whether stability, composition-consistency, and regularity are independent in the presence of the other axioms is quite challenging. We can only provide the following partial answers.

Remark 7.6.13. *Neither stability nor composition-consistency is implied by LRS, monotonicity, and regularity. In fact, there is a tournament solution that satisfies LRS, monotonicity, and regularity, but violates stability and composition-consistency (see Appendix C.3).*

Remark 7.6.14. *Neither regularity nor composition-consistency is implied by stability and monotonicity. In fact, there is a tournament solution that satisfies stability and monotonicity, but violates regularity and composition-consistency (see Appendix C.4).*

Brandt et al. [42] brought up the question whether stability implies regularity (under mild assumptions) because all stable tournaments solutions studied prior to this paper were regular.[14] Remark 7.6.14 shows that this does not hold without making assumptions that go beyond monotonicity.

Given the previous remarks, it is possible that composition-consistency and regularity are not required for Theorem 7.6.9 and Corollary 7.6.10. Indeed, our computer experiments have shown that the only stable and monotonic tournament solution satisfying LRS for all tournaments of order up to 7 is BP. This may, however, be due to the large number of automorphisms in small tournaments and composition-consistency and regularity could be required for larger tournaments. It is also noteworthy that the proof of Theorem 7.6.9 only requires a weak version of composition-consistency, where all components are tournaments in which all alternatives are returned due to automorphisms.

Since stability is implied by Samuelson's weak axiom of revealed preference or, equivalently, by transitive rationalizability, Corollary 7.6.10 can be seen as an escape from Arrow's impossibility theorem where the impossibility is turned into a possibility by weakening transitive rationalizability and (significantly) strengthening the remaining conditions (see also [41]).

7.7 Conclusion and Future Work

In this chapter, we provide several insights on the structure of stable choice functions and tournament solutions. We show that every stable choice function is generated by a unique simple choice function and completely characterize which simple choice functions give rise to stable choice functions. Furthermore, we exhibit the first concrete tournament in which the tournament equilibrium set fails to be stable and provide a characterization of the bipartisan set that uses only properties previously proposed in the literature.

An intriguing question that remains after our work is whether it is possible to strengthen our characterization of the bipartisan set. In particular, one potential strengthening of Corollary 7.6.7 is

[14] We checked on a computer that the stable tournament solution $\mathring{TC}$ (see [33, 34]) satisfies regularity for all tournaments of order 17 or less.

that the bipartisan set is the unique *inclusion-minimal* tournament solution that satisfies stability, composition-consistency, monotonicity, and regularity. Indeed, the only other tournament solutions we are aware of that satisfy all four properties are the trivial tournament solution and the minimal covering set, both of which are coarsenings of the bipartisan set. If this strengthening were to hold, it would provide a more compelling characterization of the bipartisan set.

Chapter 8

Who Can Win a Single-Elimination Tournament?

8.1 Introduction

In this chapter, we continue our study of methods for choosing the best alternatives from a tournament. One common way to select a winner among multiple alternatives is by using a *single-elimination (SE) tournament*, also known as a *knockout tournament*, a *sudden death tournament*, an *Olympic system tournament*, or a *binary-cup election*. In a SE tournament, pairs of alternatives are matched according to an initial seeding, the winners of these pairs advance to the next round, and the losers are eliminated after a single loss. Play continues according to the seeding until a single alternative, the winner, remains. While SE tournaments can be organized for political candidates or other types of alternatives, they are especially popular in sports competitions, both among fans due to their exciting "do-or-die" nature and among tournament organizers due to their efficiency.[1] In contrast to a round-robin tournament, which requires $\Theta(n^2)$ matches to be played between n players and will be our subject of study in the next chapter, the winner of a SE tournament is decided after a total of $n-1$ matches. In tournaments like the NCAA March Madness or the US Open Tennis Championships, which involve more than 64 teams each, the difference between a linear and quadratic number of matches is quite pronounced.

While the efficiency of SE tournaments is desirable, the winner of a given SE tournament can depend significantly on the initial seeding. A series of works [12, 76, 85, 97, 111, 140, 141, 158, 159] has investigated how easily the winner of SE tournaments can be manipulated simply by adjusting the seeding of the tournament. Formally, the problem is called the *tournament fixing problem (TFP)*, or the *agenda control problem for balanced knockout tournaments*. In TFP, we are given a set of

[1] For this reason, we will use the term "player" instead of "alternative" for the rest of this chapter.

players, a tournament relation between players (for each pair of players, who would win in a head-to-head matchup), and a player of interest; then, we are asked the following question: is there a seeding to a balanced SE tournament where our player of interest wins? TFP is known to be NP-hard [12], with the best-known algorithm running in $2^n \text{poly}(n)$ time [85]. Thus, unless P = NP, it is in general intractable to determine which players can win a SE tournament. Nevertheless, a number of works on TFP have produced "structural results" which argue that for certain classes of instances, one can find a winning seeding for our player of interest in polynomial (and often linear) time [85, 140, 158]. These structural results suggest that in many practical settings, the winner of a SE tournament is susceptible to manipulation, because many players have winning seedings that can be found efficiently. In addition, under reasonable probabilistic models for generating tournaments, these structural results have been shown to occur with high probability [140, 158], further suggesting that the worst-case hardness results may not apply to real-world instances. In other words, in many actual tournaments, it is feasible for SE tournament organizers to rig the outcome of the competition. Experiments have investigated this finding in practical settings [125].

In this chapter, we improve our understanding of conditions on the input tournament and player of interest that are sufficient for the player to be a SE winner. Many previous structural results involve the notion of a *king*, or a player that has distance at most 2 to every other player in the domination graph. In Section 8.3, we present a vast generalization of many of the known structural results involving kings, showing that essentially any "combination" of the known sufficient conditions for a king to be a SE winner is also sufficient for the king to be a winner (Theorem 8.3.1). Additionally, we extend the work on 3-*kings* (i.e., players who have distance at most 3 to every other player in the domination graph), introduced by Kim and Vassilevska Williams [85], and give a new set of sufficient conditions for a 3-king to win a SE tournament (Theorem 8.3.2).

In Section 8.4, we apply these and other known structural results to understand the relationship between SE winners and the winners according to other tournament solutions. In particular, Theorem 8.4.1 shows that the players selected by a number of well-studied tournament solutions are also SE winners. Another class of tournament solutions of interest was introduced by Laslier [98] as a natural extension of the Copeland set. In these "iterative matrix solutions", we consider the tournament matrix A (corresponding to the adjacency matrix of the underlying tournament graph); a player is included in the kth iterative matrix solution if they have the maximum number of "wins" in A^k. We give a new interpretation of this solution and use it to show that for sufficiently large tournaments, the players in the iterative matrix solutions will also be SE winners (Theorem 8.4.4).

In Section 8.5, we investigate probabilistic models for generating random tournaments and the resulting structure of such tournaments. In particular, we start by giving an improved result for tournaments generated by the Condorcet Random (CR) Model. The CR Model assumes an underlying order to players, where stronger players generally win against weaker players and are only upset with some small probability p. We demonstrate that in tournaments generated by the CR

Model, even when the probability of upset p is $\Theta(\ln n/n)$, with high probability every player in the tournament will have a winning seeding that can be discovered efficiently (Theorem 8.5.1). This upset rate p is optimal (up to constant factors) because a player needs to win $\log n$ matches in order to win a SE tournament.[2] Our result greatly improves on the previous best result by Vassilevska Williams [158], who proved an analogous claim for $p \in \Omega(\sqrt{\ln n/n})$. In light of this optimal result for the CR Model, we propose a new generative model for tournaments that aims to remove the structure which arises from assuming an underlying order of players and a consistent noise parameter. Despite the fact that the model may produce tournaments with largely arbitrary structure, we show a result similar to previous results on the CR Model (Theorem 8.5.2).

All of our results are constructive. In particular, we demonstrate that certain conditions are sufficient for a player to be a SE winner by giving algorithms, running in polynomial time, that output a seeding under which the player wins.

8.2 Preliminaries

We assume in this chapter that all SE tournaments are balanced and played amongst a power of two, $n = 2^k$ for some $k \geq 0$, players. The basic definitions and notation on tournaments and tournament solutions are introduced in Section 2.2. Our notation in this chapter differs slightly in that we view the tournament as a graph, and denote by V the set of players and E the set of edges in the tournament. For $i, j \in V$, we sometimes write $(i, j) \in E$ to mean $i \succ j$. Recall that the set of *kings* corresponds to the uncovered set of a tournament; it is known that this is equivalent to the set of players that have distance at most 2 to every other player [137]. Similarly, a *3-king* is a player who have distance at most 3 to every other player.

We now provide brief descriptions of tournament solutions that we consider in this chapter in addition to those introduced in Sections 2.2 and 7.4; for more details, we refer to [34, 98].

- The *Copeland set* is the set of players of maximum out-degree in the tournament.
- The *Markov set* can be thought of as the set of players who win the most matches, in expectation, in a "winner-stays" tournament, where play proceeds by repeatedly selecting a random player to play the previous winner.
- The *Slater set* of a tournament T is the set of players that are maximal elements in the strict linear orders that can be obtained from T by inverting as few edges as possible, i.e., in the strict linear orders that have the maximum number of edges in common with T.

[2] Any logarithm written without a base in this chapter is assumed to have base 2.

8.3 Structural Results

A number of results are known about conditions under which a player is guaranteed to be a SE winner [85, 141, 158]. Many of these results concern players who are kings. In particular, Vassilevska Williams [158] showed that a "superking"—a king v such that every player in $N_{in}(v)$ loses to at least $\log n$ players from $N_{out}(v)$—is always a SE winner. On the other hand, Stanton and Vassilevska Williams [140] showed a generalization they call a "king of high out-degree"—that is, a king with out-degree k who loses to fewer than k players that have out-degree greater than k—is always a SE winner. In this section, we generalize these results by combining their respective conditions. Moreover, we further explore the notion of 3-kings that was considered by Kim and Vassilevska Williams [85] and present an alternative condition under which a 3-king can win a SE tournament.

Before we proceed to the results, we make some remarks on the strength of the king condition. While the ability to reach any other player in the tournament in at most two steps might seem like a strong condition (which would limit the usefulness of our results), it is in fact not as strong as it may look at first sight. Indeed, every tournament contains a king, and in particular any player with the maximum number of wins in the tournament is always a king. In addition, if we generate a tournament by choosing the direction of each edge independently and uniformly at random, it is known that the set of kings is equal to the entire set of players with high probability [65].

Theorem 8.3.1. *Consider a tournament $T = (V, E)$ where $\mathcal{K} \in V$ is a king. Let $A = N_{out}(\mathcal{K})$ and $B = V \backslash (A \cup \{\mathcal{K}\}) = N_{in}(\mathcal{K})$. Suppose that B is a disjoint union of three (possibly empty) sets H, I, J such that*

1. *$|H| < |A|$;*
2. *$in_A(i) \geq \log |V|$ for all $i \in I$ (i.e., $out_A(i) \leq |A| - \log |V|$ for all $i \in I$);*
3. *$out(j) \leq |A|$ for all $j \in J$.*

Then $\mathcal{K}$ is a SE winner, and we can compute a winning seeding for $\mathcal{K}$ in polynomial time.

Note that the superking result [158] corresponds to the special case where $H = J = \emptyset$, while the "king of high out-degree" result [140] corresponds to the special case where $I = \emptyset$. Hence Theorem 8.3.1 is much stronger than previous results in the sense that each player in B only has to satisfy one of the three "reasons" why it is not strong, instead of having to adhere to any particular one.

Proof. We proceed by induction, arguing that we can construct a seeding where, in each round, the three properties listed above and the condition that $\mathcal{K}$ is a king are maintained as invariants. We will first take care of the cases where the tournament is small. If $|V| = 1$ or 2, B is empty and the result is clear.

Suppose that $|V| = 4$. If $|A| \geq 2$, the result follows from [140]. Otherwise $|A| = 1$, and $H = I = \emptyset$ and $|J| \leq 1$, which contradicts $|V| = 4$.

Suppose now that $|V| \geq 8$. If $|A| \leq 2$, then $|H| \leq 1$, $I = \emptyset$, and $|J| \leq 3$, which contradicts $|V| \geq 8$. If $I = \emptyset$, $H \cup J = \emptyset$, or $|A| \geq |V|/2$, the result follows from [140] and [158]. Hence we may assume from now on that $|V| \geq 8$, $3 \leq |A| < |V|/2$, $I \neq \emptyset$, and $H \cup J \neq \emptyset$.

We will present an algorithm to compute a winning seeding for $\mathcal{K}$. The algorithm matches the players for the first round of the tournament in such a way that the leftover tournament after the first round also satisfies the conditions of the theorem. The description of the algorithm is as follows.

1. Perform a maximal matching M_1 from A to H.
2. Since $|H| < |A|$, we have $A \backslash M_1 \neq \emptyset$. Perform a maximal matching M_2 (which might be an empty matching) from $A \backslash M_1$ onto $I \cup J$.
3. If A was not fully used in the two matchings, match an arbitrary unused player in A with $\mathcal{K}$. Else, choose an arbitrary player $a \in A \cap M_2$ and rematch it to $\mathcal{K}$.
4. Perform arbitrary matchings within A, H, and $I \cup J$.
5. If there are leftover players, there must be exactly two of them; match them to each other.

We prove the correctness of the algorithm by showing that the four invariants are maintained by the algorithm. Let V', A', B', H', I', J' denote the subsets of V, A, B, H, I, J that remain after the iteration.

1. $|H'| < |A'|$. We will show that $|H'| \leq |H|/2$ and $|A'| \geq |A|/2$. The claim follows since $|H| < |A|$. If $H = \emptyset$, then $|H'| < |A'|$ holds trivially, so we may assume that H is nonempty. At least one player in H is used in the matching M_1, so we have $|H'| \leq |H|/2$. We will show that the matching $M_1 \cup M_2$ consists of at least two pairs. Since there can be at most two pairs in the matching provided by the algorithm in which a player in A is beaten by a player outside of A (i.e., the pair in which a player in A is matched to $\mathcal{K}$ and the pair in which a player in A is matched in the final step of the algorithm for leftover nodes), it will follow that $|A'| \geq |A|/2$.

   If M_1 consists of at least two pairs, we are done. Suppose that M_1 consists of exactly one pair. Since $|V| \geq 8$, each player in I is beaten by at least three players in A. (Recall that I is nonempty.) One of these players is possibly used in M_1, and one is possibly used to match with $\mathcal{K}$, but that still leaves at least one player in A that beats a player in I. Hence $M_1 \cup M_2$ consists of at least two pairs, as desired.

2. $in_{A'}(i) \geq \log |V'|$ *for all* $i \in I$. Let $i \in I'$. Since M_2 is a maximal matching, every player that contributes to the in-degree of i in A survives the iteration, except possibly the player that is rematched to $\mathcal{K}$. Hence the in-degree of i in A' is at least $\log |V| - 1 = \log(|V|/2)$.

3. *$out(j) \leq |A'|$ for all $j \in J'$.* The condition is equivalent to $out_{B'}(j) < in_{A'}(j)$. Let $j \in J'$. We have either $in_{A'}(j) = in_A(j)$ or $in_{A'}(j) = in_A(j) - 1$, where the latter case occurs exactly when a player in A that beats j is rematched to $\mathcal{K}$. In the former case we immediately obtain $out_{B'}(j) < in_{A'}(j)$. In the latter case, A has been fully used in the two matchings before one player is rematched to $\mathcal{K}$. This means that j eliminates another player in B, and it follows that $out_{B'}(j) \leq out_B(j) - 1 < in_A(j) - 1 = in_{A'}(j)$.

4. *$\mathcal{K}$ is a king.* Let $b \in B'$. If $b \in H'$, then since M_1 is a maximal matching, b is beaten by some player in A'. If $b \in I'$, then since the second invariant is maintained, b is beaten by some player in A'. Otherwise $b \in J'$. Since the third invariant is maintained, b beats at most $|A'| - 1$ players in A', and hence b is also beaten by some player in A' in this case.

Hence the four invariants are maintained, and the algorithm correctly computes a winning seeding for $\mathcal{K}$. □

Thus, we have shown a very general result about kings that holds in tournaments on n players for *any* power of two n, answering an open problem posed by Stanton and Vassilevska Williams [141] to provide more general structural results that hold independently of the size of the tournament. (Some earlier results only hold for large n.)

Next, we consider the weaker notion of a 3-king. Prior work presented a set of conditions under which a 3-king is a SE winner [85]. One of their conditions is that there exists a perfect matching from the set of players that are reachable in exactly two steps from the 3-king $\mathcal{K}$ onto the set of players that are reachable in exactly three steps from $\mathcal{K}$. Here, we present a different set of conditions that does not include the requirement of a perfect matching.

Theorem 8.3.2. *Consider a tournament $T = (V, E)$ where $\mathcal{K} \in V$ is a 3-king. Let $A = N_{out}(\mathcal{K})$, $B = N_{out}(A) \cap N_{in}(\mathcal{K})$, and $C = N_{in}(\mathcal{K}) \backslash B$. Suppose that the following three conditions hold:*

1. $|A| \geq \frac{|V|}{2}$;

2. $A \succ B$;

3. $|B| \geq |C|$.

Then $\mathcal{K}$ is a SE winner, and we can compute a winning seeding for $\mathcal{K}$ in polynomial time.

Proof. If $|V| = 1, 2$, or 4, the result is clear. For $|V| \geq 8$, first perform a maximal matching from B to C and match $\mathcal{K}$ to an arbitrary player in A, and then pair off players within A. If $|A|$ is odd, then $A \cup \{\mathcal{K}\}$ matches evenly. Else, match the remaining $a \in A$ to some $b \in B$. We pair off players within each of B, C arbitrarily, and match the remaining pair between B and C if needed. After the round, $|A| \geq \frac{|V|}{4}$. Since the matching from B to C is nonempty, we still have that $|B| \geq |C|$ after the iteration. Moreover, since we applied a maximal matching, each player in C is still beaten by some player in B. Thus, the required conditions are maintained as invariant, and we can efficiently compute a winning seeding for $\mathcal{K}$. □

It would be interesting to investigate the extent to which we can weaken the strong second condition that all players in A beat all players in B. It should be noted that if any of the three conditions is removed, the theorem no longer holds. In particular, if the second condition is dropped, a counterexample by Kim and Vassilevska Williams [85] shows that for any constant $\epsilon > 0$, there is a tournament on n players where $\mathcal{K}$ is a 3-king who win against $(1-\epsilon)n$ players but cannot win a SE tournament. Given that the notion of a 3-king is significantly weaker than that of a king (recall that kings who beat at least $|V|/2$ players are SE winners), it seems reasonable to conjecture that a strong assumption such as the second condition (or the conditions seen in [85]) may be required to prove structural results for 3-kings.

8.4 Single-Elimination Winners and Tournament Solutions

In this section, we investigate the relationship between the set of SE winners and some traditional tournament solutions. Like tournament solutions, the ability to win a SE tournament provides us with a criterion with which we can distinguish between stronger and weaker players in a tournament.

Theorem 8.4.1. *A player chosen by the Copeland set, the Slater set, or the Markov set is a SE winner. A player in the bipartisan set with the highest Copeland score is also a SE winner.*

Proof. All four tournament solutions are contained in the uncovered set, meaning that a player from any of these sets is a king. Therefore, using a special case of Theorem 8.3.1 (or an earlier result of Vassilevska Williams [158]), it suffices to show that such a player wins against at least half of the remaining players. For the Copeland set this is trivial, while Laslier [98] and Laffond et al. [94] showed that any player in the Slater set, and any player in the bipartisan set with the highest Copeland score, beat at least half of the players. Next, we show that players from the Markov set win against at least half of the players.

Recall that the Markov set is defined to be the set of players of maximum probability in the stationary distribution of the Markov chain defined by the normalized Laplacian matrix $Q = (q_{ij})_{n\times n}$ of the Markov chain of the tournament, where $q_{ij} = 1/n$ if v_i beats v_j (0 otherwise) and $q_{ii} = out(v_i)/n$. Assume that the first player is in the Markov set. It follows that the probability associated with the first player in the eigenvector $p = (p_i)_{n\times 1}$ corresponding to the eigenvalue 1 is maximal.

Assume for contradiction that $q_{11} < \frac{1}{2}$. We then have

$$\begin{aligned} p_1 &= q_{11}p_1 + q_{12}p_2 + \cdots + q_{1n}p_n \\ &\leq q_{11}p_1 + q_{12}p_1 + \cdots + q_{1n}p_1 \\ &= 2q_{11}p_1 \\ &< p_1, \end{aligned}$$

a contradiction. □

It is not true that any player in the bipartisan set is always a SE winner. Indeed, consider a transitive tournament with at least four players, with the slight modification that the weakest player beats the strongest player. Then the former player is included in the bipartisan set even though she only beats one player and cannot be a SE winner.

Another family of tournament solutions is introduced by Laslier [98] as "iterative matrix solutions". Consider the tournament adjacency matrix $A = (a_{ij})$, in which $a_{ij} = 1$ if i beats j, and 0 otherwise. The Copeland score is given by $A\mathbf{1}$. For any positive integer k, we consider $A^k\mathbf{1}$ and include the player(s) with the maximum resulting score in our kth iterative tournament solution.

There is a natural interpretation of iterative matrix solutions as the number of paths of length k starting from each player. Any player in an iterative matrix solution belongs to the uncovered set. In other words, if the player v is covered by some w (i.e., $w \succ \{v\} \cup N_{out}(v)$), then v cannot be in the iterative matrix solution. Indeed, if v is covered by w, then the first steps of the paths starting from w form a superset of the first steps of the paths starting from v. On the other hand, it is not the case that any player in an iterative matrix solution always beats at least half of the remaining players, as shown by the following example.

Consider $k = 2$ and the tournament with player set $V = A \cup B \cup \{x\}$, where $A \approx rn$ and $B \approx (1-r)n$ with $\frac{1}{2} < r < \frac{1}{\sqrt{3}}$. Suppose that $A \succ x \succ B \succ A$, and A and B are close to regular. The Copeland scores of $a \in A, b \in B, x$ are $\frac{rn}{2}, \frac{(1+r)n}{2}, (1-r)n$, respectively. It follows that the iterative matrix scores of a, b, x are $\frac{r^2n^2}{4}, \frac{(1+r^2)n^2}{4}, \frac{(1-r^2)n^2}{2}$. This implies that x has the maximum iterative matrix score but beats fewer than half of the remaining players.

Nevertheless, we will show that for large enough tournaments, players in an iterative matrix solution are always SE winners. First we need the following two lemmas, the second of which immediately follows from the first.

Lemma 8.4.2. *In a tournament with n players, the minimum possible number of k-paths is $\binom{n}{k+1}$.*

Proof. In a transitive tournament, each subset of size $k+1$ gives rise to exactly one k-path. On the other hand, by a simple inductive argument, each subset of size $k+1$ gives rise to at least one k-path that goes through all $k+1$ players. The result follows immediately. □

Lemma 8.4.3. *In a tournament with n players, a player with the maximum number of k-paths originating from it is the origin of at least $\frac{1}{n}\binom{n}{k+1}$ k-paths.*

We are now ready to prove the theorem.

Theorem 8.4.4. *For any fixed k, there exists a constant N_k such that for any tournament of size at least N_k, a player with the maximum number of k-paths originating from it is a SE winner.*

Proof. Let v be a player with the maximum number of k-paths originating from it, and let A and B be the sets of players who lose to v and who beat v, respectively. From Lemma 8.4.3, v is the origin of at least $\frac{1}{n}\binom{n}{k+1} \geq \frac{n^k}{2(k+1)!}$ k-paths for large enough n. Hence it must have out-degree at least $\frac{n}{2(k+1)!}$. In other words, $|A| \geq \frac{n}{2(k+1)!}$.

If the number of players in B with in-degree from A less than $\log n$ is less than $|A|$, we can apply Theorem 8.3.1. Otherwise, there are at least $|A| \geq \frac{n}{2(k+1)!}$ players in B with in-degree from A less than $\log n$. Call this set H, and consider a player $h \in H$. Since h beats all but at most $\log n$ players in A, we can compare the number of k-paths originating from v with the number of k-paths originating from h by removing the common k-paths. The remaining number of k-paths originating from v is at most $\log n \cdot n^{k-1}$, while by Lemma 8.4.3 again, a player in H with the maximum number of k-paths within H is the origin of at least $O(n^k)$ k-paths, since $|H|$ is linear in n. This contradicts the assumption that v has the maximum number of k-paths originating from it. □

8.4.1 The Strength of Kings

Since results concerning SE winners often involve the assumption that a player is a king in the given tournament, one might hope that there is a strong relation between SE winners and the uncovered set. For example, it could be that a constant fraction of players in the uncovered set must be SE winners, or vice versa. This is not the case, however, as the following theorem shows.

Theorem 8.4.5. *Let $r \in (0,1)$. There exists a tournament such that the proportion of players in the uncovered set that are SE winners is less than r and the proportion of SE winners that are contained in the uncovered set is also less than r.*

Proof. Consider a tournament with player set $V = A \cup B \cup \{x, y\}$ such that

- $x \succ y, B$
- $y \succ B, A$
- $B \succ A$
- $A \succ x$.

The uncovered set is $A \cup \{x, y\}$.

Let $|A| = k$ and $|B| = n$. If $k < \log n$, then players in A do not win enough matches to become a SE winner. Hence the proportion of players in the uncovered set that are SE winners is at most $\frac{2}{k+2}$.

On the other hand, suppose that B is a regular tournament with all players isomorphic. By symmetry, if one player in B is a SE winner, then all of them are. In order for a player in B to be a SE winner, players x and y need to be eliminated. But this can easily be done in two rounds, with x beating y in the first round and a player in A beating x in the second round. Hence the proportion of SE winners that are contained in the uncovered set is at most $\frac{2}{n+2}$.

Taking k and n large enough with $k < \log n$, we obtain the desired result. □

8.5 Generative Models for Tournaments

Recall the Condorcet Random (CR) Model, studied by Braverman and Mossel [44], Vassilevska Williams [158], and Stanton and Vassilevska Williams [140]. In the CR Model, we assume that there is an underlying order of the players, and that, in general, stronger players win against weaker players; however, with some small probability $p < 1/2$, the weaker player upsets the stronger player. In the corresponding tournament graph, we say that for two players i, j such that i occurs before j in the ordering, $(i, j) \in E$ with probability $1 - p$ and $(j, i) \in E$ otherwise. A number of results are known about which players are SE winners in tournaments drawn from the CR Model. When $p \in \Omega(\sqrt{\ln n/n})$, then with high probability, every player in the tournament is a superking, and therefore a SE winner [158]. In fact, even when $p \geq C \ln n/n$, roughly the first half of players are SE winners, and more generally if $p = C \cdot 2^i \ln n/n$, then roughly the first $1 - 1/2^{i+1}$ fraction of players are SE winners [140]. Previous work has also studied various generalizations of the CR Model [85, 140].

In this section, we present improved results on tournaments generated by the standard CR Model. We show that with high probability, *every* player in a CR tournament is a SE winner, even with the noise $p = \Theta(\ln n/n)$ (with no dependence on the player's rank).

Theorem 8.5.1. *Let $C \geq 64$ be a constant and $p \geq C \ln n/n$. Let T be a tournament generated by the CR Model with noise parameter p on $n > n_C$ players (for some constant n_C). With probability $1 - 1/\Omega(n^2)$, every player has an efficiently-computable winning seeding over T.*

Note that this result is asymptotically optimal, as a player must have at least $\log n$ wins to be able to win a SE tournament. If $p = o(\ln n/n)$, then with high probability, the weakest player will not be able to win a SE tournament, regardless of the seeding. The case where $p \geq C\sqrt{\ln n/n}$ is covered by Vassilevska Williams [158], who showed that every player in such a tournament is a SE winner.

We give a sketch of the proof before proceeding to the full proof. First, we argue that the weakest player w will win against more than $k \log n$ players in the first half, for some constant k. We will think of "swapping" $k \log n$ of these losers, which we call S, from the first half with some arbitrary set of players from the second half (so that these losers become some of the strongest players over the second half). Then, we argue that at least one player v that w beats will be in the first $n/6$ players. This player, with high probability, will be a king over the first half of players, who wins against more than half of the players; thus, by [158], this player will be a SE winner over the first half of players. Next, we argue that for some arbitrary player u in the weaker half of players, at least $\log n$ players from the $k \log n$ that were swapped to the second half will beat u. We then take a union bound over the players in the second half, and argue that w will be a superking over the second half, and again by [158], a SE winner over the second half. Thus, w will be a SE winner over the entire tournament by winning over the weaker half, while v wins against the stronger half, and w wins against v in the final round. We take a union bound over all players to arrive at the desired result.

The detailed proof follows.

Proof of Theorem 8.5.1. Let $C \geq 64$ be a constant and $C \ln n/n \leq p \leq C\sqrt{\ln n/n}$. First, note that we expect w to win against $\frac{C}{2} \ln n = \frac{C \ln 2}{2} \log n$ players in the first half. Next, we can show that with high probability, w wins against more than $\frac{C \ln 2}{4} \log n$ players. Let $k = \frac{C \ln 2}{4}$. We have

$$\begin{aligned}
&\Pr[w \text{ wins against } > k \log n \text{ players in the first half}] \\
&\quad \geq 1 - \exp\left(-\frac{(k \log n)^2}{4k \log n}\right) \\
&\quad = 1 - \exp\left(-\frac{k \log n}{4}\right) \\
&\quad = 1 - \exp\left(-\frac{C \ln 2 \log n}{16}\right) \\
&\quad = 1 - 1/n^{C/16}.
\end{aligned}$$

We also argue that with probability at least $1 - 1/n^{C/6}$, w wins against some player v in the first $n/6$ players:

$$\begin{aligned}
&\Pr[w \text{ wins against some } v \in [1, n/6]] \\
&\quad = 1 - (1 - p)^{n/6} \\
&\quad \geq 1 - (1 - (C \ln n/6)/(n/6))^{n/6} \\
&\quad \geq 1 - \exp(-C \ln n/6) \\
&\quad = 1 - 1/n^{C/6},
\end{aligned}$$

where the inequality follows from the approximation $(1-a/x)^x \leq e^{-a}$ for $a > 0$.

In what follows, we will imagine swapping a set of $k \log n$ players, called S, whom w wins against from the first half (excluding v) with $k \log n$ arbitrary players from the second half. This allows us to argue about the "first half" and the "second half" of players independently. We will argue that v is a SE winner over the new "first half" of players, and that the inclusion of $k \log n$ strong players whom w beats makes w a superking over the new "second half".

First, we argue that it is likely that v, whose rank is at most $n/6$, will be a SE winner over the first half. In particular, with high probability, v will be a king over the first half of players who wins against at least $n/4$ players. Note that we expect v to win against at least $n/3 \cdot (1-p) + pn/6 - 1 \geq n/3 - C\sqrt{n \ln n}/6 - 1$ players from the first half. The out-degree of v is given by a random variable, which is the sum of independent random variables, so we can bound the probability that $out(v) < n/4$ using the Chernoff bound (Lemma 2.3.1):

$$\begin{aligned}\Pr[out(v) \geq n/4] &\geq 1 - \exp\left(-\frac{(n/12 - \frac{C}{6}\sqrt{n \ln n} - 1)^2}{2(n/3 - \frac{C}{6}\sqrt{n \ln n} - 1)}\right) \\ &> 1 - 1/n^4,\end{aligned}$$

where the last inequality is a very loose bound on this probability that takes effect for sufficiently large n.

Next, we consider the probability that v is a king over the first half, conditioned on its high out-degree. We take a union bound over all possible players who did not lose against v, and show that it is unlikely that any of these players beats every single player whom v beats.

$$\begin{aligned}&\Pr[v \text{ is a king over the first half } \mid out(v) \geq n/4] \\ &\geq 1 - \sum_{i=1}^{n/4-1} (1-p)^{out(v)} \\ &\geq 1 - n/4 \cdot (1-p)^{n/4} \\ &\geq 1 - n/4 \cdot \exp(-C \ln n/4) \\ &\geq 1 - 1/4n^{C/4-1},\end{aligned}$$

Finally, we argue that with high probability, w will be a superking over the second half of players. Consider some other u from the second half of players. The expected number of players from S who beat u is $\geq k \log n \cdot (1-p) = k \log n - \frac{kC \log^{3/2} n}{\sqrt{n}} \geq (k-1) \log n$ for sufficiently large n. Applying the Chernoff bound (Lemma 2.3.1) again, we obtain the following bound.

$$\begin{aligned}
&\Pr[u \text{ loses to fewer than } \log n \text{ players from } S] \\
&\quad \le \exp\left(-\frac{((k-2)\log n)^2}{2(k-1)\log n}\right) \\
&\quad = \exp\left(-\frac{(k^2-4k+4)}{2(k-1)}\log n\right) \\
&\quad = n^{-\left(\frac{k^2-4k+4}{2\ln 2\cdot(k-1)}\right)}.
\end{aligned}$$

Then, to guarantee that every u in the second half loses to at least $\log n$ players whom w beats, we take a union bound over the $n/2$ players. For any $k > 11$, this probability will be $\le 1/n^3$.

The overall probability that w beats a sufficiently strong king over the first half of players is at least the following:

$$1 - 1/n^{C/6} - 1/n^4 - 1/4n^{C/4-1} \ge 1 - 2/n^4.$$

Thus, the probability that w wins against $k \log n$ players from the (true) first half, wins against some strong king v over the first half, and is a superking over the second half, is at least the following:

$$1 - 1/n^{C/16} - 2/n^4 - 1/n^3 \ge 1 - 2/n^3.$$

Since w is the weakest player of the tournament, the probability that any other player is a SE winner can only be greater. Taking a union bound over all players, we conclude that with probability at least $1 - 1/\Omega(n^2)$, every player in the tournament will be a SE winner. □

8.5.1 Generalizing the Condorcet Random Model

As the prior claims demonstrate, in the standard CR Model, every player is a SE winner with high probability, even when upsets occur at an asymptotically minimal rate. While this result indicates the depth of our understanding of conditions under which a player is a SE winner, it also suggests that the assumption that tournaments are drawn from a CR Model—where the noise parameter p is fixed for all matchups—may be too rigid, incidentally providing structure that may not exist in practical settings. Prior work [140] proposed a Generalized CR Model, where for two players $i < j$, j upsets i with probability $p \le p(i,j) \le 1/2$, for some globally specified p. But even this model asserts that the probability of upsets for *every* edge must occur within the range of $[p, 1/2]$. We aim to relax our restrictions even further in order to disrupt this inherent structure in the CR Model.

Consider the following generative model, which is parameterized by a noise factor $p < 1/2$ and a participation factor $\Delta \le 1/2$. The tournament on n players is generated as follows: pick any set of pairs of players E' satisfying the condition that each player appears in at least $(1/2 + \Delta)n$ such pairs; then, for every pair $\{u, v\} \in E'$, pick (u, v) with probability $p_{u,v} \in [p, 1-p]$, and (v, u)

otherwise. The probabilities $p_{u,v}$ can be arbitrary as long as they are in $[p, 1-p]$. The remaining edges between players may be set arbitrarily. In this new model, many typical arguments used in analyzing CR tournaments, including those used in the proof of Theorem 8.5.1, which hinge on the precise definition of the CR Model, break down.

Note that unlike the CR Model, the new model does not start with an underlying ordering of players; however, such an ordering can easily be emulated. For instance, to emulate the CR Model, simply choose an ordering σ, set $\Delta = 1/2$, and for all u, v such that $\sigma(u) < \sigma(v)$, sample (u, v) with probability $1-p$. That said, because the model does not start with an explicit ordering, it is much more versatile. Moreover, because only a $(1/2 + \Delta)$ fraction of the edges are determined randomly, known structures can be (adversarially) hard-coded into the resulting graphs. In this sense, any results that we can show about tournaments generated from this model are extremely general and will apply broadly. Despite this generality, we are able to give a statement for our model mirroring that of Vassilevska Williams [158] for the CR Model.

Theorem 8.5.2. *Let $p > c\sqrt{\frac{\log n}{2\Delta n}}$ for some $c > 5$. Then with probability $> 1 - 1/\Omega(n^{(c-5)/2\ln 2})$, every player in a tournament T sampled from the aforementioned model has an efficiently-computable winning seeding over T.*

The proof of Theorem 8.5.2 is similar to the proof of the analogous statement about the CR Model found in [158]. It argues that with high probability, every player in the tournament is a superking.

Proof. Let $p = c\sqrt{\frac{\log n}{2\Delta n}}$. We will argue that with high probability all nodes in a randomly sampled tournament are superkings, so by [158] they will be SE winners. Let $T = (V, E)$ be a randomly sampled tournament. We will bound the probability that $v \in V$ is not a superking, namely, the probability that there exists some $u \in V \setminus \{v\}$ such that u loses to fewer than $\log n$ players whom v beats.

Let $u \in V \setminus \{v\}$. Let A_v be the set of players w for which the edge between v and w was sampled randomly with probability in the range $[p, 1-p]$. Define A_u analogously. We let $W = A_v \cap A_u$ be the players whose relation is sampled randomly for both v and u. We can lower bound the size of this intersection as $|W| \geq (1/2 + \Delta)n - 1 + (1/2 + \Delta)n - 1 - (n-2) = 2\Delta n$. Now, note that the expected number of edges from v into W is the sum of the probabilities that (v, w) is an edge for each $w \in W$, and thus is at least $2\Delta np$. Applying the Chernoff bound (Lemma 2.3.1), we can bound the probability that this set of edges into W is smaller than $c \log n/p = 2\Delta np/c$:

$$\begin{aligned}
&\Pr\left[\text{number of edges from } v \text{ into } W \le \frac{2\Delta np}{c}\right] \\
&\le \exp\left(-(1-1/c)^2 \Delta np\right) \\
&= \exp\left(-(1-1/c)^2 c\sqrt{\Delta n \log n/2}\right) \\
&= 2^{-\Omega(\sqrt{n \log n})}.
\end{aligned}$$

Now, we condition on the assumption that v beats at least $c \log n/p$ players from W. Note that each of these players beat u with probability $\ge p$, so we expect $\ge c \log n$ of these players to beat u. Thus, using the Chernoff bound (Lemma 2.3.1) again, we can bound the probability that u does not lose to at least $\log n$ of these players:

$$\begin{aligned}
&\Pr[\text{number of edges from } W \text{ into } u \le \log n] \\
&\le \exp\left(-(1-1/c)^2 c \log n/2\right) \\
&= n^{-(1-1/c)^2 c/2 \ln 2}.
\end{aligned}$$

Letting $C = (1-1/c)^2 c/2 \ln 2 - 2$, by a union bound over v's opponents, the probability that v is not a superking is at most $2^{-\Omega(\sqrt{n \log n})} + n^{-C-1}$. Applying another union bound over all players, the probability that some player is not a superking is at most $2^{\Omega(\sqrt{n \log n})} + n^{-C} \le O(n^{-C})$. Hence with probability $1 - 1/\Omega(n^C)$, all players are superkings. The result then follows from the assumption that $C \ge (c-5)/2 \ln 2$. □

8.6 Conclusion and Future Work

In this chapter, we establish results that shed light on the manipulability of SE tournaments. We show that the winner of such a tournament can be fixed when the underlying tournament graph satisfies a rather general set of conditions, and apply our results to understand the relationship between SE winners and other common tournament solutions. We also investigate probabilistic models for generating random tournaments and prove asymptotic bounds on the probabilities that allow all players to be SE winners in the resulting tournament.

It is worth noting that our results, as well as those of several works cited in this chapter, are based on a model where we know with certainty which player would win if two players were to meet. This strong assumption is not satisfied by most tournaments in practice. As we mentioned earlier, however, TFP is already NP-hard under this deterministic model [12]. As a result, in a generalization where we are given the probability distribution of the pairwise match outcomes, we cannot hope to obtain a finite approximation of the maximum winning probability for each player.

Nevertheless, it might still be interesting to see whether we reach similar conclusions, either in the form of structural or probabilistic results, in models that more closely reflect tournaments in the real world.

Chapter 9

Scheduling Asynchronous Round-Robin Tournaments

9.1 Introduction

Besides a single-elimination tournament, another popular format for organizing sports competitions is a *round-robin tournament*, also known as an *all-play-all tournament*. In a round-robin tournament, every pair of players play each other a fixed number of times during the competition. Since every player competes with every other player, the winner of a round-robin tournament is usually thought of as depending much less on luck than that of a single-elimination tournament. A series of work has investigated how to schedule a round-robin tournament when different notions are central to the organizers' consideration. One line of research has focused on time-relaxed tournaments, which takes into account the issue of time off between games involving the same player [87, 88, 130], while another has considered fairness issues [45, 46, 156, 162]. We refer the interested reader to a survey by Rasmussen and Trick [122] and a book by Anderson [8] for more details on the literature.

In this chapter, we study the problem of scheduling *asynchronous* round-robin tournaments, i.e., round-robin tournaments in which no two games take place at the same time. There are a number of reasons why it might be desirable to schedule all games at different times. Indeed, this tournament format allows spectators to follow all the games live, and the organizers can maximize revenue while having to organize the same number of games. Tournaments may even need to be asynchronous if there is only one venue where a game can take place. An example of an asynchronous round-robin tournament is the 2012 Premier League Snooker in England, in which five players in the group stage play a total of ten games in ten different weeks (albeit in ten different venues as well).

When scheduling an asynchronous round-robin tournament, the organizers may desire properties that improve the quality and fairness of the tournament. Unlike in single-elimination tournaments, for which the organizers can significantly impact the outcome of the tournament by setting up a bracket of their choice,[1] the set of games to be played in a round-robin tournament cannot be changed. Nevertheless, the order in which the games are played can still be an important factor in a round-robin tournament. For example, when players have a longer rest between games, they are more likely to have a relaxing rest and perform at their full potential in the next game. On the other hand, if some player has a long rest going into a game while her opponent has just played her previous game, the former player could have a significant advantage. Another desirable property of a schedule is that at any point during the tournament, all players should have played roughly the same number of games. This prevents the advantage of knowing too many results involving other players and the possibility of collusion. We define measures that capture all of these properties, and exhibit schedules that perform (close to) optimally with regard to these measures. In particular, we show that the schedule generated by the well-known "circle design" performs well with respect to all three measures when the number of players is even, but not so well when the number of players is odd. We also propose a different schedule that performs optimally with respect to all three measures when the number of players is odd. We hope that this schedule will be of practical interest to organizers of asynchronous round-robin tournaments.

A related problem that is worth mentioning is the problem of finding balanced tournament designs, which has been considered by some prior work [21, 75, 128]. In the setting of balanced tournament designs, it is assumed that there exist external factors that make some games different from others, and it is desirable that players receive roughly the same effect from these external factors. For instance, the tournament might involve games during different times of the day or at different venues. Since some players may be more familiar with playing in the morning than in the evening, or with playing at one venue than another, the aim of a balanced tournament design is to eliminate or minimize the potential advantage by scheduling players to play as evenly across the different times and venues as possible. By contrast, in our setting there is no inherent difference between games. Indeed, a reasonable example to keep in mind throughout this chapter is that the games of the tournament are scheduled on consecutive days, one game per day, at a single venue.

A summary of our results can be found in Table 9.1.

9.2 Preliminaries

We assume that the tournament in consideration is a single round-robin tournament, i.e., every pair of players play each other exactly once. As we will mention in Section 9.5, however, several of our results can be generalized to arbitrary round-robin tournaments as well.

Let n denote the number of players in the tournament. We divide the games of the tournament into r rounds of g games, where the first round comprises the first g games, the second round the

[1] See Chapter 8.

	Circle method, n even	Any schedule, n even	Circle method, n odd	Any schedule, n odd
Guaranteed rest time	$(n-4)/2$	$\leq (n-4)/2$	$(n-5)/2$	$\leq (n-3)/2$
Games-played difference index	1	≥ 1	2	≥ 1
Rest difference index	1 if $n = 4$; 2 if $n \geq 6$	≥ 1	$(n+1)/2$	≥ 1

Table 9.1: Summary of our results. All bounds are known to be attainable except that for the rest difference index when n is even. See also Section 9.5 for further discussion.

next g games, and so on. The parameters r and g depend on n and are given by

$$g = \left\lfloor \frac{n}{2} \right\rfloor,$$

$$r = 2 \cdot \left\lceil \frac{n}{2} \right\rceil - 1 = \begin{cases} n & \text{if } n \text{ is odd;} \\ n-1 & \text{if } n \text{ is even.} \end{cases}$$

A player is said to play in slot i in a round if she plays the ith game of that round. We emphasize that in asynchronous tournaments, rounds do not carry any particular meaning in the implementation of the tournament and are defined merely for the convenience of our analysis.

A single round-robin tournament consists of $\binom{n}{2} = \frac{n(n-1)}{2}$ games. Each player plays $n-1$ games, and we have the identity

$$r \cdot g = \left(2 \cdot \left\lceil \frac{n}{2} \right\rceil - 1\right) \cdot \left\lfloor \frac{n}{2} \right\rfloor = \frac{n(n-1)}{2}.$$

A well-known method for scheduling a round-robin tournament, described for instance by Haselgrove and Leech [75], is called the *circle design*. The method works as follows. Assume first that n is even. We arrange the players into two rows of $n/2$ players in such a way that the two rows align player by player. The games in the first round correspond to the pairs of players that are aligned in this arrangement. For asynchronous tournaments, we read the games from left to right. To generate the games in the next round, we keep the top-left player fixed and rotate the remaining players one step counterclockwise. (It is also possible to rotate the remaining players one step *clockwise*, but this results in the same schedule as rotating counterclockwise under appropriate renaming of the players.) We perform the rotation $n-2$ times to generate the games in all $n-1$ rounds. If n is odd, we simply pretend that the top-left player is a dummy player, and whichever player is matched to that player "sits out" the round (i.e., gets a bye in that round). The first three rounds for the tournaments with $n = 10$ and $n = 11$ are shown in Figures 9.1 and 9.2, respectively.

We now define three measures of a schedule for an asynchronous tournament that concern the quality and fairness of the tournament. The first measure, guaranteed rest time, considers the

1	2	3	4	5	1	10	2	3	4	1	9	10	2	3
10	9	8	7	6	9	8	7	6	5	8	7	6	5	4

Figure 9.1: The first three rounds generated by the circle design for a tournament with $n = 10$.

1	2	3	4	5	11	1	2	3	4	10	11	1	2	3
10	9	8	7	6	9	8	7	6	5	8	7	6	5	4

Figure 9.2: The first three rounds generated by the circle design for a tournament with $n = 11$. Note that one player "sits out" each round (i.e., gets a bye in that round).

minimum amount of time that the schedule allows players to take a rest before their next game.

Definition 9.2.1. *The* guaranteed rest time *of a schedule for an asynchronous tournament is the maximum integer b such that in the schedule, any two games involving a player is separated by at least b games not involving that player.*

A schedule with a high guaranteed rest time is desirable, as it allows players to take a long rest and prepare themselves for the next game. The higher the guaranteed rest time, the more likely we will see players perform at their full potential in the tournament.

The next two measures, the games-played difference index and the rest difference index, reflect the fairness of the schedule.

Definition 9.2.2. *The* games-played difference index *of a schedule for an asynchronous tournament is the minimum integer p such that at any point in the schedule, the difference between the number of games played by any two players is at most p.*

It is evident that for any tournament with at least three players, the games-played difference index is at least 1. A schedule with a low games-played difference index ensures that all players have played roughly the same number of games at any point during the tournament. This prevents the advantage that some players may have if they know the results of too many games involving other players. Indeed, with this knowledge the players can adjust their strategy to achieve their desired position in the tournament and may even conspire with one another to do so.

Definition 9.2.3. *The* rest difference index *of a schedule for an asynchronous tournament is the minimum integer d such that for any game in the schedule, if one player has not played in i_1 consecutive games since her last game and the other player has not played in i_2 consecutive games since her last game, then $|i_1 - i_2| \leq d$. (To handle the situation in which a player is playing her first game of the tournament, we assume that all players are involved in an imaginary game that takes place one slot before the first game of the schedule.)*

It is again evident that for any tournament with at least three players, the rest difference index is at least 1. A schedule with a low rest difference index guarantees that the two players involved in a game have approximately the same amount of rest time going into the game.

9.3 Even Number of Players

In this section, we assume that the tournament in consideration consists of an even number of players. We will show that under this assumption, the schedule generated by the circle design fares extremely well with respect to all of the measures introduced in Section 9.2. Since the round-robin tournament with two players consists of a single game, we will only consider $n \geq 4$.

We begin by showing an upper bound on the guaranteed rest time.

Proposition 9.3.1. *Let $n \geq 4$. Any schedule for a tournament with $n = 2k$ players has guaranteed rest time at most $k - 2$.*

Proof. Assume for the sake of contradiction that the guaranteed rest time is at least $k - 1$. This means that all players play at most once in any k consecutive games. Since there are $2k$ players, each player plays exactly once in the first round. Hence the first game in the second round must involve the same two players as the first game in the first round, a contradiction. □

Next, we analyze the schedule generated by the circle design.

Proposition 9.3.2. *Let $n \geq 4$. The schedule generated by the circle design for a tournament with $n = 2k$ players has guaranteed rest time $k - 2$, games-played difference index 1, and rest difference index 2 if $n \geq 6$ and 1 if $n = 4$.*

Proof. We verify each of the measures separately.

- *Guaranteed rest time*: Note that each player plays exactly once in every round. Since the slot of a player is shifted by at most 1 from one round to the next, and each round consists of k slots, the player has a rest of at least $k - 2$ games. On the other hand, a player whose slot is shifted to the left has a rest of exactly $k - 2$ games.

- *Games-played difference index*: Since each player plays exactly once in every round, all of the players have played the same number of games at the end of each round. It follows that the index is 1.

- *Rest difference index*: One can directly verify that the index is 1 if $n = 4$. Assume now that $n \geq 6$, and consider the second slot in the second round. One of the players in that slot is shifted from the third slot in the first round and the other player from the first slot in the first round. Hence the index is at least 2. On the other hand, the slot of a player is shifted by at most 1 from one round to the next, so the index is exactly 2.

We have verified all three measures. □

Propositions 9.3.1 and 9.3.2 together imply that the schedule generated by the circle design has an optimal guaranteed rest time and an optimal games-played difference index. Moreover, the rest

1	3	1	2	1	4	1	2	5	1	2	3	2	3	4
2	4	5	6	3	5	6	3	6	4	5	6	4	5	6

1	3	5	1	1	3	1	2	1	2	3	2	2	4	4
2	4	6	3	5	6	6	4	4	6	5	3	5	6	5

Figure 9.3: Two schedules with rest difference index 1 for a tournament with $n = 6$.

difference index can be improved by at most 1. We now show that unless $n = 4$, it is impossible to simultaneously obtain a guaranteed rest time of $k-2$ and games-played difference and rest difference indices of 1.

Theorem 9.3.3. *Let $n \geq 6$. No schedule for a tournament with $n = 2k$ players has guaranteed rest time $k-2$, games-played difference index 1, and rest difference index 1.*

Proof. Assume for the sake of contradiction that such a schedule exists. We first claim that in the schedule, each player plays exactly once in every round. This can be shown by induction on the number of rounds. Suppose that each player plays exactly once in every round up to round $i \geq 0$. In round $i+1$, if some player plays twice, then some other player does not play at all, contradicting the fact that the games-played difference index is 1. Hence each player also plays exactly once in round $i+1$, completing the induction.

Suppose now that in the first round, players 1 and 2 play in the first game, players 3 and 4 in the second, and players 5 and 6 in the third. Since the guaranteed rest time is $k-2$, the first game in the second round can only involve players from the first two games in the first round. Assume without loss of generality that player 1 and 3 play in that game. Similarly, the second game in the second round can only involve players from the first three games in the first round. The game cannot be played between player 4 and one of players 5 and 6, since the game involving player 2 in the second round would violate the rest difference index condition. Hence the game is played between players 2 and 4.

By the same reasoning, the first game in the third round must be played by players 1 and 4, and the second by players 2 and 3. But then no player can play against player 1 in the fourth round without violating the rest difference index condition. Thus we have the desired contradiction. □

Theorem 9.3.3 implies that if a schedule were to have rest difference index 1, it would have to sacrifice either the guaranteed rest time or the games-played difference index. Nevertheless, it is interesting to ask whether there exists for all even n a schedule with rest difference index 1. For $n = 6$, two such schedules are shown in Figure 9.3. The first schedule also has an optimal guaranteed rest time of 1, but makes the necessary sacrifice by having a games-played difference index of 2. On the other hand, the second schedule is worse off in both measures, having guaranteed rest time 0 and games-played difference index 3.

9.4 Odd Number of Players

In this section, we assume that the tournament in consideration consists of an odd number of players. We will show that unlike in the case where the number of players is even, the schedule generated by the circle design does not fare so well with respect to the measures introduced in Section 9.2. Nevertheless, we will exhibit a different schedule that performs optimally with respect to all of the measures.

The round-robin tournament with three players consists of three games, and any two schedules of the three games are equivalent under renaming of the players, so we have no choice to make in this case.

We begin by showing an upper bound of $k-1$ for the guaranteed rest time of any schedule.

Proposition 9.4.1. *Let $n \geq 3$. Any schedule for a tournament with $n = 2k+1$ players has guaranteed rest time at most $k-1$.*

Proof. Consider the first $k+1$ games of the tournament. Since they involve the participation of $2k+2$ players (counting multiplicity), the pigeonhole principle implies that some player plays at least twice among those games. Such a player has a rest of at most $k-1$ games. □

Next, we analyze the schedule generated by the circle design. Even though this schedule does not match the bound in Proposition 9.4.1, we will later exhibit a different schedule that does attain the upper bound.

Proposition 9.4.2. *Let $n \geq 5$. The schedule generated by the circle design for a tournament with $n = 2k+1$ players has guaranteed rest time $k-2$, games-played difference index 2, and rest difference index $k+1$.*

Proof. We verify each of the measures separately.

- *Guaranteed rest time*: Note that each player plays at most once in every round. Since the slot of a player is shifted by at most 1 from one round to the next, and each round consists of k slots, the player has a rest of at least $k-2$ games. On the other hand, a player whose slot is shifted to the left has a rest of exactly $k-2$ games.

- *Games-played difference index*: Since the player that sits out each round is distinct, the difference between the highest and lowest number of games played by a player at the end of each round is 1. Each player plays at most once in every round, so the difference increases by at most 1 during a round. Hence the index is at most 2.

  On the other hand, consider the point after the first game in the third round has just finished. The player that sat out the second round has played once, while a player involved in the first game of the third round has played three times. It follows that the index is 2.

1	3	1	2	4	1	2	3	1	2
2	4	5	3	5	3	4	5	4	5

Figure 9.4: The schedule as described in Theorem 9.4.4 with guaranteed rest time 1, games-played difference index 1, and rest difference index 1 for a tournament with $n = 5$.

1	3	5	1	2	4	6	1	2	4	3	1	2	4	3	1	2	5	3	1	2
2	4	6	7	3	5	7	3	5	6	7	5	6	7	5	6	4	7	6	4	7

1	3	5	1	2	4	6	1	2	4	1	3	2	4	1	3	2	1	5	3	2
2	4	6	7	3	5	7	3	5	7	6	5	7	6	5	7	6	4	7	6	4

Figure 9.5: Two schedules with guaranteed rest time 2, games-played difference index 1, and rest difference index 1 for a tournament with $n = 7$. The first schedule corresponds to the one described in Theorem 9.4.4.

- *Rest difference index*: Consider the first slot in the third round. One of the players in that slot was last involved in the first game of the first round, while the other player played in the second slot of the second round. Hence the index is at least $k + 1$.

  On the other hand, consider any two players involved in a game. If the two players also played in the previous round, the difference in their rest time is at most $k - 1$. Otherwise, one of the player sat out the previous round. This implies that the player played the first game of the round before the previous round, while the other player played the second game of the previous round. Hence the index is exactly $k + 1$.

We have verified all three measures. □

We now show that if a schedule attains the upper bound on the guaranteed rest time, it will also fare optimally with respect to the rest difference index.

Lemma 9.4.3. *Let $n \geq 3$. Any schedule for a tournament with $n = 2k + 1$ players with guaranteed rest time $k - 1$ has rest difference index 1.*

Proof. Suppose that a schedule for a tournament with $2k + 1$ players has guaranteed rest time $k - 1$. This means that any k consecutive games in the schedule are played by $2k$ distinct players.

We show that the rest difference index is 1. Consider an arbitrary game after the kth game. This game cannot involve a player that played in one of the previous $k - 1$ games. Moreover, the game cannot be played between the two players that played each other k games ago. Hence the only possibility is that the game is played between the player that sat out the previous k games and one of the two players that played k games ago. In particular, all $2k + 1$ players appear in any block of $k + 1$ consecutive games. This implies that the player that sat out the previous k games played $k + 1$ games ago (if this game exists). Hence the rest difference index is 1, as desired. □

Proposition 9.4.1 and Lemma 9.4.3 do not carry much meaning on their own. Indeed, without an example to show that the bounds can be achieved, it is difficult to tell how useful the bounds are. In particular, the rest difference index of the schedule generated by the circle design ($k+1$) is quite far from the bound we have so far (1). All of these observations raise the natural question of whether there exist other schedules that perform better on some or all measures. The next theorem gives the most satisfying answer possible to this question; it shows that there exists a schedule that fare optimally—and strictly better than the circle-design schedule—with respect to all three measures.

Theorem 9.4.4. *Let $n \geq 3$. There exists a schedule for a tournament with $n = 2k+1$ players with guaranteed rest time $k-1$, games-played difference index 1, and rest difference index 1.*

Proof. In light of Lemma 9.4.3, it suffices to show the existence of a schedule for a tournament with $2k+1$ players with guaranteed rest time $k-1$ and games-played difference index 1. We exhibit the schedule by specifying the slot that the players play in each round. Slots are taken modulo $k+1$, and slot 0 means that a player sits out that round. The schedule is defined as follows.

- For $1 \leq i \leq k$, player $2i-1$ is placed in slot i in the first $2i$ rounds. After that, the player moves forward by one slot in each round.
- For $1 \leq i \leq k$, player $2i$ is placed in slot i in the first round. The player moves forward by one slot in each round until round $2k+3-2i$. After that, she stays in the same slot until the last round.
- Player $2k+1$ is placed in slot $\lfloor j/2 \rfloor$ in the jth round.

The resulting schedules for the tournaments with $n = 5$ and $n = 7$ can be seen in Figures 9.4 and 9.5, respectively.

We show that the schedule is well-defined by demonstrating that every pair of players play each other exactly once. We divide the verification into cases.

- For $1 \leq i \neq j \leq k$, players $2i-1$ and $2j-1$ play each other in round $i+j$.
- For $1 \leq i \neq j \leq k$, players $2i$ and $2j$ play each other in round $2k+3-i-j$.
- For $1 \leq i \leq j \leq k$, players $2i-1$ and $2j$ play each other in round 1 if $i = j$ and round $2k+2+i-j$ otherwise.
- For $1 \leq i < j \leq k$, players $2i$ and $2j-1$ play each other in round $j-i+1$.
- For $1 \leq i \leq k$, players $2i-1$ and $2k+1$ play each other in round $2i$.
- For $1 \leq i \leq k$, players $2i$ and $2k+1$ play each other in round $2k+2-i$.

Next, we show that the guaranteed rest time is $k-1$. If a player sits out a round between two of her games, then the two games are separated by at least k other games. Otherwise, a player either stays in the same slot or moves one slot forward in the next round. In both cases, the player has a rest of at least $k-1$ games in between.

Finally, we show that the games-played difference index is 1. At the end of each round, the difference between the highest and lowest number of games played by a player is at most 1. The players with a lower number of games played are exactly those that already sat out a round. Hence it suffices to show that in any round, a player that already sat out a round appears no later than a player that participated in every round. One can check that player $2k+1$, who sat out the first round, appears no later than any player that did not sit out, and any other player that already sat out appears no later than her. This completes the proof of the claim, and therefore the theorem. □

The schedule described in Theorem 9.4.4 is not necessarily the unique schedule satisfying the desired properties. Indeed, for $n = 7$, another schedule satisfying the desired properties is shown in Figure 9.5. To see that the two schedules cannot be obtained from each other by permuting the player indices, observe that the first two rounds of games uniquely determine the identity of the players: player 1 plays in games 1 and 4, player 2 plays in games 1 and 5, player 3 plays in games 2 and 5, and so on. Since the two schedules differ in the second game of the fourth round, no permutation of player indices in one schedule results in the other schedule.

Now that Theorem 9.4.4 gives us a schedule that fare optimally on all three measures, we may demand a stronger notion of fairness. In particular, while the rest difference index of 1 guarantees that two players going into a game have roughly the same amount of rest, it seems fairer if all players sometimes get a longer rest than their opponent and sometimes a shorter one than if some players always get a longer rest than their opponent. Nevertheless, the following proposition shows that as long as we insist on maximal guaranteed rest time, this goal cannot be achieved.

Proposition 9.4.5. *Let $n \geq 3$. For any schedule for a tournament with $n = 2k+1$ players with guaranteed rest time $k-1$, there exists a player that has a longer rest time than her opponent in every game after her first game.*

Proof. Consider a schedule for a tournament with $n = 2k+1$ players with guaranteed rest time $k-1$. As in the proof of Lemma 9.4.3, we find that any game after the kth game is played between the unique player that sat out the previous k games and one of the two players that played k games ago. This implies that if a player just played a game and still has more games left in the tournament, then she will have a rest of either $k-1$ or k games before her next game. Put differently using the terminology in the proof of Theorem 9.4.4, a player either stays in the same slot or moves one slot forward in the next round. Since the number of rounds, $2k+1$, is equal to the number of players, each player sits out exactly one round.

Suppose that players 1 and 2 play each other in the first game of the tournament, and player 2 has a rest of k games before her next game. We claim that player 2 moves one slot forward in every round. It suffices to prove this claim in order to establish the theorem, since the claim implies that player 2 has a longer rest time than her opponent in every game after her first game.

Assume for the sake of contradiction that player 2 stays in the same slot at some point during the tournament. Consider the first instance in which this occurs. Since every player sits out exactly one slot, the slot is not slot 0.

Suppose that player 2 repeats a slot in rounds i and $i+1$. This means that the player that plays against player 2 in round $i+1$ (say, player t) played in the slot before player 2 in round i. Since the two players play each other only once during the tournament, player t also played in the slot before player 2 in round $i-1$, round $i-2$, and so on down to round 2. Hence player t sat out the first round, played against player 1 in the first slot of the second round, and is in the slot ahead of player 1 in the third round. This implies that player 1 cannot "overtake" player t in the slot position for the rest of the tournament. But since player t already sat out while player 1 did not, this means that player 1 cannot sit out for the rest of the tournament, a contradiction. □

9.5 Conclusion and Future Work

In this chapter, we define three measures that capture quality and fairness properties of a schedule for an asynchronous round-robin tournament, and we exhibit schedules that perform (close to) optimally with respect to all of these measures. Here we give some comments and directions for future work.

Several of our results can be generalized to arbitrary round-robin tournaments in which every pair of players play each other a fixed number of times. Indeed, we can turn a single round-robin tournament into an arbitrary round-robin tournament by duplicating each round a desired number of times. This method preserves the guaranteed rest time and rest difference index, and for Proposition 9.3.2 it also preserves the games-played difference index. Moreover, Propositions 9.3.1 and 9.4.1 can be generalized to arbitrary round-robin tournaments as well.

As mentioned in Section 9.3, an interesting open question is whether there exists a schedule with rest difference index 1 when there are an even number of players. Such schedules are shown in Figure 9.3 for the case $n = 6$. If the answer turns out to be affirmative, one could also ask for a schedule with a "balanced" rest difference in the sense described before Proposition 9.4.5, i.e., players sometimes get a longer rest than their opponent and sometimes a shorter one. In addition, one could ask for the optimal value of one measure when the remaining two are forced to achieve their optimal values. From Proposition 9.3.2 and Theorem 9.3.3, we know that when the guaranteed rest time and the games-played difference index achieve their optimal values, the minimum rest difference index is 2. When the number of players is odd, it would be interesting to explore whether it is possible to achieve a rest difference index of 1 with a balanced rest difference if we are willing to sacrifice on other measures.

Finally, it might be worth investigating the structure of "optimal" schedules: how many there are, and whether they differ between themselves in some other meaningful way for the players. This could potentially yield new insights into the fascinating study of scheduling round-robin tournaments.

Appendix A

Omitted Proofs from Chapter 3

A.1 Direct Proof of Theorem 3.3.1 for Two Agents

We give a proof of Theorem 3.3.1 for the case $n = 2$ that does not rely on Kneser's conjecture. Denote by $\succeq_1$ and $\succeq_2$ the preferences on $\mathcal{S}$ of the two agents. We establish the existence of a set of size at most $\lfloor \frac{m+2}{2} \rfloor$ that is agreeable to both agents; the tightness of the bound follows in the same way as in our proof of Theorem 3.3.1 for any number of agents.

Assume first that $m = 2k+1$ is odd. Suppose for contradiction that no subset of size at most $k+1$ is agreeable to both agents. Let $T \subseteq S$ be such that $|T| = k$. We begin by proving the following claim.

Claim: If $T \succ_1 S \backslash T$, then

$$(T \cup \{x\}) \backslash \{x'\} \succ_1 ((S \backslash T) \backslash \{x\}) \cup \{x'\}$$

for any $x \in S \backslash T$ and $x' \in T$.

Proof of Claim: Suppose that $T \succ_1 S \backslash T$, $x \in S \backslash T$, and $x' \in T$. It follows from monotonicity that $T \cup \{x\} \succ_1 (S \backslash T) \backslash \{x\}$. Since no subset of size $k+1$ is agreeable to both agents, we have $(S \backslash T) \backslash \{x\} \succ_2 T \cup \{x\}$. By monotonicity again, we have

$$((S \backslash T) \backslash \{x\}) \cup \{x'\} \succ_2 (T \cup \{x\}) \backslash \{x'\}.$$

Using again the assumption that no subset of size $k+1$ is agreeable to both agents, it follows that

$$(T \cup \{x\}) \backslash \{x'\} \succ_1 ((S \backslash T) \backslash \{x\}) \cup \{x'\},$$

and our claim is proved. □

We now use our claim to obtain the desired contradiction. Assume without loss of generality that

$\{x_1, x_2, \ldots, x_k\} \succ_1 \{x_{k+1}, x_{k+2}, \ldots, x_{2k+1}\}$. Applying our claim repeatedly to move items between the two sets, we find that

$$\{x_{k+1}, x_2, \ldots, x_k\} \succ_1 \{x_1, x_{k+2}, \ldots, x_{2k+1}\},$$

$$\{x_{k+1}, x_{k+2}, x_3, \ldots, x_k\} \succ_1 \{x_1, x_2, x_{k+3}, \ldots, x_{2k+1}\},$$

and so on, until finally

$$\{x_{k+1}, x_{k+2}, \ldots, x_{2k}\} \succ_1 \{x_1, x_2, \ldots, x_k, x_{2k+1}\}.$$

By monotonicity, we have $\{x_{k+1}, x_{k+2}, \ldots, x_{2k+1}\} \succ_1 \{x_1, x_2, \ldots, x_k\}$, which contradicts our assumption that $\{x_1, x_2, \ldots, x_k\} \succ_1 \{x_{k+1}, x_{k+2}, \ldots, x_{2k+1}\}$.

Assume now that $m = 2k$ is even. Let S' be the set of all items in S except x_1. We know from the case of m odd that there exists a subset $T \subseteq S'$ of size at most k such that $T \succeq_1 S' \backslash T$ and $T \succeq_2 S' \backslash T$. Since preferences are monotonic, we have that $T \cup \{x_1\} \succeq_1 S' \backslash T$ and $T \cup \{x_1\} \succeq_2 S' \backslash T$. This means that the set $T \cup \{x_1\}$ of size at most $k + 1$ is our desired subset, completing the proof.

Note that this proof also yields a polynomial-time algorithm to compute an agreeable set of size at most $\left\lfloor \frac{m+2}{2} \right\rfloor$ that is agreeable to both agents. Assume that $m = 2k+1$ is odd; the case m even can be handled similarly. Let $T \subseteq S$ be an arbitrary subset of size k. If $S \backslash T \succeq_1 T$ and $S \backslash T \succeq_2 T$, we are done. Otherwise, assume without loss of generality that $T \succ_1 S \backslash T$, and choose arbitrarily $x \in S \backslash T$ and $x' \in T$. As in the proof of the claim, if $T \cup \{x\} \succeq_2 (S \backslash T) \backslash \{x\}$, or if $(S \backslash T) \backslash \{x\} \succ_2 T \cup \{x\}$ and $((S \backslash T) \backslash \{x\}) \cup \{x'\} \succeq_1 (T \cup \{x\}) \backslash \{x'\}$, we are done. Hence we may assume as in the conclusion of the claim that $(T \cup \{x\}) \backslash \{x'\} \succ_1 ((S \backslash T) \backslash \{x\}) \cup \{x'\}$. This means that we can find an agreeable subset by moving elements repeatedly between the two sets as in the continuation of the proof. Since we need to move elements at most k times, our algorithm runs in polynomial time.

A.2 NP-hardness of Balanced 2-Partition

We show that Balanced 2-Partition is NP-hard via a reduction from 2-Partition, a well-known NP-hard problem.

Lemma A.2.1. Balanced 2-Partition *is NP-hard.*

Proof. We reduce from 2-Partition, a problem in which a multiset B of positive integers is given and the goal is to decide whether there exists a multiset $T \subseteq B$ such that $\sum_{b \in T} b = \sum_{b \in B \backslash T} b$. 2-Partition is known to be NP-complete (see, e.g., [70]).

Given a 2-Partition instance, we create a Balanced 2-Partition instance as follows. Let A be the multiset containing all elements of B and $|B|$ additional zeros. Clearly, the reduction runs in polynomial time. We show that B is a YES instance of 2-PARTITION if and only if A is a YES instance of BALANCED 2-PARTITION.

(YES Case) Suppose that B is a YES instance of 2-PARTITION, i.e., there exists $T \subseteq B$ such that $\sum_{b \in T} b = \sum_{b \in B \setminus T} b$. Let $S \subseteq A$ be the multiset containing all elements of T and $|B| - |T|$ additional zeros. Clearly, $|S| = |B| = |A|/2$ and $\sum_{a \in S} a = \sum_{b \in T} b = \sum_{b \in B} b/2 = \sum_{a \in A} a/2$, meaning that A is a YES instance of BALANCED 2-PARTITION as desired.

(NO Case) We prove the contrapositive; suppose that A is a YES instance of BALANCED 2-PARTITION. This means that there exists $S \subseteq A$ of size $|A|/2 = |B|$ such that $\sum_{a \in S} a = \sum_{a \in A \setminus S} a$. Let T be the subset of B containing all elements of B whose corresponding elements are included in S. Clearly, we have $\sum_{b \in T} b = \sum_{a \in S} a = \sum_{a \in A \setminus S} a = \sum_{b \in B \setminus T} b$. Hence B is a YES instance of 2-PARTITION. $\square$

Appendix B

Omitted Proofs from Chapter 4

B.1 Proof of Theorem 4.3.1

First we list the following well-known fact, which allows us to easily determine the mean of a random variable from its cumulative density function.

Lemma B.1.1. *Let X be a non-negative random variable. Then*

$$\mathbb{E}[X] = \int_0^\infty \Pr[X \geq x] dx.$$

To analyze the algorithm, consider any agent a_{ij} and any group $A_{i'} \neq A_i$. We will first bound the probability that $u_{ij}(G_{i'}) > u_{ij}(G_i)$. To do this, for each good $g \in G$, define $B_{ij,g}$ as

$$B_{ij,g} = u_{ij}(g) \cdot \mathbf{1}\left[i = \underset{q=1,2,\ldots,k}{\arg\max} \sum_{p=1}^{n'} u_{qp}(g)\right],$$

where $1\left[i = \arg\max_{q=1,2,\ldots,k} \sum_{p=1}^{n'} u_{qp}(g)\right]$ is an indicator random variable that indicates whether $i = \arg\max_{q=1,2,\ldots,k} \sum_{p=1}^{n'} u_{qp}(g)$. Similarly, define $C_{ij,g}^{i'}$ as

$$C_{ij,g}^{i'} = u_{ij}(g) \cdot \mathbf{1}\left[i' = \underset{q=1,2,\ldots,k}{\arg\max} \sum_{p=1}^{n'} u_{qp}(g)\right].$$

Moreover, suppose that $\mathcal{D}_g$ has mean μ_g and variance σ_g^2.

Notice that, with respect to agent a_{ij}, $B_{ij,g}$ is the utility that good g contributes to A_i whereas $C_{ij,g}^{i'}$ is the utility that good g contributes to $A_{i'}$. In other words, $u_{ij}(G_{i'}) > u_{ij}(G_i)$ if and only if $\sum_{g\in G} B_{ij,g} < \sum_{j\in M} C_{ij,g}^{i'}$. To bound $\Pr\left[u_{ij}(G_{i'}) > u_{ij}(G_i)\right]$, we will first bound $\mathbb{E}\left[B_{ij,g}\right]$ and

$\mathbb{E}\left[C^{i'}_{ij,g}\right]$. Then, we will use the Chernoff bound to bound $\Pr\left[\sum_{g\in G} B_{ij,g} < \sum_{j\in M} C^{i'}_{ij,g}\right]$.

Observe that, due to symmetry, we can conclude that

$$\mathbb{E}\left[u_{ij}(g)\cdot\mathbf{1}\left[i' = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right] = \mathbb{E}\left[u_{ij}(g)\cdot\mathbf{1}\left[i'' = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right]$$

for any $i'' \neq i$. Thus, we can now rearrange $C^{i'}_{ij,g}$ as follows:

$$\begin{aligned}
\mathbb{E}\left[C^{i'}_{ij,g}\right] &= \frac{1}{k-1}\left(\sum_{i''\neq i}\mathbb{E}\left[u_{ij}(g)\cdot\mathbf{1}\left[i'' = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right]\right)\\
&= \frac{1}{k-1}\left(\mathbb{E}\left[u_{ij}(g)\sum_{i''\neq i}\mathbf{1}\left[i'' = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right]\right)\\
&= \frac{1}{k-1}\left(\mathbb{E}\left[u_{ij}(g)\left(1-\mathbf{1}\left[i = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right)\right]\right).
\end{aligned}$$

Hence, we have

$$\mathbb{E}\left[C^{i'}_{ij,g}\right] = \frac{1}{k-1}\left(\mu_g - \mathbb{E}\left[B_{ij,g}\right]\right). \tag{B.1}$$

Now, consider $B_{ij,g}$. Again, due to symmetry, we have

$$\begin{aligned}
\mathbb{E}\left[B_{ij,g}\right] &= \frac{1}{n'}\left(\sum_{j=1}^{n'}\mathbb{E}\left[u_{ij}(g)\cdot\mathbf{1}\left[i = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right]\right)\\
&= \frac{1}{n'}\mathbb{E}\left[\left(\sum_{j=1}^{n'} u_{ij}(g)\right)\cdot\mathbf{1}\left[i = \operatorname*{arg\,max}_{q=1,2,\dots,k}\sum_{p=1}^{n'} u_{qp}(g)\right]\right].
\end{aligned}$$

Let $\mathcal{S}$ denote the distribution of the sum of n' independent random variables, each drawn from $\mathcal{D}_g$. It is obvious that $\sum_{p=1}^{n'} u_{qp}(g)$ is drawn from $\mathcal{S}$ independently for each q. In other words, $\mathbb{E}\left[B_{ij,g}\right]$ can be written as

$$\mathbb{E}\left[B_{ij,g}\right] = \frac{1}{n'}\mathbb{E}\left[X_1\cdot\mathbf{1}\left[X_1 = \max\{X_1, X_2, \dots, X_k\}\right]\right].$$

The expectation on the right is taken over $X_1, X_2, \dots, X_k$ sampled independently from $\mathcal{S}$.

From symmetry among $X_1, X_2, \ldots, X_k$, we can further derive the following:

$$\begin{aligned}\mathbb{E}\left[B_{ij,g}\right] &= \frac{1}{n'} \Pr\left[X_1 = \max\{X_1, X_2, \ldots, X_k\}\right] \mathbb{E}\left[X_1 \mid X_1 = \max\{X_1, X_2, \ldots, X_k\}\right] \\ &= \frac{1}{n'k} \mathbb{E}\left[X_1 \mid X_1 = \max\{X_1, X_2, \ldots, X_k\}\right] \\ &= \frac{1}{n'k} \mathbb{E}\left[\max\{X_1, X_2, \ldots, X_k\}\right].\end{aligned}$$

Consider the distribution of $\max\{X_1, X_2, \ldots, X_k\}$. Let us call this distribution $\mathcal{Y}$. Notice that $\mathbb{E}\left[\max\{X_1, X_2, \ldots, X_k\}\right]$ is just the mean of $\mathcal{Y}$, i.e.,

$$\mathbb{E}\left[B_{ij,g}\right] = \frac{1}{n'k} \mathop{\mathbb{E}}_{Y\sim\mathcal{Y}}[Y]. \tag{B.2}$$

To bound this, let F_S and F_Y be the cumulative density functions of $\mathcal{S}$ and $\mathcal{Y}$ respectively. Notice that $F_Y(x) = F_S(x)^k$ for all x. Applying Lemma B.1.1 to $\mathcal{S}$ and $\mathcal{Y}$ yields the following:

$$\mathop{\mathbb{E}}_{S\sim\mathcal{S}}[S] = \int_0^\infty (1 - F_S(x))dx,$$

and,

$$\mathop{\mathbb{E}}_{Y\sim\mathcal{Y}}[Y] = \int_0^\infty (1 - F_S(x)^k)dx.$$

By taking the difference of the two, we have

$$\mathop{\mathbb{E}}_{Y\sim\mathcal{Y}}[Y] = \mathop{\mathbb{E}}_{S\sim\mathcal{S}}[S] + \int_0^\infty F_S(x)\left(1 - F_S(x)^{k-1}\right)dx.$$

To bound the right hand side, recall that $\mathcal{S}$ is just the distribution of the sum of n' independent random variables sampled according to $\mathcal{D}_g$. Note that the third moment of $\mathcal{D}_g$ is at most 1 because it is bounded in $[0,1]$. Thus, by applying the Berry-Esseen Theorem (Lemma 2.3.3), we have

$$\left| F_S(x) - \Pr_{y\sim\mathcal{N}(\mu_g n', \sigma_g^2 n')}[y \le x] \right| \le \frac{C_{BE}}{\sigma_g^3\sqrt{n'}}.$$

for all $x \in \mathbb{R}$. When n' is sufficiently large, the right hand side is at most 0.1.

Moreover, one can check that for every $x \in \left[\mu_g n', \mu_g n' + \sigma_g\sqrt{n'}\right]$, we have $\Pr_{y\sim\mathcal{N}(\mu_g n', \sigma_g^2 n')}[y \le x] \in [0.5, 0.85]$. Hence, $F_S(x) \in [0.4, 0.95]$ for every $x \in \left[\mu_g n', \mu_g n' + \sigma_g\sqrt{n'}\right]$.

Now, we can bound $\mathbb{E}_{Y\sim\mathcal{Y}}[Y]$ as follows:

$$\begin{aligned}
\mathop{\mathbb{E}}_{Y\sim\mathcal{Y}}[Y] &= \mathop{\mathbb{E}}_{S\sim\mathcal{S}}[S] + \int_0^\infty F_S(x)\left(1 - F_S(x)^{k-1}\right) dx \\
&= \mu_g n' + \int_0^\infty F_S(x)\left(1 - F_S(x)^{k-1}\right) dx \\
\text{(Since } F_S(x)\left(1 - F_S(x)^{k-1}\right) \geq 0) &\geq \mu_g n' + \int_{\mu_g n'}^{\mu_g n' + \sigma_g\sqrt{n'}} F_S(x)\left(1 - F_S(x)^{k-1}\right) dx \\
&\geq \mu_g n' + \int_{\mu_g n'}^{\mu_g n' + \sigma_g\sqrt{n'}} (0.4)(0.05) dx \\
&= \mu_g n' + \sigma_g\sqrt{n'}/50 \\
\text{(Since } \sigma_g \geq \sigma_{min}) &\geq \mu_g n' + \sigma_{min}\sqrt{n'}/50.
\end{aligned}$$

Plugging the above inequality into equation (B.2), we can conclude that

$$\mathbb{E}\left[B_{ij,g}\right] = \frac{1}{n'k}\mathop{\mathbb{E}}_{Y\in\mathcal{Y}}[Y] \geq \frac{\mu_g}{k} + \frac{\sigma_{min}}{50k\sqrt{n'}}.$$

From this and equation (B.1), we have

$$\mathbb{E}\left[C_{ij,g}^{i'}\right] = \frac{1}{k-1}\left(\mu_g - \mathbb{E}\left[B_{ij,g}\right]\right) \leq \frac{1}{k-1}\left(\mu_g - \frac{\mu_g}{k}\right) = \frac{\mu_g}{k}.$$

Now, define $Z_{ij,g}^{i'}$ as $Z_{ij,g}^{i'} = C_{ij,g}^{i'} + \left(\mu_g/k - \mathbb{E}\left[C_{ij,g}^{i'}\right]\right)$. Notice $\mathbb{E}\left[Z_{ij,g}^{i'}\right] = \mu_g/k$.

As stated earlier, $u_{ij}(G_{i'}) > u_{ij}(G_i)$ if and only if $\sum_{g\in G} B_{ij,g} < \sum_{j\in M} C_{ij,g}^{i'}$. Let $S_B = \sum_{g\in G} B_{ij,g}, S_C = \sum_{g\in G} C_{ij,g}^{i'}, S_Z = \sum_{g\in G} Z_{ij,g}^{i'}$ and let $\delta = \frac{\sigma_{min}}{200\mu_g\sqrt{n'}}$. Notice that, since we assume that the variance of $\mathcal{D}_g$ is positive, μ_g is also non-zero, which means that δ is well-defined. Using the Chernoff bound (Lemma 2.3.1) on S_B and S_Z, we have

$$\Pr[S_B \leq (1-\delta)\,\mathbb{E}[S_B]] \leq \exp\left(\frac{-\delta^2\,\mathbb{E}[S_B]}{2}\right),$$

and,

$$\Pr[S_Z \geq (1+\delta)\,\mathbb{E}[S_Z]] \leq \exp\left(\frac{-\delta^2\,\mathbb{E}[S_Z]}{3}\right).$$

Moreover, when n' is large enough, we have $(1-\delta)\,\mathbb{E}[S_B] \geq (1+\delta)\,\mathbb{E}[S_Z]$. Thus, we have

$$\begin{aligned}
\Pr[S_B < S_Z] &\leq \exp\left(\frac{-\delta^2\,\mathbb{E}[S_B]}{2}\right) + \exp\left(\frac{-\delta^2\,\mathbb{E}[S_Z]}{3}\right) \\
&\leq \exp\left(\frac{-\delta^2 m\mu_g}{2k}\right) + \exp\left(\frac{-\delta^2 m\mu_g}{3k}\right) \\
&\leq 2\exp\left(\frac{-\sigma_{min}^2 m}{120000kn'\mu_g}\right) \\
\text{(Since } \mu_g \leq 1) &\leq 2\exp\left(\frac{-\sigma_{min}^2 m}{120000n}\right).
\end{aligned}$$

Due to how $Z_{ij,g}^{i'}$ is defined, we have $\Pr[S_B < S_Z] \geq \Pr[S_B < S_C] = \Pr[u_{ij}(G_{i'}) > u_{ij}(G_i)]$. Using the union bound for all agents a_{ij} and all groups $A_{i'} \neq A_i$, the probability that the allocation output by the algorithm is not envy-free is at most

$$2n(k-1)\exp\left(\frac{-\sigma_{min}^2 m}{120000n}\right),$$

which is at most $1/m$ when $m \geq Cn\log n$ for some sufficiently large C. This completes the proof of the theorem.

Appendix C

Examples for Chapter 7

C.1 Examples for Remark 7.5.1

Brandt and Harrenstein [41, p. 1729] mention that $\widehat{\alpha}$ and $\widehat{\gamma}$ are independent from each other in the context of general choice functions. Here, we prove that the same holds even in the context of tournament solutions.

Proposition C.1.1. *There exists a tournament solution that satisfies $\widehat{\alpha}$, but not $\widehat{\gamma}$.*

Proof. Let S be a stable tournament solution. As mentioned in Section 7.3, $\lceil S \rceil$ satisfies $\widehat{\alpha}$. However, it is easily seen that, unless S is trivial, $\lceil S \rceil$ violates $\widehat{\gamma}$. Hence, the statement follows from the existence of non-trivial stable tournament solutions (such as BP). ☐

Proposition C.1.2. *There exists a tournament solution that satisfies $\widehat{\gamma}$, but not $\widehat{\alpha}$.*

Proof. Let S be a stable tournament solution. Define the tournament solution S' such that for each tournament $T = (A, \succ)$,

$$S'(T) = \begin{cases} S(T) & \text{if } |A \setminus S(T)| > 1 \\ A & \text{otherwise.} \end{cases}$$

It can be shown that S' satisfies $\widehat{\gamma}$, but may violate $\widehat{\alpha}$. For the latter, let $S = TC$ and consider a transitive tournament $(\{a, b, c\}, \succ)$ such that $a \succ b$, $b \succ c$, and $a \succ c$. By definition, $S'(\{a, b, c\}) = \{a\}$, but $S'(\{a, b\}) = \{a, b\}$. ☐

C.2 Examples for Remark 7.5.3

BA satisfies local $\widehat{\alpha}$, but $\widehat{BA} = ME$ violates $\widehat{\alpha}$ [42].

Similarly, there exists a tournament solution S for which $\widehat{S}$ is well-defined, but $\widehat{S}$ is not stable. For a stable tournament solution S, we have by definition that $S = \widehat{S}$ and hence that $\widehat{S}$ is also stable. The following proposition shows that $\widehat{\alpha}$ does not carry over from S to $\widehat{S}$ even if S is simple and $\widehat{S}$ is well-defined.

Proposition C.2.1. *There exists a simple tournament solution S satisfying $\widehat{\alpha}$ such that $\widehat{S}$ is well-defined but $\widehat{S}$ does not satisfy $\widehat{\alpha}$.*

Proof. Let S be the tournament solution that always chooses all alternatives, with two exceptions:

- If the tournament is of order 2, then S chooses only the Condorcet winner.
- If the tournament is the tournament T_4 given in Figure C.1, then S chooses alternatives a, b, and c.

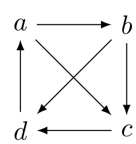

Figure C.1: Tournament T_4

Clearly, S is simple and satisfies $\widehat{\alpha}$. Since $\widehat{S}$ chooses alternatives a, b, and c from T_4, but chooses only the Condorcet winner from the transitive tournament of order 3, it does not satisfy $\widehat{\alpha}$.

It remains to show that $\widehat{S}$ is well-defined. One can check that every tournament contains an S-stable set. Suppose for contradiction that some tournament T contains two distinct minimal S-stable sets, which we denote by B and C. Then B and C are also S-stable in $B \cup C$. If B is a singleton, then B is the Condorcet winner in $B \cup C$, which means C cannot be S-stable, a contradiction. Hence both B and C are transitive tournaments of order 3, and $4 \leq |B \cup C| \leq 6$. One can check all the possibilities of $B \cup C$ to conclude that this case is also impossible. □

For the tournament solution S defined in the proof of Proposition C.2.1, we have that $\widehat{\widehat{S}}$ is not well-defined. Even though no tournament contains two distinct minimal $\widehat{S}$-stable sets, T_4 does not contain any $\widehat{S}$-stable set. This example also shows that for a tournament solution S', $\widehat{S'}$ may fail to be well-defined not because it allows two distinct minimal S'-stable sets in a tournament but because some tournament contains no S'-stable set.

C.3 Examples for Remark 7.6.13

We show that there exists a tournament solution different from BP that satisfies LRS, monotonicity, regularity, and Condorcet consistency.

To this end, we define a new tournament solution called *POS* which chooses all alternatives with positive relative degree. More precisely, an alternative is chosen by *POS* if it dominates strictly more than half of the remaining alternatives, is not chosen if it dominates strictly less than half of the remaining alternatives, and goes to a "tie-break" to determine whether it is chosen if it dominates exactly half of the remaining alternatives.

For tournaments of even size, *POS* chooses exactly the alternatives that dominate at least (or equivalently, more than) half of the remaining alternatives. Hence we do not need a tie-break for tournaments of even size. The tie-breaking rule for tournaments of odd size $2n+1$ is as follows: For any (unlabeled) tournament T of order $2n$ and any partition of it into two sets B and C of size n, consider two tournaments T_1 and T_2 of order $2n+1$. The tournament T_1 contains T and another alternative a that dominates B but is dominated by C, while the tournament T_2 contains T and another alternative a that dominates C but is dominated by B. If T_1 or T_2 is regular, *POS* chooses a in that tournament and not in the other one. Otherwise, *POS* arbitrarily chooses a in exactly one of T_1 and T_2.

Proposition C.3.1. *POS satisfies LRS, monotonicity, and regularity.*

Proof. We need to show that the tie-breaking rule in the definition of *POS* is well-defined. First, we show that if we perform a local reversal on alternative a, we do not get an isomorphic tournament with alternative a mapped to itself. Indeed, if a were mapped to itself, it would mean that no tournament solution satisfies *LRS*, which we know is not true since *BP* satisfies *LRS*. Secondly, the tournament obtained by performing a local reversal on an alternative in a regular tournament is not regular. Hence we do not obtain a conflict within the tie-breaking rule.

It follows directly from the definition that *POS* satisfies *LRS*, monotonicity, and regularity. □

The tournament T_4 given in Figure C.1 shows that *POS* violates composition-consistency and $\widehat{\alpha}$ (and hence stability).

Interestingly, *BP* (and all of its coarsenings) always intersect with *POS* while there exists a tournament for which *BA* (and all of its refinements such as *TEQ* and *ME*) do *not* overlap with *POS*. This follows from results on the *Copeland value* by Laffond et al. [93, 95].

C.4 Examples for Remark 7.6.14

We construct a tournament solution that satisfies monotonicity and stability, but violates regularity and composition-consistency.

Every tournament solution has to be regular on tournaments of order 5 or less because of nontrivial automorphisms. Consider the tournament T_7 shown in Figure C.2. T_7 admits a unique nontrivial automorphism that maps each of the six alternatives in the two 3-cycles to the next

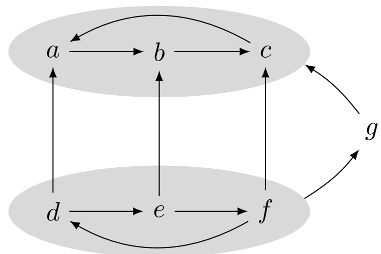

Figure C.2: Tournament T_7. $g \succ \{a, b, c\}$, $\{d, e, f\} \succ g$, and all omitted edges point downwards.

alternative in its 3-cycle and maps alternative g to itself.[1]

Now, define the simple tournament solution S_7, which always returns all alternatives unless the tournament is T_7 or it can be modified from T_7 by weakening alternative g. In the latter case, S_7 returns all alternatives except g.

We check that this definition is sound. First, we know that in T_7, there is no automorphism that maps alternative g to another alternative. When we weaken g, it is the unique alternative with the smallest out-degree, and hence cannot be mapped by an automorphism to another alternative. Now, the alternatives a, b, c form an orbit, and S_7 excludes g whenever it is dominated by d, e, f (and has any dominance relationship to a, b, c). This yields four non-isomorphic tournaments for which S_7 excludes g.

Proposition C.4.1. *$\widehat{S_7}$ satisfies stability and monotonicity.*

Proof. First, observe that S_7 trivially satisfies local $\widehat{\alpha}$ because S_7 only excludes an alternative in tournaments of order 7. By virtue of Theorem 7.3.4, it therefore suffices to show that $\widehat{S_7}$ is well-defined.

One can check that every tournament contains an S_7-stable set. Let T_6 denote the tournament obtained by removing alternative g from T_7. Suppose for contradiction that there exists a tournament T that contains two distinct minimal S_7-stable sets, which we denote by B and C. Then B and C are also S_7-stable in $B \cup C$. Moreover, $T|_B$ must correspond to the tournament T_6, and each alternative in $C \backslash B$ either has the same dominance relation to B as the alternative g does to T_6 or has a dominance relation that is a weakening of g. The same statement holds for C. We consider the following cases.

Case 1: $10 \leq |B \cup C| \leq 11$. The tournament $T|_B$ has one of its alternatives corresponding to alternatives d, e, and f in Figure C.2 outside of $B \cap C$, and this alternative must dominate all of the alternatives in C. Similarly, there exists an alternative in $C \backslash B$ that dominates all of the alternatives in B. But this implies that some two alternatives dominate each other, a contradiction.

Case 2: $7 \leq |B \cup C| \leq 9$. At least one of the two tournaments $T|_B$ and $T|_C$ must have all of its alternatives corresponding to alternatives d, e, and f in Figure C.2 in the intersection $B \cap C$,

[1]Note that $(T_7)^g$ is the smallest tournament in which BA and UC differ [39].

for otherwise we obtain a contradiction in the same way as in Case 1. Assume without loss of generality that $T|_B$ has its alternatives corresponding to d, e, and f in the intersection. Hence three alternatives in $B \cap C$ that form a cycle dominate the same alternative in C. But this does not occur in T_6, a contradiction.

It follows from Theorem 7.5.2 that $\widehat{S_7}$ satisfies monotonicity. □

Clearly, $\widehat{S_7}$ is not regular since it excludes an alternative from the regular tournament T_7. Moreover, it is not a coarsening of BP since BP selects all of the alternatives in T_7. Hence we have that stable and monotonic tournament solutions are not necessarily coarsenings of BP.

Bibliography

[1] Mark Aizerman and Fuad Aleskerov. *Theory of Choice*, volume 38 of *Studies in Mathematical and Managerial Economics*. North-Holland, 1995.

[2] Stefano Allesina and Jonathan M. Levine. A competitive network theory of species diversity. *Proceedings of the National Academy of Sciences*, 108(14):5638–5642, 2011.

[3] Noga Alon. Splitting necklaces. *Advances in Mathematics*, 63(3):247–253, 1987.

[4] Noga Alon, Dana Moshkovitz, and Shmuel Safra. Algorithmic construction of sets for k-restrictions. *ACM Transactions on Algorithms*, 2(2):153–177, 2006.

[5] Georgios Amanatidis, Georgios Birmpas, George Christodoulou, and Evangelos Markakis. Truthful allocation mechanisms without payments: Characterization and implications on fairness. In *Proceedings of the 18th ACM Conference on Economics and Computation*, pages 545–562, 2017.

[6] Georgios Amanatidis, Georgios Birmpas, and Evangelos Markakis. On truthful mechanisms for maximin share allocations. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 31–37, 2016.

[7] Georgios Amanatidis, Evangelos Markakis, Afshin Nikzad, and Amin Saberi. Approximation algorithms for computing maximin share allocations. *ACM Transactions on Algorithms*, 13(4):52, 2017.

[8] Ian Anderson. *Combinatorial Design and Tournaments*. Oxford University Press, 1997.

[9] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

[10] Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 1. North-Holland, 2002.

[11] Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 2. North-Holland, 2010.

[12] Haris Aziz, Serge Gaspers, Simon Mackenzie, Nicholas Mattei, Paul Stursberg, and Toby Walsh. Fixing a balanced knockout tournament. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pages 552–558, 2014.

[13] Haris Aziz, Serge Gaspers, Simon Mackenzie, and Toby Walsh. Fair assignment of indivisible objects under ordinal preferences. *Artificial Intelligence*, 227:71–92, 2015.

[14] Imre Bárány. A short proof of Kneser's conjecture. *Journal of Combinatorial Theory, Series A*, 25(3):325–326, 1978.

[15] Salvador Barberà, Walter Bossert, and Prasanta K. Pattanaik. Ranking sets of objects. In Salvador Barberà, Peter J. Hammond, and Christian Seidl, editors, *Handbook of Utility Theory, Volume 2*, pages 893–977. Kluwer Academic Publishers, 2004.

[16] Siddharth Barman and Sanath Kumar Krishna Murthy. Approximation algorithms for maximin fair division. In *Proceedings of the 18th ACM Conference on Economics and Computation*, pages 647–664, 2017.

[17] Roberto Barrera, Kathryn Nyman, Amanda Ruiz, Francis Edward Su, and Yan Zhang. Discrete envy-free division of necklaces and maps. *CoRR*, abs/1510.02132, 2015.

[18] Xiaohui Bei, Guangda Huzhang, and Warut Suksompong. Truthful fair division without free disposal. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 63–69, 2018.

[19] Andrew C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49(1):122–136, 1941.

[20] Kshipra Bhawalkar and Tim Roughgarden. Welfare guarantees for combinatorial auctions with item bidding. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 700–709, 2011.

[21] D. C. Blest and D. G. Fitzgerald. Scheduling sports competitions with a given distribution of times. *Discrete Applied Mathematics*, 22(1):9–19, 1988.

[22] Géza Bohus. On the discrepancy of 3 permutations. *Random Structures and Algorithms*, 1(2):215–220, 1990.

[23] Georges Bordes. Some more results on consistency, rationality and collective choice. In Jean-Jacques Laffont, editor, *Aggregation and Revelation of Preferences*, chapter 10, pages 175–197. North-Holland, 1979.

[24] Sylvain Bouveret, Katarína Cechlárová, Edith Elkind, Ayumi Igarashi, and Dominik Peters. Fair division of a graph. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 135–141, 2017.

[25] Sylvain Bouveret, Ulle Endriss, and Jérôme Lang. Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In *Proceedings of the 19th European Conference on Artificial Intelligence*, pages 387–392, 2010.

[26] Sylvain Bouveret and Michel Lemaître. Characterizing conflicts in fair division of indivisible goods using a scale of criteria. *Autonomous Agents and Multi-Agent Systems*, 30(2):259–290, 2016.

[27] Steven J. Brams, D. Marc Kilgour, and Christian Klamler. The undercut procedure: an algorithm for the envy-free division of indivisible items. *Social Choice and Welfare*, 39(2–3):615–631, 2012.

[28] Steven J. Brams, D. Marc Kilgour, and Christian Klamler. Two-person fair division of indivisible items: an efficient, envy-free algorithm. *Notices of the AMS*, 61(2):130–141, 2014.

[29] Florian Brandl, Felix Brandt, and Hans Georg Seedig. Consistent probabilistic social choice. *Econometrica*, 84(5):1839–1880, 2016.

[30] Florian Brandl, Felix Brandt, and Warut Suksompong. The impossibility of extending random dictatorship to weak preferences. *Economics Letters*, 141:44–47, 2016.

[31] Felix Brandt. Minimal stable sets in tournaments. *Journal of Economic Theory*, 146(1):1481–1499, 2011.

[32] Felix Brandt. Set-monotonicity implies Kelly-strategyproofness. *Social Choice and Welfare*, 45(4):793–804, 2015.

[33] Felix Brandt, Markus Brill, Felix Fischer, and Paul Harrenstein. Minimal retentive sets in tournaments. *Social Choice and Welfare*, 42(3):551–574, 2014.

[34] Felix Brandt, Markus Brill, and Paul Harrenstein. Tournament solutions. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 3, pages 57–84. Cambridge University Press, 2016.

[35] Felix Brandt, Markus Brill, Hans Georg Seedig, and Warut Suksompong. On the structure of stable tournament solutions. *Economic Theory*, 65(2):483–507, 2018.

[36] Felix Brandt, Markus Brill, and Warut Suksompong. An ordinal minimax theorem. *Games and Economic Behavior*, 95:107–112, 2016.

[37] Felix Brandt, Maria Chudnovsky, Ilhee Kim, Gaku Liu, Sergey Norin, Alex Scott, Paul Seymour, and Stephan Thomassé. A counterexample to a conjecture of Schwartz. *Social Choice and Welfare*, 40(3):739–743, 2013.

[38] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[39] Felix Brandt, Andre Dau, and Hans Georg Seedig. Bounds on the disparity and separation of tournament solutions. *Discrete Applied Mathematics*, 187:41–49, 2015.

[40] Felix Brandt, Felix Fischer, Paul Harrenstein, and Maximilian Mair. A computational analysis of the tournament equilibrium set. *Social Choice and Welfare*, 34(4):597–609, 2010.

[41] Felix Brandt and Paul Harrenstein. Set-rationalizable choice and self-stability. *Journal of Economic Theory*, 146(4):1721–1731, 2011.

[42] Felix Brandt, Paul Harrenstein, and Hans Georg Seedig. Minimal extending sets in tournaments. *Mathematical Social Sciences*, 87:55–63, 2017.

[43] Felix Brandt and Hans Georg Seedig. On the discriminative power of tournament solutions. In *Selected Papers of the International Conference on Operations Research, OR2014*, Operations Research Proceedings, pages 53–58. Springer, 2016.

[44] Mark Braverman and Elchanan Mossel. Noisy sorting without resampling. In *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 268–276, 2008.

[45] Dirk Briskorn. Combinatorial properties of strength groups in round robin tournaments. *European Journal of Operational Research*, 192(3):744–754, 2009.

[46] Dirk Briskorn and Sigrid Knust. Constructing fair sports league schedules with regard to strength groups. *Discrete Applied Mathematics*, 158(2):123–135, 2010.

[47] Eric Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011.

[48] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D. Procaccia, Nisarg Shah, and Junxing Wang. The unreasonable fairness of maximum Nash welfare. In *Proceedings of the 17th ACM Conference on Economics and Computation*, pages 305–322, 2016.

[49] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, 23(4):493–507, 1952.

[50] Herman Chernoff. Rational selection of decision functions. *Econometrica*, 22(4):422–443, 1954.

[51] Marie Jean Antoine Nicolas de Caritat, Marquis de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* Imprimerie Royale, 1785.

[52] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms.* MIT Press, 3rd edition, 2009.

[53] Sebastian A. Csar, Rik Sengupta, and Warut Suksompong. On a subposet of the Tamari lattice. *Order*, 31(3):337–363, 2014.

[54] John P. Dickerson, Jonathan Goldman, Jeremy Karp, Ariel D. Procaccia, and Tuomas Sandholm. The computational rise and fall of fairness. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, pages 1405–1411, 2014.

[55] Irit Dinur and David Steurer. Analytical approach to parallel repetition. In *Proceedings of the 46th Symposium on Theory of Computing*, pages 624–633, 2014.

[56] Conal Duddy, Nicolas Houy, Jérôme Lang, Ashley Piggins, and William S. Zwicker. Social dichotomy functions. Working paper, 2014.

[57] Bhaskar Dutta. Covering sets and a new Condorcet choice correspondence. *Journal of Economic Theory*, 44(1):63–80, 1988.

[58] Paul Erdős. On a combinatorial problem. II. *Acta Mathematica Academiae Scientiarum Hungarica*, 15(3–4):445–447, 1964.

[59] Carl-Gustav Esseen. On the Liapunoff limit of error in the theory of probability. *Arkiv för matematik, astronomi och fysik*, A28:1–19, 1942.

[60] Tomer Ezra, Michal Feldman, Tim Roughgarden, and Warut Suksompong. Pricing multi-unit markets. *CoRR*, abs/1705.06623, 2018.

[61] Uriel Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[62] Uriel Feige. On maximizing welfare when utility functions are subadditive. *SIAM Journal on Computing*, 39(1):122–142, 2009.

[63] Uriel Feige, Vahab S. Mirrokni, and Jan Vondrák. Maximizing non-monotone submodular functions. *SIAM Journal on Computing*, 40(4):1133–1153, 2011.

[64] Dan S. Felsenthal and Moshé Machover. After two centuries should Condorcet's voting procedure be implemented? *Behavioral Science*, 37(4):250–274, 1992.

[65] Mark Fey. Choosing from a large tournament. *Social Choice and Welfare*, 31(2):301–309, 2008.

[66] Peter C. Fishburn. Probabilistic social choice based on simple voting comparisons. *Review of Economic Studies*, 51(4):683–692, 1984.

[67] David C. Fisher and Richard B. Reeves. Optimal strategies for random tournament games. *Linear Algebra and its Applications*, 217:83–85, 1995.

[68] David C. Fisher and Jennifer Ryan. Tournament games and positive tournaments. *Journal of Graph Theory*, 19(2):217–236, 1995.

[69] Duncan K. Foley. Resource allocation and the public sector. *Yale Economics Essays*, 7(1):45–98, 1967.

[70] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[71] Mohammad Ghodsi, MohammadTaghi HajiAghayi, Masoud Seddighin, Saeed Seddighin, and Hadi Yami. Fair allocation of indivisible goods: Improvement and generalization. In *Proceedings of the 19th ACM Conference on Economics and Computation*, pages 539–556, 2018.

[72] Jonathan Goldman and Ariel D. Procaccia. Spliddit: Unleashing fair division algorithms. *ACM SIGecom Exchanges*, 13(2):41–46, 2014.

[73] Laurent Gourvès and Jérôme Monnot. Approximate maximin share allocations in matroids. In *Proceedings of the 10th International Conference on Algorithms and Complexity*, pages 310–321, 2017.

[74] Joshua E. Greene. A new short proof of Kneser's conjecture. *American Mathematical Monthly*, 109(10):918–920, 2002.

[75] Jenifer Haselgrove and John Leech. A tournament design problem. *American Mathematical Monthly*, 84(3):198–201, 1977.

[76] Noam Hazon, Paul E. Dunne, Sarit Kraus, and Michael Wooldridge. How to rig elections and competitions. In *Proceedings of the 2nd International Workshop on Computational Social Choice*, pages 301–312, 2008.

[77] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[78] Nicolas Houy. A few new results on TEQ. Mimeo, 2009.

[79] Nicolas Houy. Still more on the tournament equilibrium set. *Social Choice and Welfare*, 32(1):93–99, 2009.

[80] Olivier Hudry. A survey on the complexity of tournament solutions. *Mathematical Social Sciences*, 57(3):292–303, 2009.

[81] David S. Johnson. Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9(3):256–278, 1974.

[82] Richard M. Karp. Reducibility among combinatorial problems. In *Proceedings of Complexity of Computer Computations*, pages 85–103, 1972.

[83] Ian A. Kash, Peter Key, and Warut Suksompong. Simple pricing schemes for the cloud. In *Proceedings of the 13th Conference on Web and Internet Economics*, pages 311–324, 2017.

[84] Michael P. Kim, Warut Suksompong, and Virginia Vassilevska Williams. Who can win a single-elimination tournament? *SIAM Journal on Discrete Mathematics*, 31(3):1751–1764, 2017.

[85] Michael P. Kim and Virginia Vassilevska Williams. Fixing tournaments for kings, chokers, and more. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 561–567, 2015.

[86] Martin Kneser. Aufgabe 360. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 58(2):27, 1955.

[87] Sigrid Knust. Scheduling sports tournaments on a single court minimizing waiting times. *Operations Research Letters*, 36(4):471–476, 2008.

[88] Sigrid Knust. Scheduling non-professional table-tennis leagues. *European Journal of Operations Research*, 200(2):358–367, 2010.

[89] David A. Kohler and R. Chandrasekaran. A class of sequential games. *Operations Research*, 19(2):270–277, 1971.

[90] Stavros G. Kolliopoulos and Neal E. Young. Approximation algorithms for covering/packing integer programs. *Journal of Computer and System Sciences*, 71(4):495–505, 2005.

[91] Germain Kreweras. Aggregation of preference orderings. In *Mathematics and Social Sciences I: Proceedings of the seminars of Menthon-Saint-Bernard, France (1–27 July 1960) and of Gösing, Austria (3–27 July 1962)*, pages 73–79, 1965.

[92] David Kurokawa, Ariel D. Procaccia, and Junxing Wang. When can the maximin share guarantee be guaranteed? In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 523–529, 2016.

[93] Gilbert Laffond, Jean-François Laslier, and Michel Le Breton. The bipartisan set of a tournament game. *Games and Economic Behavior*, 5(1):182–201, 1993.

[94] Gilbert Laffond, Jean-François Laslier, and Michel Le Breton. More on the tournament equilibrium set. *Mathématiques et sciences humaines*, 31(123):37–44, 1993.

[95] Gilbert Laffond, Jean-François Laslier, and Michel Le Breton. The Copeland measure of Condorcet choice functions. *Discrete Applied Mathematics*, 55(3):273–279, 1994.

[96] Hyman G. Landau. On dominance relations and the structure of animal societies: I. Effect of inherent characteristics. *Bulletin of Mathematical Biophysics*, 13(1):1–19, 1951.

[97] Jérôme Lang, Maria Silvia Pini, Francesca Rossi, Kristen Brent Venable, and Toby Walsh. Winner determination in sequential majority voting. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pages 1372–1377, 2007.

[98] Jean-François Laslier. *Tournament Solutions and Majority Voting*. Springer, 1997.

[99] Charles E. Leiserson, Tao B. Schardl, and Warut Suksompong. Upper bounds on number of steals in rooted trees. *Theory of Computing Systems*, 58(2):223–240, 2016.

[100] Paul Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villers, 1937.

[101] Richard J. Lipton, Evangelos Markakis, Elchanan Mossel, and Amin Saberi. On approximately fair allocations of indivisible goods. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, pages 125–131, 2004.

[102] László Lovász. On the ratio of optimal integral and fractional covers. *Discrete Mathematics*, 13(4):383–390, 1975.

[103] László Lovász. Kneser's conjecture, chromatic number and homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[104] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, 1994.

[105] Pasin Manurangsi and Warut Suksompong. Asymptotic existence of fair divisions for groups. *Mathematical Social Sciences*, 89:100–108, 2017.

[106] Pasin Manurangsi and Warut Suksompong. Computing an approximately optimal agreeable set of items. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 338–344, 2017.

[107] Pasin Manurangsi and Warut Suksompong. Computing a small agreeable set of indivisible items. Working paper, 2018.

[108] Pasin Manurangsi and Warut Suksompong. When do envy-free allocations exist? Working paper, 2018.

[109] Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Y. Ozbay. Revealed attention. *American Economic Review*, 102(5):2183–2205, 2012.

[110] Jiří Matoušek. A combinatorial proof of Kneser's conjecture. *Combinatorica*, 24(1):163–170, 2004.

[111] Nicholas Mattei, Judy Goldsmith, Andrew Klapper, and Martin Mundhenk. On the complexity of bribery and manipulation in tournaments with uncertain information. *Journal of Applied Logic*, 13(4):549–554, 2015.

[112] David C. McGarvey. A theorem on the construction of voting paradoxes. *Econometrica*, 21(4):608–610, 1953.

[113] Matthias Mnich, Yash Raj Shrestha, and Yongjie Yang. When does Schwartz conjecture hold? In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 603–609, 2015.

[114] Bernard Monjardet. Statement of precedence and a comment on IIA terminology. *Games and Economic Behavior*, 62:736–738, 2008.

[115] Scott Moser. Majority rule and tournament solutions. In Jac C. Heckelman and Nicholas R. Miller, editors, *Handbook of Social Choice and Voting*, chapter 6, pages 83–101. Edward Elgar, 2015.

[116] Dana Moshkovitz. The projection games conjecture and the NP-hardness of $\ln n$-approximating set-cover. *Theory of Computing*, 11:221–235, 2015.

[117] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, 171(1):295–341, 2010.

[118] Alantha Newman, Ofer Neiman, and Aleksandar Nikolov. Beck's three permutations conjecture: a counterexample and some consequences. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 253–262, 2012.

[119] Hoon Oh, Ariel D. Procaccia, and Warut Suksompong. Fairly allocating many goods with few queries. *CoRR*, abs/1807.11367, 2018.

[120] Ariel D. Procaccia and Junxing Wang. Fair enough: Guaranteeing approximate maximin shares. In *Proceedings of the 15th ACM Conference on Economics and Computation*, pages 675–692, 2014.

[121] Kirk Pruhs and Gerhard J. Woeginger. Divorcing made easy. In *Proceedings of the 6th International Conference on Fun with Algorithms*, pages 305–314, 2012.

[122] Rasmus V. Rasmussen and Michael A. Trick. Round robin scheduling – a survey. *European Journal of Operations Research*, 188:617–636, 2008.

[123] Ran Raz and Shmuel Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, pages 475–484, 1997.

[124] Alvin E. Roth. *Who Gets What — and Why: The New Economics of Matchmaking and Market Design*. Mariner Books, 2016.

[125] Tyrel Russell. *A Computational Study of Problems in Sports*. PhD thesis, University of Waterloo, 2010.

[126] Donald G. Saari and Steven Barney. Consequences of reversing preferences. *Mathematical Intelligencer*, 25:17–31, 2003.

[127] Christian Saile and Warut Suksompong. Robust bounds on choosing from large tournaments. In *Proceedings of the 7th International Workshop on Computational Social Choice*, 2018.

[128] P. J. Schellenberg, G. H. J. Van Rees, and S. A. Vanstone. The existence of balanced tournament designs. *Ars Combinatoria*, 3:303–318, 1977.

[129] Thorleif Schjelderup-Ebbe. Beiträge zur Sozialpsychologie des Haushuhns. *Zeitschrift für Psychologie*, 88:225–252, 1922.

[130] Jörn Schönberger, Dirk C. Mattfeld, and Hubert Kopfer. Memetic algorithm timetabling for non-commercial sport leagues. *European Journal of Operational Research*, 153(1):102–116, 2004.

[131] Thomas Schwartz. Cyclic tournaments and cooperative majority voting: a solution. *Social Choice and Welfare*, 7(1):19–29, 1990.

[132] Alex Scott and Mark Fey. The minimal covering set in large tournaments. *Social Choice and Welfare*, 38(1):1–9, 2012.

[133] Erel Segal-Halevi and Shmuel Nitzan. Proportional cake-cutting among families. *CoRR*, abs/1510.03903, 2015.

[134] Erel Segal-Halevi and Shmuel Nitzan. Envy-free cake-cutting among families. *CoRR*, abs/1607.01517, 2016.

[135] Erel Segal-Halevi and Warut Suksompong. Democratic fair allocation of indivisible goods. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 482–488, 2018.

[136] Amartya K. Sen. Choice functions and revealed preference. *Review of Economics Studies*, 38(3):307–317, 1971.

[137] Kenneth A. Shepsle and Barry R. Weingast. Uncovered sets and sophisticated voting outcomes with implications for agenda institutions. *American Journal of Political Science*, 28(1):49–74, 1984.

[138] Forest W. Simmons and Francis Edward Su. Consensus-halving via theorems of Borsuk-Ulam and Tucker. *Mathematical Social Sciences*, 45(1):15–25, 2003.

[139] Patrick Slater. Inconsistencies in a schedule of paired comparisons. *Biometrika*, 48(3–4):303–312, 1961.

[140] Isabelle Stanton and Virginia Vassilevska Williams. Manipulating stochastically generated single-elimination tournaments for nearly all players. In *Proceedings of the 7th International Workshop of Internet and Network Economics*, pages 326–337, 2011.

[141] Isabelle Stanton and Virginia Vassilevska Williams. Rigging tournament brackets for weaker players. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 357–364, 2011.

[142] Walter Stromquist. How to cut a cake fairly. *The American Mathematical Monthly*, 87(8):640–644, 1980.

[143] Walter Stromquist. Envy-free cake divisions cannot be found by finite protocols. *Electronic Journal of Combinatorics*, 15:#R11, 2008.

[144] Francis Edward Su. Rental harmony: Sperner's lemma in fair division. *The American Mathematical Monthly*, 106(10):930–942, 1999.

[145] Warut Suksompong. Apples, oranges, and bananas. *Crux Mathematicorum*, 40(3):116–119, 2014.

[146] Warut Suksompong. Bounds on multithreaded computations by work stealing. Master's thesis, Massachusetts Institute of Technology, 2014.

[147] Warut Suksompong. Individual and group stability in neutral restrictions of hedonic games. *Mathematical Social Sciences*, 78:1–5, 2015.

[148] Warut Suksompong. Assigning a small agreeable set of indivisible items to multiple players. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 489–495, 2016.

[149] Warut Suksompong. Asymptotic existence of proportionally fair allocations. *Mathematical Social Sciences*, 81:62–65, 2016.

[150] Warut Suksompong. Scheduling asynchronous round-robin tournaments. *Operations Research Letters*, 44(1):96–100, 2016.

[151] Warut Suksompong. Fairly allocating contiguous blocks of indivisible items. In *Proceedings of the 10th International Symposium on Algorithmic Game Theory*, pages 333–344, 2017.

[152] Warut Suksompong. On black-box transformations in downward-closed environments. In *Proceedings of the 10th International Symposium on Algorithmic Game Theory*, page XV, 2017.

[153] Warut Suksompong. Approximate maximin shares for groups of agents. *Mathematical Social Sciences*, 92:40–47, 2018.

[154] Warut Suksompong, Charles E. Leiserson, and Tao B. Schardl. On the efficiency of localized work stealing. *Information Processing Letters*, 116(2):100–106, 2016.

[155] William Thomson. Introduction to the theory of fair allocation. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 11, pages 261–283. Cambridge University Press, 2016.

[156] Pim van 't Hof, Gerhard Post, and Dirk Briskorn. Constructing fair round robin tournaments with a minimum number of breaks. *Operations Research Letters*, 38(6):592–596, 2010.

[157] Hal R. Varian. Equity, envy, and efficiency. *Journal of Economic Theory*, 9:63–91, 1974.

[158] Virginia Vassilevska Williams. Fixing a tournament. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence*, pages 895–900, 2010.

[159] Thuc Vu, Alon Altman, and Yoav Shoham. On the complexity of schedule control problems for knockout tournaments. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 225–232, 2009.

[160] Gerhard J. Woeginger. A polynomial-time approximation scheme for maximizing the minimum machine completion time. *Operations Research Letters*, 20(4):149–154, 1997.

[161] Yongjie Yang. A further step towards an understanding of the tournament equilibrium set. *CoRR*, abs/1611.03991, 2016.

[162] Lishun Zeng and Shinji Mizuno. Constructing fair single round robin tournaments regarding strength groups with a minimum number of breaks. *Operations Research Letters*, 41(5):506–510, 2013.