Incentive Compatible Mechanisms without Money

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Nicos Protopapas

April 2022

Abstract

Mechanism design arises in environments where a set of strategic agents should achieve a common goal, but this goal may be affected by the selfish behavior of the agents. A popular tool to mitigate this impact is *incentive compatibility*, the design of mechanisms in such a way that strategic agents are motivated to act honestly. Many times this can be done using *payments*: monetary transactions can be implemented by the mechanism, which provide the agents with the right incentives to reveal their true colors. However, there are cases where such payments are not applicable for various reasons, moral, legal, or practical. In this thesis, we focus on problems where payments are prohibited, and we propose incentive compatible solutions, respecting this constraint.

We concentrate on two main problems: the problem of *impartial selection* and the problem of *truthful budget aggregation*. In both problems, strategic agents need to come up with a joint decision, but their selfish behavior may lead them to highly sub-optimal solutions. Our goal is to design mechanisms providing the agents with proper incentives to act sincerely. Unfortunately, we are only able to achieve this by sacrificing the quality of the solution, in the sense that the solutions we get are not as good as the solutions we could get in an environment where the agents would not be strategic. Therefore, we compare our mechanisms with ideal, non-strategic outcomes, providing worst-case approximation guarantees.

The first problem we confront, impartial selection, involves the selection of an influential member of a community of individuals. This community can be described by a directed graph, where the nodes represent the individuals and the directed edges represent nominations. The task is given this graph to select the node with the highest number of nominations. However, the community members are selfish agents; hence, their reported nominations are not trusted, and this seemingly trivial task is now challenging. Impartiality, a property requiring no single node to influence her selection probability, provides proper incentives to the agents to act honestly. Recent progress in the literature has identified impartial selection rules with optimal approximation ratios, i.e., the ratio between the maximum in-degree and the in-degree of the selected node. However, it was noted that worst-case instances are graphs with small in-degrees. Motivated by this fact, we deviate from the trend and propose the study of *additive approximation*: the difference between the highest number of nominations and the number of nominations of the selected member, as an alternative

measure of the quality of impartial selection mechanisms.

The first part of this thesis is concerned with the design of impartial selection mechanisms with small additive approximation guarantees. On the positive side, we were able to design two randomized impartial selection mechanisms with additive approximation guarantees of $\mathcal{O}(\sqrt{n})$ and $\mathcal{O}(n^{2/3}\ln^{1/3} n)$ for two well-studied models in the literature, where n denotes the community size. We complement our positive results by providing negative results for various cases. First, we provide a characterization for the interesting class of *strong sample* mechanisms, which allows us to obtain lower bounds of $n-2$, and of $\Omega(\sqrt{n})$ for their deterministic and randomized variants, respectively. Also, we present a general lower bound of 3 for all deterministic impartial mechanisms. The best-known upper bound is the trivial one, of $n-1$, which leaves us with a tantalizing open question. Finally, we provide additive approximation bounds for some known mechanisms.

We continue our investigation of the impartial selection problem from another direction. Getting our inspiration from the design of auction and posted pricing mechanisms with good approximation guarantees for welfare and profit maximization, we follow up our work with an enhanced model, where we study the extent to which prior information on voters' preferences could be helpful in the design of efficient deterministic impartial selection mechanisms with good additive approximation guarantees. First, we define a hierarchy of three models of prior information, which we call the *opinion poll*, the *a priori popularity*, and the *uniform* model. Then, we analyze the performance of a natural selection mechanism that we call *Approval Voting with Default* and show that it achieves a $\mathcal{O}(\sqrt{n\ln n})$ additive guarantee for opinion poll and a $\mathcal{O}(\ln^2 n)$ for a priori popularity inputs, where n is the number of individuals. We consider the polylogarithmic bound as the leading technical contribution of this part. Finally, we complement this last result by showing that our analysis is close to tight, showing an $\Omega(\ln n)$ lower bound.

We then turn our attention to the *truthful budget aggregation* problem. In this problem, strategic voters wish to split a divisible budget among different projects by aggregating their proposals into a single budget division. Unfortunately, it is well-known that the straightforward rule that divides the budget proportionally is susceptible to manipulation, and while sophisticated incentive compatible mechanisms have been proposed in the literature, their outcomes are often far from fair. To capture this loss of fairness imposed by the need for truthfulness, we propose a quantitative framework that evaluates a budget aggregation mechanism according to its worst-case distance from the proportional allocation. We study this measure in the recently proposed class of incentive compatible mechanisms, called the *moving phantom* mechanisms, and we provide approximation guarantees. For two projects, we show that the well-known *Uniform Phantom* mechanism is optimal among all truthful mechanisms. For three projects, we propose the proportional, *Piecewise Uniform* mechanism that is optimal among all moving phantom mechanisms. Finally, we provide impossibility results regarding the approximability of moving phantom mechanisms, and budget aggregation mechanisms, in general.

Acknowledgements

I am deeply grateful to have George Christodoulou as my supervisor. George, it was an honor and pleasure having you as my mentor during this challenging yet remarkable journey. I am incredibly thankful for your enthusiastic encouragement, patient guidance, constructive criticism, and all the lively discussions we had during my Ph.D. studies, both on an academic and personal level.

Secondly, I would like to thank my second supervisor Piotr Krysta, and my academic advisors, John Fearnley and Martin Gairing, for evaluating my progress and helping me improve as a researcher. Finally, I am also very grateful to the member of my thesis committee, John Fearnley and Felix Fischer, for their productive feedback during the viva examination and the helpful comments on the thesis itself.

I would also like to extend my deepest gratitude to Ioannis Caragiannis for introducing me to Computer Science research and for being an inspiring collaborator throughout the years. I also would like to thank my early collaborators, Xenophon Chatzigeorgiou, Panagiotis Kanellopoulos, George Krimpas, and Alexandros Voudouris. Finally, I would like to thank Vasilis Gketzelis for our inspiring discussions.

I should also thank all the people of the Economics and Computations group of the department. It was a privilege and a joy to be part of this inspiring research group, an authentic research incubator. Special thanks to Boris Konev, Paul Spirakis, Darek Kowalski, and Roberto Ferrero for allowing me to get valuable experience as a demonstrator and academic instructor. I cannot leave the University of Liverpool without thanking all the people in the department's support staff for being helpful whenever I needed them. I must also thank the University of Liverpool for its financial support, making my Ph.D. studies possible.

My life in my Ph.D. studies would have not been the same without the people I met during those years: Themistoklis, Michael, Katerina, Eleni, Elektra, Yiannis, Argyris, Giorgos, Peter, Xenophon, Matina, Xia, Matoula, Katerina, Manos, Artem, Bart, Aris, Alkmini, Demetris, Estelle, Kostas, Christos, and all those I might have forgotten.

Last but not least, I want to thank my closest friends and my family, Frixos, Mary, Eleni, and Aggelos. Without your unwavering support, I would not be able to follow this career path. Most importantly, my heartfelt thanks to Eleni, my life companion, for brightening my days.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

We start the presentation of this work with two illustrative examples, to provide the main motivation for this thesis.

Suppose first, that we have a prestigious position that should be filled by a single individual. What makes this example more intriguing is that all eligible candidates should collectively decide which one of them will fill the position. All of them have a name (or more) in their minds for the person they genuinely believe to be the right one to take this position. However, since all of them are eligible, they all prefer to be selected for the position, than anyone else. If we are allocated with the task to find the most popular candidate among them and award her the position, what is the best possible way to learn who the most popular candidate is? Indeed, if we ask them directly to name someone and then award the position to the candidate with the maximum number of nominations, we might end up in a situation where some individuals, *strategically* thinking, would not nominate an individual worthy of the position (according to their own beliefs, of course), to increase their own chance of winning. In an extreme case of selfish behavior, we may end up in a satirical condition where everyone will vote for the person they believe to have the least chance to win.

In a second example, consider a small start-up company, owned by a few college friends equally. The company has gained a considerable amount of profit in its first years. The owners have decided to invest a large share of that money to grow their company further. Some of them believe they should invest their money in commodities, others believe they should expand the company by investing in new equipment and hiring more personnel, while others believe they should make some high-risk/high-reward investments in other startup companies. Some of them would like to do all of these but on different levels. A crucial question here is how they should take an *joint* decision about these investments. A simple procedure is to ask each of them to cast a proposal on how they believe this budget should be split over the three possible options, and the joint decision will be the average proposal. However, as we will see later on, this procedure is easily manipulable. Any one of the partners may be tempted to exaggerate their proposals in order to guide the joint decision towards a *personally* more appealing outcome.

The above two examples summarize the motivation behind this thesis. In both cases, a set of individuals need to come up with a *joint decision* based on private, subjective preferences. An ideal decision process would elicit these preferences from the individuals and would use them to compute a good joint decision. This decision, however, affects the individuals directly, and as such, they might be tempted to manipulate the decision process to their benefit. The main goal of this thesis is to propose procedures, known to the literature as *mechanisms* which will provide proper incentives to the individuals to reveal their true preferences, and ensure that the joint decision fulfills some quality standards.

In many cases, we can design efficient mechanisms with using *payments*, which work as compensation or penalties for the participants. Maybe the most well-known example is the idea of the second-prize auction for the allocation of a single item, on which we will elaborate later on. However, in the scenarios described above, the use of monetary transactions is not possible, due to ethical or practical reasons. In the first example, our goal is to fill the position based on popularity and merit. In the second example, the partners may want to avoid using monetary transactions to keep their decision power equal.

If our goal is simply to design some mechanisms which incentivize the individuals to act truthfully, then this thesis can be very short. Indeed, we can trivially devise a mechanism where no agent can change the outcome at all, by just ignoring their proposals. While one such mechanism would be enough to guarantee incentive compatibility, it is a useless mechanism in most cases. For this reason, we propose and use some measures to quantify how good each mechanism is, according to some efficiency or fairness criteria.

As we will later see, incentive compatibility comes at a price. In all of our problems, we cannot impose incentive compatibility without losing something on the quality of the solution. To quantify this loss we compare our solutions with the solution of an ideal algorithm, which already knows the preferences of the individuals.

The main goals of this thesis are three. First, to provide measures, capturing notions of fairness and efficiency in environments with selfish behaviors. Second, to provide and analyze mechanisms that achieve these goals as best as possible. Finally, to explore the limits of the proposed mechanisms and measures with impossibility results.

Roadmap

In the following section, we give a brief introduction to Mechanism Design, focused mainly on incentive compatibility. In Section 1.2 we present the problem of impartial selection, which is connected to our first example and it is the prime focus of Chapters 3 and 4. In Section 5 we introduce the problem of truthful budget aggregation, presented through the emerging example of participatory budgeting, which is connected with our second example and is the focal point of Chapter 5. In Section 1.4 we summarize our results. Finally, Section 1.5 presents a roadmap for the rest of the thesis.

1.1 Mechanism Design

Mechanism design is a sub-field of Game Theory and Microeconomics which came into close interaction with Computer Science from the late 1990s and forth, mainly due to the Internet and its vast growth.

From the computer scientist's point of view, a mechanism is an algorithm, whose input is provided by *self-interested* parties. Under this glance, the algorithm designer, need only not consider the usual desiderata of algorithm design (e.g. correctness, complexity e.t.c) but also needs to consider how the *incentives* of the parties providing the input will influence the outcome.

To show how important is to care for incentives when designing algorithms with self-interested parties, we present an entertaining example of a failed tournament design. A famous incident in Association Football, happened in a 1994 Caribbean Cup qualifications match, between the national teams of Barbados and Grenada [59]. For a brief time, both teams' best interest was to score an own goal, to qualify for the finals. At the time, the organizers had decided to not allow any match to end in a tie. Whenever a game was tied after the usual 90 minutes of play, would continue for 60 more minutes of *extra time*. In the extra time, there was an unusual golden goal rule, where a goal, if scored ends the game, and is worth two goals. These rules can be seen as part of an algorithm, which gets as input the state of the game and decides the score. Note that this input comes from two self-interest parties. The algorithm's designer goal is to enhance competition. The two teams, on the other side, want to advance to the next stage of the competition.

Before the start of the game, due to previous results, there was a single way for Barbados' team to advance to the finals: to win by at least two goals. Grenada would advance otherwise. Barbados had a 2-0 advantage until the 83rd minute of the game when Grenadas' team scored and become again the front-runner. After a few minutes where the score did not change, two Barbados' players decided that they have a better chance to get their two-goal difference in the 30 minutes of extra time, rather than in the following few minutes they had until the game finish. Hence, at the 87-th minute, they *intentionally* scored an own goal, making the score 2-2. In the remaining minutes, Grenada's players realized that they have to score, on any side of the pitch, to either lose or win by one goal and qualify for the finals. For a few minutes, Grenada's players were attacking both goals, while Barbados' players were defending both goals, making a mockery of the basic rules of the game. No more goals were scored in this chaotic situation, and the 90 minutes of play ended as a 2-2 draw. Eventually, Barbados' team gamble worked, as they scored the golden goal in extra time and won by 4-2, advancing to the finals.

It is no surprise that both rules, the one preventing ties and the double golden goal rule are not in use today. This is a situation where the rule designers failed to align the participants' incentives with their goal, to have a competitive tournament where all teams will try to score more than their opponents. For another interesting situation of a failed tournament design from sports, see Chapter 1 of [91].

1.1.1 Incentive Compatibility

From the economics perspective, the foundations of Mechanism Design were laid down in the 1960s with the work of Leonid Hurwicz, who asked the simple question: How should a social planner come up with a decision involving information which is spread among a number of people? [63]. Hurwicz, which was awarded the 2007 Nobel Prize along with Eric Maskin and Roger Myerson for "having laid the foundations of mechanism design theory", is also attributed to the idea of *incentive compatibility*.

A mechanism is said to be incentive compatible when no participant can make a personal improvement by revealing her private information non-truthfully. This high-level idea describes a robustness property of a mechanism, which is a focal point in this thesis. Incentive compatibility, many times referred to as *truthfulness* or *strategyproofness*, sometimes can be better appreciated due to its absence. To illustrate this consider the following example.

Example 1. Consider a very simplistic search engine that ranks web pages according to their incoming hyperlinks from other pages, breaking ties according to lexicographical order. Suppose that n webpages are participating in this ranking system, and all of them want to be in the first place[1]. Let that $N = \{1, ..., n\}$ be the set of these websites. Consider now the following delicate construction: the i-th web page has i incoming hyperlinks, from the web-pages $1, 2, ..., i$. Under this construction, the n-th web page is located first placed in the ranking and gains the benefits of the highest click-through rate. At this point the webmasters of the $(n-1)$-th website notice that by removing its hyperlink to the n-th web page, they can claim the top position for their own website. By doing this, they remove a valuable link towards the most popular website. At this point this does not seem very harmful: the most popular n-th website is located in the second position and the first position shows a website with very similar popularity.

However, after the hyperlink is removed, the webmasters of the $(n-2)$-th website realize they have a chance to climb on top on their own: they can remove two hyperlinks and their web page will reach the top of the ranking. It's not hard to see that this behavior will cascade, and eventually, the first lexicographically web page will be in the top position, with no incoming hyperlinks and a totally useless network.

While this example depicts an extreme case, it illustrates the various disadvantages we have due to a lack of strategyproofness. Under this manipulable mechanism, the webmasters have serious reasons to be strategic and try to manipulate the mechanism to their advantage. Even when they don't know the exact structure of the network they have serious incentives to devote some of their resources to find out when they will have a good chance to gain by being strategic. Furthermore, it shows that a possible case where the incentives of the participants can be aligned in a catastrophic way for the network. On the contrary, if we can replace the ranking mechanism with a strategyproof one, the web pages administrators will have no reason to use the hyperlinks strategically. Instead, they can use hyperlinks solely to add to the quality of the information on their websites.

[1]At least this is an advice one can get by various Search Engine Optimization oriented websites [14, 34]. These websites claim that the average click-through rate for the first place is about two times the average click-through rate for the second place.

1.1.2 Mechanism Design with Money

Maybe the most celebrated area of mechanism design is the design of *auctions*. In its simplest form, an auction involves a set of *strategic* individuals who are interested in buying a product from a seller. Each bidder has a valuation for that product, which one can interpret as the maximum amount of money this bidder is willing to spend to acquire this product. This is, of course, a *private* information of the bidder, and it is at its own discretion whether she will reveal it truthfully to the seller, or not. The perfect scenario for the seller is to know these valuations. This way, she can propose a price equal to the maximum valuation of the bidders and sell the price to the higher possible price. This also carries the property that the bidder who wants the item the most, gets it.

At first glance, a scenario where the bidder would wilfully reveal their true valuations of the product looks unrealistic. No player would like the price to be as high as their valuation, not even the winner, who would prefer to get the item at a lower price. Quite surprisingly, however, a very simple procedure incentivizes the bidders to reveal their true valuation for the product. This procedure, known as the *second-price* auction, asks the bidders to reveal their bids, and sells the product to the highest bidder, charging her the amount the second-highest bid proposed. This auction has truth-telling as a *dominant strategy*. When bidding truthfully, no bidder that is not getting the product has any incentive to change her bid: if she bid less than the maximum bidder, nothing changes. If she bids over the maximum bid, she will get the item but must pay the bid of the (previously) maximum-bidder, hence she will be charged more than she is willing to pay. Also, the bidder that won the item has no reason to change her bid: if she over-bids nothing happens, if she underbid, she will lose the item, for which now pays a prize lower at most equal to her true valuation.

The second-price auction, proposed by Vickrey [102], is an example of an incentive-compatible mechanism. Remarkably, this mechanism is able to identify the bidder who wants the item the most, while this information is not available in advance.

The area of *Algorithmic Mechanism Design* is established in the seminal work by Nisan and Ronen [80] and the original framework proposed uses payments as a means to provide truthfulness. The mechanism can be described with two elements, an *output* function and a tuple of *payments*. Each individual usually called an *agent*, has a bundle of available strategies. The mechanism will use the strategies proposed by the agents to decide the outcome and a payment for each agent (this can be positive or negative depending on the application). In such environments, the agents have *cardinal* utilities over the possible outcomes, which are expressed with a *valuation* function. In other words, the gain or loss of an agent is quantified. In this setting, a mechanism designer can use the payments in order to make truthtelling a *dominant strategy*[2] for the agents.

The most well-known example of successful mechanism design is the VCG mechanism by Vickrey [102], Clarke [31] and Groves [53], which extends the idea of the second-price auction. The VCG mechanism is a general method for designing truthful mechanisms. The main idea behind it is that it maximizes the *social welfare*, i.e. the sum of the valuations of all agents, and compensates the agents according to the gain of the other players. See Nisan [79] for an introduction on this method.

1.1.3 Mechanism Design without Money

The canonical example in algorithmic mechanism design (i.e. as it was introduced in [80]) uses payments as a means to impose truthfulness. In many cases in Mechanism Design, payments are not an option. This comes due to various reasons, cultural, ethical, and sometimes even practical.

A common example where payments are not allowed is the case of *elections*. When we want to elect the leader of a state, we want to find the candidate which is supported by the most people, regardless of their ability to pay. Scenarios like this are connected with *social choice* [22]. In an election, a set of voters is selecting one winner between a set of candidates. Under this paradigm, the voters have ranked preferences over the candidates and a mechanism uses the proposals of the voters to decide the winner of the election. These kinds of preferences are usually called *ordinal*. One such scenario is an election where there exist three candidates, namely Alice, Bob, and Carol. A single voter might prefer Bob to be the winner, she is fine with Carol as the winner, but she cannot accept Alice as the winner. A different voter might rank Alice on top, Carol second, and Bob in the third place. A mechanism that receives rankings of the alternatives as an input and returns a single alternative as the winner is called a *social choice* function.

For the case where only two candidates are running for office, there is a well-known strategyproof mechanism, namely the *majority* mechanism: each voter proposes one of the

[2] A dominant strategy is a strategy that is best for an agent, independently from what the other agents do. When we refer to dominant strategies we usually mean that this strategy is at least as good as any other strategy, which is sometimes referred to as *weakly dominant strategy*. When a strategy is strictly better than any other strategy we refer to it as *strongly dominant strategy*.

candidates to win the election, and the candidate with the most proposals is the winner. It is not hard to see why this mechanism is strategyproof. A voter who supports the candidate who loses the election can only misreport by voting for the winner. The outcome of the election can not be changed, and in fact, the winning candidate increases her winning margin.

This however cannot extend to cases with more than 2 candidates. The well-known *Gibbard-Satterthwaite* theorem [47, 93], states that for more than 2 candidates the only deterministic social choice function that cannot be strategically manipulated is the *dictatorship*. A dictatorship is a mechanism in which the outcome is dictated by a single voter, and in many cases, it is not an acceptable mechanism.

While the Gibbard-Satterthwaite theorem limits the choices of a mechanism designer, some escape routes have been used in the literature. The Gibbard-Satterthwaite theorem takes as an assumption that possible inputs are *unrestricted*; any ranking of the alternatives is a possible input. In many applications, however, this might not be the case. By restricting our attention to cases where the preferences of the voters have a specific structure, we can find meaningful incentive compatible mechanisms.

The case of single-peaked preferences The most well-known example of domain restriction is the use of *single-peaked* preferences. Under the single-peaked preferences assumption, each voter has a most preferred point in a one-dimensional space, which we refer to as the voter's *peak*[3]. Then each voter's preference for the other points decreases as these points move away from the mean. As an example, consider the following oversimplified scenario. In an election, the candidates can be ranked according to a *left*$\backslash$*right* measure, simulated in the $[0, 1]$ line, where 0 means extremely left and 1 means extremely right. Suppose now that we have 4 candidates: candidate a is measured at 0.1, candidate b is measured at 0.5 candidate c is measured at 0.6 and candidate d is measured at 1. The voters have also preferences located in the same $[0, 1]$ line. Consider for example a voter i with a peak at 0.4. Voter i cannot prefer d to c or c to b. Note that we don't know whether voter i prefers a to b.

Strategyproof voting rules for such scenarios exist in the literature as back as the 1940s [16]. In a seminal work Moulin [77] showed that any incentive compatible and anonymous[4] mechanism is a *generalized median rule*. One such mechanism returns the

[3]also known in the literature as *top* or *bliss point*.
[4]symmetric with respect to the voters.

median value of the voters' peaks and $n + 1$ values (where n is the number of voters) chosen by the mechanism. When only $n - 1$ phantom values are used, these mechanisms are also Pareto Optimal[5]. To see how these mechanisms work, we provide the following example.

Example 2. Consider an office where multiple employees work and the temperature of the room is controlled by an air-conditioning system. The employees want to decide on where to set the temperature of the room. In a situation like this, we can assume, without much oversimplification, that each employee has a most preferred preference (a peak) in the range $[10, 30]$ of degrees Celsius, and their preferences deteriorate as the temperature of the room moves away from their peak. Let that Alice has her peak at 20 degrees, Bob has his peak at 17 degrees and Carol has her peak at 24 degrees. A simple strategyproof mechanism is to use the median between these three values, i.e. 20 degrees. This mechanism can be implemented as a generalized median rule by using two phantom values equal to 10 and two phantom values equal to 30. Note that no voter can gain by misreporting; for example, Bob, whose peak is below the selected temperature, can only change the outcome by reporting a value higher than 20 and making himself even worse. Another strategyproof mechanism is to set the temperature to the minimum of these values, i.e to 17 degrees. This is also a generalized median rule: using 3 phantoms equal to 0 and 1 phantom equal to 30. Again, either Alice or Carol can only decrease the temperature and Bob is clearly happy with this situation. As a side note, notice that by using all 4 phantoms with a value of 10, we still have a strategyproof mechanism that sets the temperature at 10 degrees. Notice that all voters would prefer any temperature in the $(10, 17]$ interval. This happens since this mechanism is not Pareto Optimal.

Various works generalize the idea of single-peaked preferences in higher dimensions. See Section 6 from Barberà [12] for a survey. In Chapter 5 we are concerned with a family of preferences similar to single-peaked preferences, defined on the standard simplex, and we analyze a class of mechanisms influenced by the generalized median mechanisms. Elkind et al. [40] present various examples where the special structure in the voters' preferences leads to non-dictatorial truthful mechanisms.

The mechanism design literature has various other examples of successful truthful mechanism design without money. One well-known example is its usage in kidney exchange. Kidney failure is a medical condition which in many times requires a kidney transplant.

[5] This property, sometimes called *efficiency* in the Economics literature demands that whenever any voter can improve the situation for herself, the situation worsens for some other voter.

Unlike other organs, kidneys have the special feature that healthy, living persons can be donors. The market for kidneys is illegal in most countries, hence a mechanism designer cannot use payments to allocate donors to patients. In many cases, family members or close friends of a patient are willing to give one of their own kidneys, but the donors are not always compatible with their patient. Roth et al. [90] build on previous ideas from mechanism design and provided a strategyproof mechanism to re-allocate the patient-donors pairs, accommodating various constraints which arise due to the distinctive nature of the problem. This application triggered further research in the area. See Chapter 10 from Roughgarden [91] for an overview of this application.

Closely related is the Stable Matching problem introduced by Gale and Shapley [46] and similar problems, treated in [89]. For some introductory notes and more examples see the survey by Schummer and Vohra [94].

1.1.4 Approximation

In many cases, truthful mechanisms cannot be designed optimally. In the following, we will provide two examples. In the first one, the main reason is computational intractability: Truthful mechanisms that solve the problem optimally, but it's computationally hard to get the solution. The second one hits an even harder rock: sometimes truthfulness cannot co-exists with the desired goal.

As we have mentioned earlier, we can use the VCG mechanism as a tool to devise truthful mechanisms, when payments are available. In order to do this, the mechanism should compute the maximum social welfare, and using this, it can devise payments for the participants which ensures truthfulness. In some cases, however, the computation of social welfare is an NP-hard problem. The prominent example of combinatorial auctions belongs to this case. Approximate calculations for the social welfare cannot provide payments that guarantee the required truthfulness, hence a different approach is needed. To circumvent this, the literature turned on the design of truthful mechanisms which approximate the optimization objective. Some notable examples include [64, 65, 36].

In other cases, approximation enters naturally into the picture, as truthfulness might be incompatible with the maximization of the objective. One prominent example is the problem of scheduling tasks in unrelated machines: A set of n unrelated machines should process a set of tasks, and a scheduling mechanism is called to allocate the tasks to the machines. Each machine has a processing cost for each job, which is private information.

The mechanism should elicit this information from the machines. The goal of the designer is to minimize the makespan, i.e. the maximum load allocated in a single machine. A truthful mechanism in this setting is possible, by using the VCG mechanism. This mechanism, however, leads to a makespan that fares n-times worse than the makespan of the optimal allocation. Nisan and Ronen who proposed this problem in [80], conjecture that there is no deterministic truthful mechanism that can do better. A series of papers worked on this conjecture and the most recent result [30] shows that no truthful mechanism exists with makespan better than $\Omega(\sqrt{n})$.

The above examples[6] belong to the class of mechanism design *with money*. Clearly, when monetary incentives are not available, the design of optimal mechanisms becomes even more restrictive. For such restrictive environments, approximate solutions become very attractive. The agenda of approximate mechanism design without money was introduced by Procaccia and Tennenholtz [84], although the idea of using approximation for the design of truthful mechanism without money preexisted (see for example [35]). They demonstrate this concept providing truthful approximate mechanisms without money for a *facility location* problem: a facility must be located in the $[0, 1]$ line, and the voters have a cost equal to the distance of the facility (which is a special case of single-peaked preferences).

Additive approximation The "default" tool for approximation used in computer science literature is the multiplicative approximation, i.e. an α-approximate mechanism returns a solution within an α factor from the optimal solution. In this thesis, we provide additive approximation guarantees, which fit our problems better. Under this definition, an α-additive mechanism returns a solution that differs from the optimal solution at most by a term α. While less frequently used, additive approximation guarantees have been used in the *approximation algorithms* literature [106, 1, 49], in the mechanism design literature mechanisms [92, 95, 32] and even in the context of approximation mechanism without money [24, 81].

In this thesis, we present two problems for which we argue that additive approximation guarantees are better suited for the problem. The first one is the problem of impartial selection, a problem we investigate in Chapters 3 and 4. The second problem investigates the problem of allocating a budget over various alternatives, by getting proposals from various strategic individuals. We present this problem from the perspective of *participatory*

[6]both of these examples are used by Procaccia and Tennenholtz [84] to showcase the need for approximate mechanism design.

budgeting, an emerging democratic process that engages citizens in public decision-making. Both problems admit additive approximation guarantees for different reasons. In the problem of impartial selection, the prominent multiplicative approximation over exaggerates cases where even the optimal solution is not so good. In the problem of budget aggregation, the budget allocations are points in a metric space and the distance between the two points arises as a natural measure.

1.2 Impartial Selection

Impartial selection arises in communities of individuals that need to select a community member to receive an award. This is a standard social choice problem typically encountered in scientific and sports communities but has also found important applications in distributed multi-agent systems. To give an entertaining example, consider the award for the best player of the year by the Professional Footballers Association (PFA), a.k.a. the "Players' Player of the Year" award[7]. Every year, each PFA member nominates the two players they consider the best for the award, and the player with the maximum number of nominations receives the award. This award gains much prestige precisely because the members of the PFA themselves decide the winner, and past winners have highlighted this fact in their speeches [100, 83]. In distributed multi-agent systems, *leader election* (e.g., see [4]) can be thought of as a selection problem of similar flavor. Other notable examples include (see [43]) the selection of a representative in a group, funding decisions based on peer-reviewing, (see [2]) discovering the most popular node on a social network or even the Papal election (see [72] for a connection of this concept with the procedure). Example 1 is an instance of this problem and indicates how harmful a non-truthful mechanism can be in this situation.

The input of the problem can be represented as a directed graph, which we usually call *nomination profile*. Each node represents a strategic agent and a directed edge indicates a nomination (or vote) by one agent for another[8].The agents have informed opinions on who should be the winner. While the set of agents is publicly known to everyone, these opinions are private information of each agent. A first assumption is that the agents cannot nominate themselves as winners. A second assumption, crucial for the existence of strategyproof mechanisms, is that each agent cares only about winning. Alternatively, we

[7]Some basic information for the award can be found in `https://en.wikipedia.org/wiki/PFA_Players%27_Player_of_the_Year`

[8]When we refer to the problem of impartial selection we use the terms *agent* and *node* interchangeably.

say that each node is willing to be strategic if this can increase its own chance to win, but will not bother to manipulate the mechanism for another node to win. This assumption restricts the preferences of the nodes, hence allowing us to escape the Gibbard-Satterthwaite theorem [2, 57].

A *selection mechanism* (or *selection rule*) receives a nomination profile as an input and returns a single node as the winner. There is a highly desirable selection mechanism: always select the highest in-degree node as the winner. Given that every node acts honestly, the task is trivial. The selection becomes, however, challenging when dealing with strategic agents. In particular, we seek selection rules which are *impartial*[9]. An impartial selection mechanism, is robust to strategic manipulation, by not allowing any agent to affect its chance of being selected. Given the assumption that the agents may try to manipulate the mechanism only to receive the award for themselves, an impartial mechanism facilitates honest behavior. Unfortunately, the ideal selection rule mentioned above is not impartial. Recall that this rule awards the prize to the node with the most nominations, and consider the case where two nodes, say a and b, have both maximum in-degree, while they both nominate each other; let a be the winner according to some tie-breaking rule. Observe that b has a clear incentive to decrease a's in-degree by removing its outgoing edge towards a and become the sole maximum in-degree node, and thus the winner.

The problem, known as *impartial selection*, was introduced independently by Alon et al. [2] and Holzman and Moulin [57] under two different models. In this thesis, we work on both models. The first model introduced by Alon et al. [2], allows the nodes to nominate as many other nodes as they wish and even abstain. Hence, this model considers all directed graphs for a given number of nodes, excluding only self-loops and parallel edges. We refer to this model as the *multiple-nomination* model. This model is quite general and covers various practical cases where the number of nominations is not under the mechanism designer's control. For example, it could be used to find the more popular page on the web or a social network. The second main model we study is introduced by Holzman and Moulin [57]. Under this model, each node nominates a single node, and no abstentions are allowed. We call this the *single-nomination* model. Note that this model is a special case of the *multiple-nomination* model. Our results in Section 3 indicate that this model generally yields more efficient solutions than the case of multiple nominations, and in cases where it is possible to control the out-degrees (e.g., for the selection of a committee president or the

[9] In this setting, impartiality is another word for strategyproofness.

Papal election), this model is preferable.

1.2.1 Related Work and Mechanisms' Examples

In this subsection, we introduce various mechanisms already proposed in the literature, for the selection of a single agent. We start with the more intuitive case of deterministic mechanisms, and we continue with randomized mechanisms. Since we have not yet introduced notation, formal definitions of these mechanisms are given in Section 2.2 in Chapter 2.

Deterministic Mechanisms

First, we introduce deterministic impartial mechanisms. Deterministic mechanisms have not been extensively studied in the literature, since impartiality places many limitations on their design (see Section 1.2.2). However such mechanisms are better suited to build intuition and give insight into the limitations of the problem. A deterministic impartial mechanism prevents any node to become the winner, by simply changing its outgoing edges.

Fixed Winner and Fixed Dictator mechanisms The most naive mechanism of all, the FIXED WINNER mechanism awards the prize to a given node, regardless of the graph. This mechanism is trivially impartial. While undesirable, this mechanism gives us evidence of the restrictive nature of impartiality: these mechanisms are the only impartial selection mechanisms when we want the nodes to vote anonymously [57].

A similar mechanism, the FIXED DICTATOR uses the same nodes as a dictator i.e. a node that decides the winner. This mechanism is well-defined only in models where abstentions are not allowed[10]. This mechanism is also trivially impartial since the dictator cannot vote for herself.

The Plurality with Default mechanism The PLURALITY WITH DEFAULT MECHANISM is introduced by Holzman and Moulin [57]. This mechanism is an impartial adaptation of the well-known plurality rule[11]. This mechanism firstly defines (arbitrarily) a node as a *default winner*. Then the mechanism compares pairwise the in-degrees of all agents, crucially excluding the outgoing edges of the default node and any the edges between the players.

[10] In the case where multiple nominations are allowed the winner can be decided with an arbitrary tie-breaking rule.

[11] The simple rule where the maximum in-degree wins and ties are broken arbitrarily.

If there exists a single node that beats all the others in this process, then this node is the winner. Otherwise, the winner is the default node.

An example of this mechanism is depicted in Figure 1.1 and a formal description for this mechanism is given in Algorithm 2. Finally, we discuss a version of this mechanism in detail in Chapter 4. There, we show that this mechanism guarantees to select a close-to-maximum in-degree node in environments where prior information is available.

It is not hard to see why this mechanism is impartial. In every pairwise comparison, whether a node beats another node has nothing to do with the pairs' out-degrees. Hence, when any node which is not the default wins, this is not affected by its outgoing edges. The situation is even more straightforward for the default node, as this node can never affect the outcome.

The Majority with Default mechanism A similar mechanism to Plurality with Default is the Majority with Default mechanism, also proposed in [57]. This mechanism defines, again arbitrarily, a node as a default winner, as well, whose vote is ignored. Then, if there exists any node voted by the majority of the voters (a node v is voted by the majority if $\lceil n/2 \rceil$ nodes, excluding the default winner, vote for v when n is the number of nodes). Otherwise, the default agent is the winner. An example of the mechanism is depicted in Figure 1.2 and a detailed presentation of this mechanism appears in Algorithm 1. We note here that this mechanism is well-defined only for the single-nomination model.

This mechanism is impartial, for any graph in the single-nomination model. Indeed, when a node's in-degree is higher than $\lceil n/2 \rceil$, no other node can make itself the winner, since there exist no two nodes with in-degree at least $\lceil n/2 \rceil$.

The Majority with Default-Maker mechanism We have seen that a crucial feature for the impartiality of the above mechanisms is the use of a default winner. An obvious drawback of this technique is that the message of the default agent is always ignored. Under the single-nomination models, we can select instead of the default winner, a *default-maker* node. For example, the *Majority with Default-Maker* mechanism will declare as a winner any node, except the default maker, with in-degree at least $\lceil n/2 \rceil$, excluding the outgoing edge of the node voted by the default-maker.

This mechanism is impartial. To see this, let v_0 be the default-maker agent and let u be the node voted by v_0. If there exists any node u^* with an in-degree higher than $\lceil n/2 \rceil$

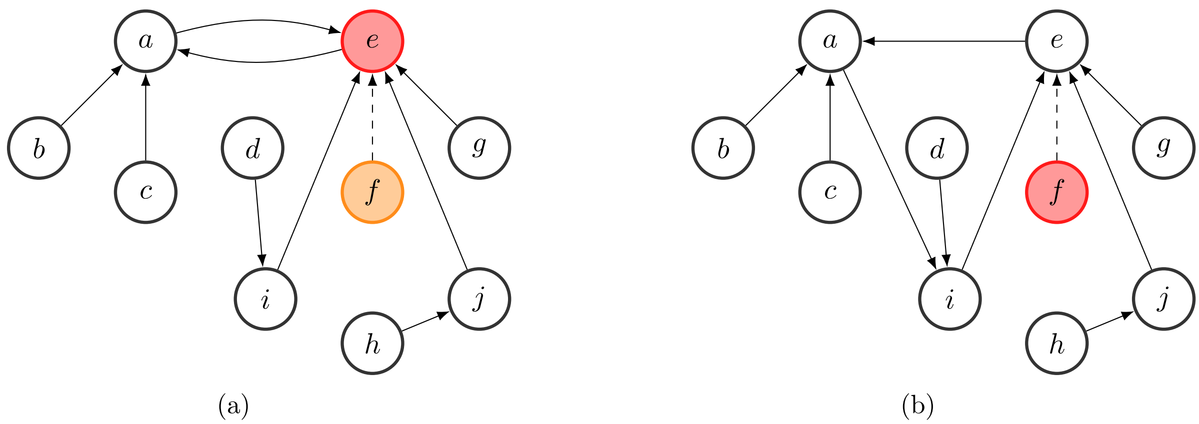


Figure 1.1: An example of the PLURALITY WITH DEFAULT mechanism. Node f is the default winner and its vote is ignored by the mechanism. Node e is the winner in profile 1.1a, as it pairwise beats all the other nodes. For example, when compared with node a, node e has in-degree 3 (excluding the votes of f and a) and a has in-degree 2 (excluding the votes of f and e). In profile 1.1b, node a re-direct its vote to node i. In a pairwise comparison between i and e, they both have in-degree of 2 (since we are ignoring votes from f and votes between i and e) and neither of them can be the winner. Node a is still beaten by node e. Hence f wins, as the default winner.

(excluding the vote of u) a single voter can decrease the in-degree of u^* by 1, but even if this breaks the majority, the winner will be u, the node voted by the default-maker. Node u cannot influence the mechanism at all. An example of the mechanism is depicted in Figure 1.3.

We note here that the default-maker node v_0 can not be allowed to win for the mechanism to be impartial. We elaborate on this with a simple example: consider a profile with 6 nodes. The default-maker v_0 is voted by 3 nodes, including node u which is voted by v_0. Vote u is also voted by 3 nodes and is the winner (its in-degree is $3 = \lceil 6/2 \rceil$ and it is not voted by itself). Now, if node v_0 changes its vote towards a node voting for u, node v_0 will become the winner, breaking impartiality.

Partition Methods The PARTITION METHODS mechanisms are a family of mechanisms proposed by [57], suited to the single-nomination model. For the sake of exposition, we present a simplified version here. The nodes are partitioned in at least 3 sets of similar sizes, called *districts*. A node in the first district is selected as the default winner, and its outgoing edge is ignored. Then, a two-step process determines the winner. Any node

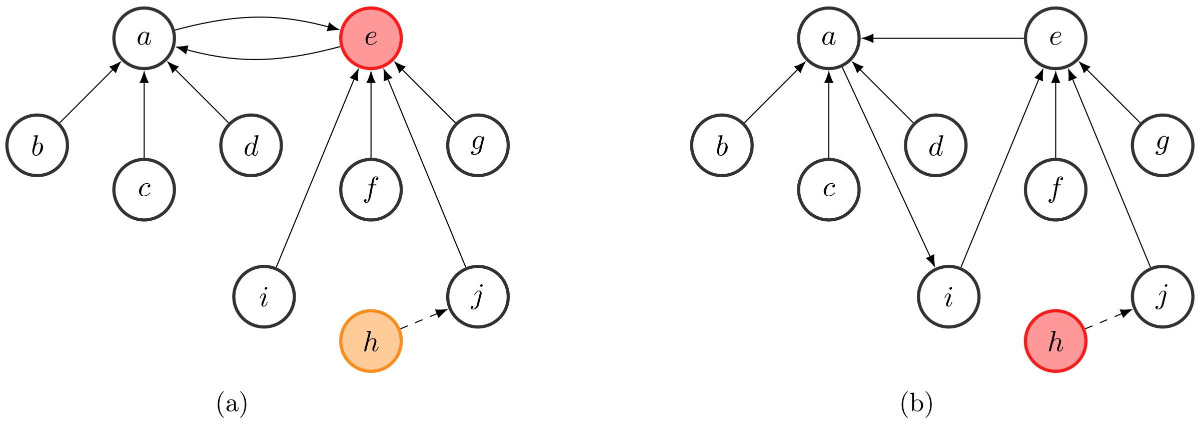


Figure 1.2: An example of the MAJORITY WITH DEFAULT mechanism. Node h is the default winner. Node e is the winner in profile 1.2a, since its in-degree, excluding the vote of the default, is 5, exactly equal to the majority quota. When node a changes its vote towards i, the majority quota is not satisfied and the default node becomes the winner.

voted by an absolute majority in each district becomes a *local winner*. If no local winner exists, the default agent becomes the winner. If only one local winner exists, this is the winner. If at least two local winners exist, the local winners become candidates. One of them with maximum in-degree from the nodes other than the local winners, becomes the winner, breaking any ties arbitrarily. In Figure 1.4 we present an example of the mechanism with 3 districts. A detailed presentation of this mechanism can be found in Algorithm 3.

Cross-Partition Methods The CROSS-PARTITION METHODS are a family of mechanisms, following the idea of partitioning the nodes in districts. This method uses a two-level partitioning of the nodes. At first, the nodes are partitioned into $K + 1$ districts. The first K of these, have nearly equal size, and the final district includes only the default winner. The nodes in the districts are partitioned again in two ways. The second partitioning has a priority on proposing eligible winners. If this fails, the districts have a chance to propose eligible winners. If this fails again, a default node is called the winner. We will illustrate this mechanism with an example. For more detail, we refer to Algorithm 5 and the discussion (i.e. regarding impartiality) of this mechanism in Chapter 2.

Figure 1.5 illustrates an example with 15 nodes. At the first level, these nodes are partitioned in 3 districts, which form a cycle. The final node is the default winner. At the second level, the nodes are partitioned in two different ways, by shape, and by color. Each

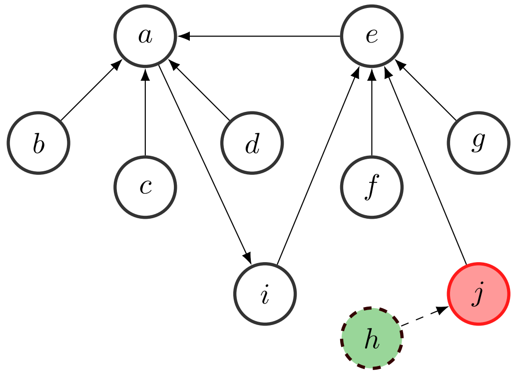


Figure 1.3: An example for the Majority with Default Maker mechanism. No node garners an absolute majority, hence the node nominated by the default-maker node h is the winner.

district must include all four combinations of shapes and colors. The mechanism first checks for a class of eligible winners called *outer heroes*. A node is an outer hero if it is nominated by a full shape component of all districts, except the district it belongs to and the previous one, and if it is also voted by a color component from the previous district. In Figure 1.5, node f is an outer hero, since it is voted by the full light color component from district 1 and the full diamond component from district 3. If outer heroes exist, they advance to the second level of voting, where all voters par the outer heroes vote for the winner. When no outer heroes exist, the mechanism searches for *inner heroes*: these are nodes that are voted by all the other nodes in their own district. If inner heroes exist, they advance to the second level where one of them is declared the winner (again, by ignoring their own votes). Finally, if no inner heroes exist, the mechanism selects the default winner.

While we will not analyze a lot this mechanism in this thesis, we present it here for two reasons. First, it shows how rich is the class of impartial mechanisms, even when we are restricted in the family of deterministic mechanisms. Second, this mechanism is proven to have an interesting property: For every triple of nodes i, j', j, there is always a nomination graph where agent i can change the winner from j to j'. This property described by Holzman and Moulin [57] as *Full Pivots*, implies every node can have the decision power to dictate the winner. This property is not satisfied by any of the previous mechanisms we have presented. Partition methods satisfy a weaker version of this property, called *Full Influence*.

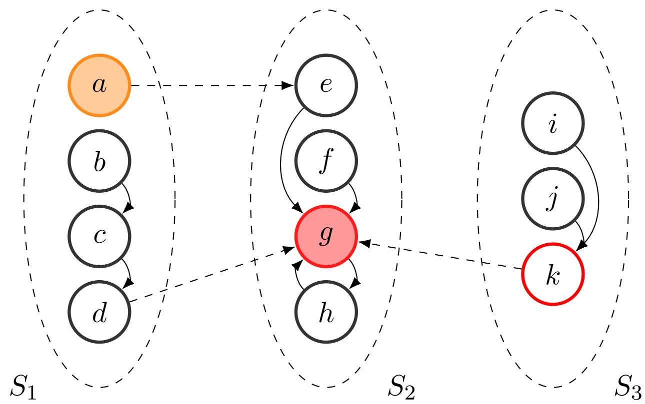


Figure 1.4: An example of a Partition Methods mechanism. In this example the nodes are partitioned in 3 districts. District S_1 must have at least 4 nodes and includes the default node a. In the first step, only votes towards the same district counts. In districts 1 and 3, any node with in-degree at least 2, becomes a local winner. In district 2 a local winner needs at least 3 incoming edges. (see Algorithm 3 for details) In the first step the local winner g and k are identified. Note that node's a vote is ignored in that step. In the second step, Node g receives 4 nominations, by nodes d, e, h and f. The node from k is ignored since k is a candidate to win. Node k receives 2 nominations, by nodes i and j.

Randomized Mechanisms We turn now our attention on randomized mechanisms, proposed in the literature. These mechanisms do not return a definitive winner, but a probability distribution over the set of individuals. In such mechanisms, a source of randomness is used for the selection of a single individual as the winner. When randomization is involved, an impartial mechanism prevents a node to change its winning probability by a change on its outgoing edges.

Uniform Random Winner and Uniform Random Dictator mechanisms These are simple mechanisms that expand the idea of the Fixed Winner and Fixed Dictator mechanisms. In the first case, the winner is selected with a probability of $1/n$. The Uniform Random Dictator, discussed in [57] and [71] is suited only with models where abstentions are not allowed. A dictator is selected with a probability equal to $1/n$ and one of its nominees becomes the winner. Under the single-nomination model, this rule assigns the prize to each individual proportionally to the number of votes it takes. Impartiality is again very easy to verify: The probability of a dictator being selected is independent of the outgoing edges of the nodes. Later on, in Chapter 3, we will show that impartiality

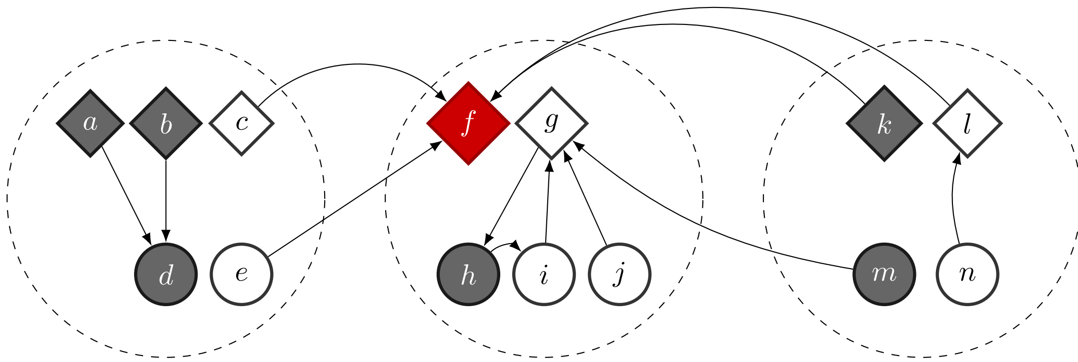


Figure 1.5: An example for the Cross Partition methods. The nodes are partitioned in 3 districts of sizes 4 and 5. Another node exists (not depicted), which acts as the default winner. The nodes are partitioned also by shape: *circles* and *diamonds*, and by color, *light* and *dark*. Each district must include all combinations of shape and color. Node f is an *outer hero*, since it is voted by the light component from district 1 and by the diamond component of district 3. Since f is the sole outer hero, it is selected a the winner by the mechanism.

demands that the dictator must be selected independently of the nomination graph for a family of mechanisms, including the Uniform Random Dictator.

The 2-Partition mechanism The 2-Partition mechanism, proposed by Alon et al. [2] is a fairly simple mechanism: It partitions the set of agents into two sets S and W. A node belongs to set S with probability 1/2. Then, a node in W with maximum in-degree from nodes in S is the winner. Any ties are resolved arbitrarily. An example of this mechanism is depicted in Figure 1.6 and the mechanism is presented in detail in Algorithm 4.

Again, it is not hard to see why the 2-Partition mechanism is impartial. Only nodes in set W have a chance to win. The messages of these nodes are ignored, hence they have no means to manipulate the outcome. A node in S can change the outcome, but cannot change it in a way to become a winner.

The Permutation mechanism The Permutation mechanism by Fischer and Klimm [43] examines the nodes sequentially following their order in a random permutation and selects as the winner the node of highest degree counting only edges with direction from "left" to "right." More specifically, the mechanism identifies the first node as a *provisional*

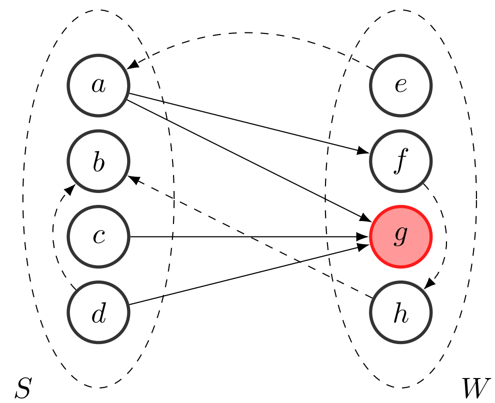


Figure 1.6: An example of the 2-Partition mechanism. The winning node g is selected as the maximum in-degree node in W from nodes in S.

winner. Then, it examines the nodes sequentially and counts the in-degree of the nodes, only from nodes that have already been examined, excluding nominations from the provisional winner. Whenever a node has an in-degree equal to the in-degree of the provisional winner, it becomes the new provisional winner. The last provisional winner is the final winner of the mechanism. An example is depicted in Figure 1.7. We present this mechanism in detail in Algorithm 7.

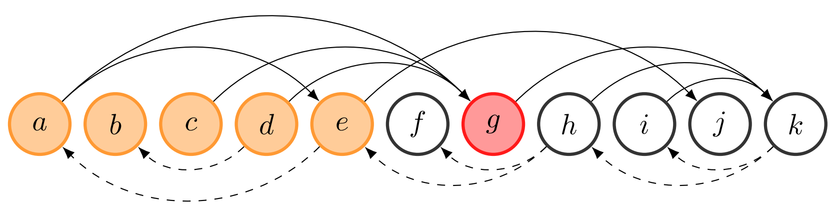


Figure 1.7: An example of the Permutation mechanism. Nodes a to e are selected as provisional winners at some point. Node g is the last provisional winner, and as a result node selected by the mechanism. Node k would be the winner if any other node votes for k.

The k-Partition mechanism A middle ground between the 2-Partition mechanism and the Permutation mechanism is the k-Partition mechanism. This mechanism first assigns each node in one of k sets $S_1, ..., S_k$, uniformly at random. The mechanism examines the sets linearly, and the node with maximum in-degree from the already examined sets is declared a provisional winner. The final provisional winner is selected by the mechanism. We present an example of this mechanism in Figure 1.8 and we present in more detail the mechanism in Algorithm 6.

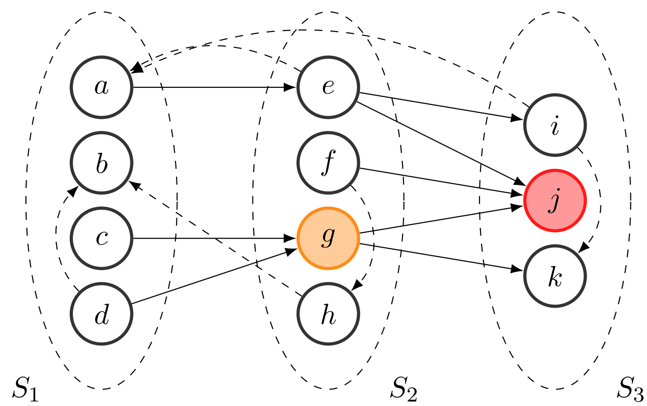


Figure 1.8: An example of the k-Partition mechanism. Note that $k = 3$. Node g becomes the winner when the second set is examined. The third provisional winner is node j, which is the winner of the mechanism. Note that when at least 1 of the edges (e, j) or (f, j) is removed the final winner would be node g.

The Slicing mechanism The Slicing mechanism of Bousquet et al. [19] is a more involved randomized mechanism, which uses ideas from the mechanisms we have presented. We provide a high-level description of the mechanism in the following. The mechanism is presented in detail in Algorithm 8.

This mechanism runs in three phases. In the first phase, a set of nodes is sampled with probability ϵ in a sample set X. Then, the mechanism calculates an estimated in-degree for the unsampled nodes voted by the sample set X, as ϵ times the in-degree of each node from the sample. In the second phase, the mechanism allocates the unsampled nodes into ϵ^{-2} sets called *slices* according to their estimated in-degree. The nodes with the highest estimated in-degree will be assigned in the last slice, while the lower in-degrees will be assigned in the lower sets. Hence, all nodes that are voted by the nodes in the sample set X are placed in slices of similarly estimated in-degree. At the final phase, the slices are examined and the mechanism maintains and updates a provisional winner, as the node with maximum in-degree from the already examined slices. The final provisional winner is selected by the mechanism, in a similar manner to Permutation and k-Partition mechanisms. A simple example of this mechanism is pictured in Figure 1.9.

While this mechanism is quite involved, impartiality is achieved by the same means as the PERMUTATION and k-PARTITION mechanisms: at each point, a node who has revealed its nominations has already lost.

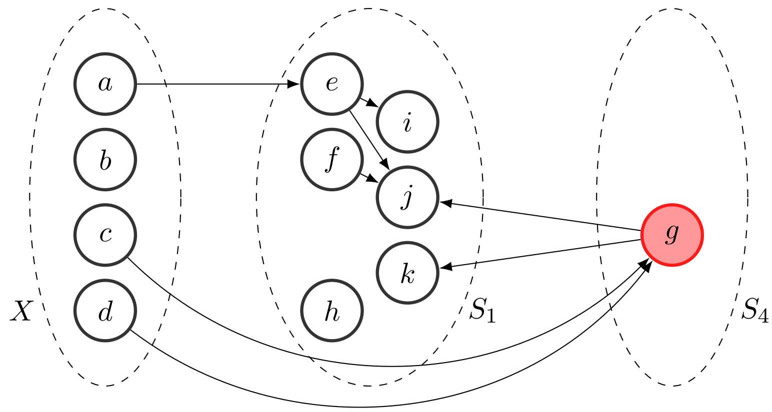


Figure 1.9: An example of the SLICING mechanism, with $\epsilon = 1/2$. The nodes in set X are selected with probability ϵ. The estimated in-degree's are 2 for nodes e, and 4 for node g, and all the rest nodes have an estimated in-degree of 0. The maximum estimated in-degree is 4. 4 slices will be created. All nodes with estimated in-degree at most 1 get in the first slice. Node g will enter the slice 4. Slices 2 and 3 (not depicted) are empty. After the nodes in slice 1 are revealed, the provisional winner does not change. Node g is the eventual winner.

1.2.2 Mechanisms' Evaluation and Limitations

In the literature involving impartial mechanisms, there are two main lines of research. One line of research, initiated by the work of [57] considers various desirable axioms and examines whether these axioms are satisfied by the mechanisms. The second line of research initiated by Alon et al. [2] defines an efficiency loss measure, namely the *approximation ratio* and compares the mechanism according to this measure. In the following, we start from the axiomatic approach, where we present various positive and negative results for impartial mechanisms. Then we continue by presenting various results involving the approximation ratio.

Axiomatic Approach

The axiomatic approach was explored mainly in the single-nomination model, hence we present the axioms for this model. We note that the presentation here is rather indicative and not precise since we have not yet introduced any notation. We will present the axioms formally in Chapter 2.

In section 1.2.1, we have referred to two axioms regarding the ability of a node to influence the selection of the winner: Full Influence and Full Pivots. A weaker axiom regarding the decision power, is the *No Dummy* axiom. This axiom demands from a mechanism to not preclude any node from being able to influence the winner at all. Both the MAJORITY WITH DEFAULT and PLURALITY WITH DEFAULT fail this axiom. One would be tempted to say that this is true for all mechanism using a default winner as last resort, but this is not true: Both PARTITION METHODS and CROSS-PARTITION METHODS satisfy this axiom: in both mechanisms, the default node has a chance to influence the selection when more than 1 eligible winners pass to the second round of voting.

A second basic axiom is *No Exclusion*. This demands a mechanism to never preclude any node from being a winner. The MAJORITY WITH DEFAULT-MAKER mechanism fails to satisfy this axiom since the default-maker can never win for impartiality to hold.

A third useful axiom is *Monotonicity*. A mechanism is monotonic when a node cannot be transformed from a winner to a loser by gaining votes. An example that fails this axiom is PLURALITY WITH DEFAULT. To see this consider the following nomination graph: Two nodes a and b, have maximum in-degree and they are voting for each other. The default node votes for someone else, hence the winner of the mechanism is the default node (note that the default node will remain the winner even if it changes its votes). If some node voting for node a decides to support the default node, then the default node turns from winner to loser. In comparison, the MAJORITY WITH DEFAULT mechanism is monotonic: the winner is either a node with an absolute majority of the votes, who will still be the winner by receiving more votes. The default node wins in cases where no majority winner exists. When the default winner gets a vote, clearly no majority can be formed.

Positive Unanimity is the natural axiom that demands a mechanism to award a node voted by all other nodes. Clearly, any mechanism which does not satisfy No Exclusion will not satisfy this axiom, too. Hence MAJORITY WITH DEFAULT-MAKER fails this axiom. On the other hand, both MAJORITY WITH DEFAULT and PLURALITY WITH DEFAULT satisfy the axiom.

We note here that Monotonicity and No Exclusion, when are both satisfied, imply Positive Unanimity. Indeed, with No Exclusion, for any node $i \in N$, there exists a nomination profile $\mathbf{x}$ where node i should be the winner. Hence by starting from the nomination profile $\mathbf{x}$, we can increase the in-degree of i until i has an in-degree of $n-1$. Due to Monotonicity, i will be the winner.

Negative Unanimity is the "reverse" axiom of Positive Unanimity. This axiom demands from a mechanism to never award a node with in-degree 0. This is satisfied by MAJORITY WITH DEFAULT-MAKER, but it cannot be satisfied by both PLURALITY WITH DEFAULT and MAJORITY WITH DEFAULT, PARTITION METHODS and CROSS PARTITION METHODS, due to the presence of the default winner.

Holzman and Moulin [57] have shown a systematic impossibility result regarding the axioms of Positive Unanimity and Negative Unanimity. No deterministic impartial mechanism exists, satisfying both axioms.

Theorem 1.1 ([57]). *There exists no selection mechanism which satisfies impartiality, positive unanimity, and negative unanimity.*

The implications of this result in the design of efficient impartial mechanisms are serious. If we decide to sacrifice negative unanimity, we may end up in scenarios where someone wins with no votes at all. If this happens when the maximum in-degree is quite small, then this could be a tolerable loss. It could be the case, however, where the maximum in-degree is quite large. The MAJORITY WITH DEFAULT mechanism, for example, may return as a winner a node with in-degree 0 where some other node is supported by nearly half the nodes (see Section 3.5 for an example). If we decide to sacrifice positive unanimity, the situation can be even worse: Since some node in-degree $n-1$ must lose at some nomination profile, then the winner will have in-degree at most 1. In Section 3.4.1 we show how a natural class of impartial mechanisms falls under this class.

An important aspect of impartial selection is concerned with the various symmetries of a problem. A mechanism is called to be *symmetric* if it is invariant in the renaming of the players. Only randomized mechanisms can be symmetric: Consider for example a deterministic mechanism and an input where all nodes vote for each other in a directed cycle. The deterministic mechanism will select a given node, say node a, as a winner with a probability of 1. Any renaming on the directed circle, however, will return the same nomination graph, and the only symmetric outcome is $(1/n, ..., 1/n)$. As noted by Holzman and Moulin [57], any randomized mechanism can be turned into a symmetric mechanism, and symmetry can be assumed without loss of generality.

The axiom of anonymous ballots treats symmetrically the nodes only as voters. One way that such a mechanism can be implemented, for example, is for the individuals to write the name of their nominee on a ballot. A sealed signature should also be present on the ballot, in order to verify that the ballot is valid (i.e. the individual did not vote for herself). Hence, when the anonymous ballots axiom is satisfied we can use, instead of a graph as an input, a tuple where each entry will show how many votes a candidate receives. It turns out however that the only deterministic and impartial such mechanism is the Fixed Winner rule [57].

In a follow-up work, Mackenzie [71] characterized all rules satisfying anonymous ballots when randomization is possible, as a convex combination between Fixed Winner mechanisms and the Uniform Random Dictatorship mechanism. For example, let $N = \{a, b, c, d\}$. Then the following mechanism is impartial and satisfies anonymous ballots: Node a wins with probability $1/3$, node b wins with probability $1/3$, and all nodes have a probability of $1/9$ to become dictators. They also characterize the Uniform Random Dictatorship as the unique mechanism satisfying negative unanimity, impartiality, and anonymous ballots. In the same work, Mackenzie [71] characterizes impartial rules which treat symmetrically the nodes only as candidates: these are the convex combinations over Fixed Dictator mechanisms. We need to note however that in [71], the incoming edges are not always treated as positive nominations, but they may be considered as negative nominations, too. Hence some of the results do not qualify directly for our models.

Approximation Ratio

A series of papers, starting from [2], evaluate mechanisms using the idea of the *approximation ratio*. The approximation ratio is defined as the maximum ratio between the maximum in-degree of a graph and the (expected) in-degree of the winner in that graph for a given mechanism. As an example, consider the profile in Figure 1.3, for the Majority with Default Maker mechanism. The winner has an in-degree of 1, while the maximum in-degree is 4, hence the approximation ratio is at least 4 for this mechanism.

Unfortunately, the above example is far from the worst-case scenario. Indeed, no deterministic mechanism can have a finite approximation ratio [2]. Consider a simple graph with just two nodes, a, and b. If we would like to have a non-infinite approximation ratio, it must be that the winner in the cases where a votes for b and b abstains, the winner must be b; and when b votes for a and a abstains the winner must be a. Yet, impartiality demand that the winner in the case where b votes for a and a votes for b to be a; and in the case where a votes for b and b votes for a, the winner to be b, which is absurd.

A similar construction provides a lower bound for any randomized impartial mechanism. We recall here that a randomized selection mechanism is a mapping between the various possible profiles and a probability distribution over all nodes. In Theorem 1.2 which follows we present this construction.

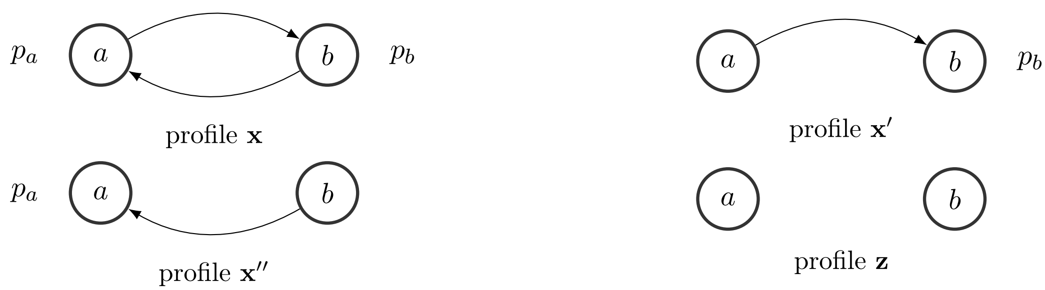


Figure 1.10: The lower bound of 2 for the approximation ratio in the multiple-nomination model. When nodes a and b change their outgoing edges, they cannot change their probability of winning.

Theorem 1.2 ([2])**.** *No randomized impartial mechanism achieves an approximation ratio smaller than* 2.

Proof. Consider any randomized impartial mechanism over the set $N = \{a, b\}$ of nodes. Let $\mathbf{x}$ be the graph where both nodes vote for each other. Let $\mathbf{x}'$ be the graph where only a votes for b and let $\mathbf{x}''$ be the graph where only b vote for a. Let (p_a, p_b) be the probabilities of winning for nodes a and b in graph $\mathbf{x}$. Clearly $p_a + p_b \leq 1$, since we cannot assign more than 1 prize. Due to impartiality the winning probability b in $\mathbf{x}'$ is p_b, and the winning probability of a in $\mathbf{x}''$ is p_a. The approximation ratio, defined as the maximum in-degree versus the probability of winning is at least $\max\left\{\frac{1}{p_a}, \frac{1}{p_b}\right\}$ which cannot be lower than 2. Figure 1.10 illustrates the proof. □

A similar proof for the single-nomination model results in a 6/5 lower bound [43].

The 2-Partition mechanism is the first mechanism evaluated with this measure and achieves an approximation ratio of 4. While the 4 approximation of the 2-Partition mechanism can be achieved by a neat argumentation, a rather more tedious approach led Fischer and Klimm [43] to the design of the k-Partition mechanism, which implies a $\frac{k-1}{2k}$

approximation, where k is the number of partitions. The PERMUTATION mechanism was revealed as a version of the k-PARTITION mechanism with infinitely many partitions, and it turns out to be the optimal mechanism with respect to approximation ratio, achieving the best possible upper bound of 2.

Bousquet et al. [19] proposed the SLICING mechanism, which achieves near-optimal approximation, given that the maximum in-degree is large enough. More precisely, they show that for any $\epsilon > 0$, there exists some constant Δ_ϵ, for which if the maximum in-degree is higher than Δ_ϵ, then SLICING mechanism yields an approximation ratio of $1 + \epsilon$.

1.2.3 Further Related Work

The divisible case Besides the papers by Holzman and Moulin [57] and Alon et al. [2], which introduced impartial selection as we study it here, de Clippel et al. [33] considered a different version with a divisible award. The goal in this version is to divide a divisible source (the *dollar*) into various partners. Each partner has an honest division in mind but would prefer to get the whole dollar for herself. An impartial mechanism, in this case, ignores any claim from an agent to herself and decides her proportion of the dollar, based on the proposals of the other agents. This model can be seen as a weighted version of the model by Holzman and Moulin [57]. Apart from impartiality, they are interested in *consesuality*: under this property, whenever all agents agree on the proposed division, the mechanism should return this division. They show that there exists a unique impartial and consensual mechanism for three agents, while for four and more agents they propose a plethora of impartial, consensual, and anonymous mechanisms. The unique mechanism for the three agents divides the dollar exactly among the three agents only when all of them are consensual. Otherwise, a small amount of the dollar is dispensed (see also follow up work of Tideman and Plassmann [101] for this mechanism). For four agents and more, their mechanisms use aggregation rules which respect the consensuality properties (for example the arithmetic mean or the median) and they show how to design consensual and impartial mechanisms, that can divide the dollar exactly among the agents. One of these rules is implemented in the fair division website `http://www.spliddit.org` (see the companion paper [50] for more details).

Multiple winners The problem of selecting exactly k nodes was also considered by Alon et al. [2]. Rather surprisingly, they show that no deterministic mechanism with a finite approximation ratio exists, for any $k \in \{0, .., n-1\}$. This implies that even when $n-1$ voters should be selected, there exists a nomination graph where only 1 node has a positive in-degree, and this node is not selected.

When randomization is allowed, they proposed the m-Random Partition mechanism, which is a generalization of the 2-Partition mechanism for the selection of 1 agent. The m-Random Partition mechanism partitions the set of agents in m subsets, and selects the (roughly) k/m maximum in-degree nodes in each subset from the other subsets, in order to return k selected nodes. They provide an approximation ratio of $1 + \mathcal{O}(1/k^{1/3})$ for this mechanism. They also provide a $1 + \Omega(1/k^2)$ lower bound for any impartial mechanism.

Tamura and Ohseto [98] first observed that when the demand for only one winner is relaxed, then impartial, negative unanimous, and positive unanimous mechanisms do exist. They propose the Plurality with runners-up mechanism which is impartial, positive unanimous and negative unanimous for at least 4 agents[12]. This rule always awards the maximum in-degree node; any other node that could be the winner by misreporting is declared the winner too. In a follow-up work, Tamura [97] characterized this rule as the minimal (in the number of winners) impartial, anonymous, symmetric, and monotonic rule.

On the same agenda, Bjelde et al. [15] proposed a deterministic version of the permutation mechanism, called the Bidirectional Permutation mechanism, which achieves a 2 approximation ratio upper bound by allowing at most two winners. This mechanism explores the nodes using a pre-determined ordering of the nodes. It selects firstly a node with maximum in-degree from the left-to-right directed nodes and then it selects a node with maximum in-degree from the right-to-left directed nodes. Note that this mechanism can select a single node, when both directions select the same node as the winner.

The randomized version of the same mechanisms yields a 3/2 upper bound on the approximation ratio for the selection of up to 2 nodes, and a combination of 2-Partition with Permutation mechanism yields an approximation ratio of 12/7 for the selection of exactly 2 nodes.

Impartiality in peer reviewing Impartiality has been investigated as a solution in peer-reviewing, a process used commonly for evaluating grand proposals, paper admission in scientific conferences, and large scale grading in Massive Open Online Courses (MOOCs)[13].

[12]For at most three agents, no impartial mechanism is positive unanimous, and negative unanimous even when multiple winners are allowed.

[13]The interested reader is referred to [25] for an introduction to the mechanism design related research in the field.

The main model here is that a set of agents, the reviewers, should evaluate various proposals. Impartiality in this context allows the reviewers to submit proposals themselves, while traditional peer reviewing methods prohibit this, for the sake of conflict of interest. A special feature of peer reviewing is that the agents cannot rate all proposals, but only a small subset of them, and also there should be multiple winners.

Kurokawa et al. [60] propose the impartial Credible Subset mechanism, which they compare with a non-impartial mechanism that assigns proposals to reviewers uniformly at random. The Credible Subset mechanism is asymptotically the best possible mechanism compared to this procedure. The Credible Subset mechanism, in a nutshell, works as follows: Each agent is assigned m proposals to rate. The top k proposals, plus any agents' proposal that may have a chance to be in the top k proposals by a change on her score to the other proposals are called *credible*. Then the mechanism returns a subset of the credible agents as the winners. To verify impartiality, firstly notice any agent that could manipulate the mechanism has a non-zero probability of being selected, which is not affected by her rates.

Aziz et al. [8] propose the Exact Dollar Partition mechanism, inspired by the work of de Clippel et al. [33]. This mechanism partitions the agents into clusters of similar sizes. The agents review proposals outside of their own cluster, for which they assign scores. The scores are normalized so that each agent assigns a total score of 1 to all the proposals they review. Based on the score each cluster j get a *dollar share* x_j. Then each cluster will return as winners either the $\lfloor x_j \cdot k \rfloor$ or the $\lceil x_j \cdot j \rceil$ top-ranked agents, according to a carefully designed procedure. The impartiality of this mechanism is not trivial. An agent cannot manipulate the mechanism by affecting her score or the score of others in the same cluster. It may however manipulate the number of winners returned from her cluster. For this, the authors propose a randomized allocation process that decides whether the j-th cluster will be allocated $\lfloor x_j \cdot k \rfloor$ or $\lceil x_j \cdot j \rceil$ winner and the agents cannot influence the expected number of winners.

In a slightly different setting, Kahng et al. [58] design impartial ranking aggregation mechanisms, i.e. mechanisms where the position of an agent in a ranking is independent of the agent's report. Other notable mechanisms have been proposed in [73, 107].

Influential agent selection In a context similar to ours, Babichenko et al. [11] proposed impartial mechanisms where the goal is not the maximum in-degree, but the node with maximum *progeny*[14]. They focus on directed forests and provide randomized mechanisms for the selection of a single node and they provide approximation ratio guarantees. They are also interested in a notion of fairness, which demands that higher progeny leads to higher selection probability and that the ratio between selection probabilities should depend only on their progenies. Interestingly, any impartial mechanism which satisfies this fairness notion either has an infinite approximation ratio or does not always return a winner. Other works in the same field include [10, 108].

1.3 Participatory Budgeting

Participatory budgeting is an emerging democratic process that engages community members with public decision-making, particularly when public expenditure should be allocated to various public projects. After its initial adoption in the Brazilian city of Porto Alegre in the late 1980s [23], its usage has been spread in over 1500 municipalities across the world [5]. Its adoption by large urban centers like New York [96], Madrid [5], Paris [5], San Francisco [48], and Toronto [104] provides an indication of the trend. The process is being used at various levels of government, from neighborhood level up to national government level, but is most widely used at district and city levels. It gains support from a diverse audience, including activists, politicians, and international institutions like the World Bank and the European Union [104]. The well-implemented Participatory budgeting processes are attributed to increased social justice and citizens' welfare, improved citizen empowerment and engagement in politics, and are generally considered a tool for a better working democracy [105, 51, 103, 74].

Participatory budgeting is a long and quite complicated process. Usually, the provided budget is decided at first, and then a series of various meetings start, where ideas are shared and preliminary proposals are examined. The proposals go through several rounds of deliberation and a small number of them make the ballot for the final vote. At this stage, the eligible voters cast their vote on how the budget should be divided [5]. While each of these stages is of its own interest, we focus mainly on the final stage, where the residents vote to decide the outcome.

[14]The progeny of a node in a directed graph is the number of nodes with paths to that node.

1.3.1 Truthful Budget Aggregation

In this thesis, we focus on a specific problem in participatory budgeting, where a set of voters need to decide how a budget can be divided among a set of candidate projects. Under this model, it is assumed that all projects can be implemented partially and each voter provides a division of the budget. Then a *budget aggregation mechanism* will receive these divisions as an input and return an *aggregated division*. To build intuition, we present a stylized example.

Example 3. Consider the local council of a small town, which wants to implement a participatory budget procedure. A budget of 1 million pounds has been approved to be spent for improvements in local schools, maintenance costs in mass transit, and for financial programs to support the development of local businesses. There are 10000 residents in this town and the local council wants to use proposals from the residents to decide on how to spend the budget among these three sectors. Hence, each resident is asked to propose a division of 1 million dollars over the three projects. In Table 1.1 we present a summary of the proposals the council gets[15]. A simple way that the council can decide the final budget assignment is to use the *arithmetic mean*. This is the simple rule where each project is assigned the normalized sum of the proposals regarding this project. For example, the proposals in Table 1.1 lead to an aggregate division equal to $(0.2555, 0.5375, 0.2075)$ million dollars for the 3 different projects.

No of Voters	Schools	Mass Transit	Local Businesses
1500	0.5	0.25	0.25
2000	0	0.5	0.5
500	0	0	1
5000	0.3	0.7	0
1000	0.3	0.5	0.2

Table 1.1: Some proposals from various voters for our example. The first column refers to the number of voters for that proposal. Columns 2 to 4 refer to the percentage of the 1 million pounds each voter assigns to each project.

Splitting the budget using the arithmetic mean (or proportionally) has various positive features. At first, it is easy to be implemented and computationally not demanding. It is also defined naturally for an arbitrary number of projects. Even more, the general public is

[15] Note that it is possible to have 1 different proposal from each voter.

familiar with the concept of the arithmetic mean, an aspect that can help in communicating the process. It has, however, a serious drawback. Any voter can manipulate the outcome. To see this, we will use an example with fewer voters[16]. Consider an example where two voters are asked to divide a budget among two projects. Suppose now that one of these votes is a keen supporter of the first of the two projects, i.e. she proposes the division $(1, 0)$. Voters like this are called *single-minded* voters. The second voter proposes the division $(0.5, 0.5)$. The arithmetic mean in such mechanisms splits the budget to $(0.75, 0.25)$. The second voter can benefit by proposing the division $(0, 1)$ which leads to her preferred division $(0.5, 0.5)$.

In this thesis, we use the arithmetic mean as a benchmark and we propose the design of truthful mechanisms, which are *as close to the arithmetic mean* as possible. We describe this in detail in Chapter 5. In the following, we present some important truthful mechanisms for budget aggregation from the literature. This presentation acts as an introduction to our work.

1.3.2 Related Work and Mechanisms' Examples

We start our presentation of incentive compatible mechanism from the simplest case, where the budget should be divided among *two* different projects. In this case, the generalized median rules from [77], tailored to our setting, are a suitable family of incentive compatible mechanisms [44]. The generalized median rules are explained in more detail in Section 1.1.3. We can define one such mechanism in our setting by using n arbitrary phantom values in $[0, 1]$ for the first project, and their complementary values for the second project. Then the outcome of the mechanism is the coordinate-wise median between the voters' proposals and the phantom values for each project. The incentive compatibility of these mechanisms is a direct consequence of Moulin's famous characterization. Note that this scheme is *anonymous* in the sense that all voters are treated symmetrically. Under the mild assumption of continuity[17], the generalized median rules are the only anonymous and continuous incentive compatible budget aggregation mechanisms [44, 77]. One such mechanism is also *neutral*, i.e., it treats the projects symmetrically, when the set of phantom values is the same for both projects [44].

Two prominent examples are the *Uniform Phantom* mechanism and the *Coordinate-Wise*

[16] The mechanism remains non-truthful for any number of voters. It is however easier to illustrate this effect in small instances.

[17] in the sense that the mapping between the preference profiles to the outcomes is continuous.

(a) Uniform Phantom (b) Coordinate-Wise Median

Figure 1.11: Examples of the Uniform Phantom and the Coordinate-Wise Median mechanisms, in an input with 3 voters. The input is the same for both examples: the first voter proposes the division $(1, 0)$, the second voter proposes the division $(0.70, 0.3)$ and the last voter proposes the division $(0.5, 0.5)$. The Uniform Phantom mechanism uses 4 phantom values uniformly distributed in the $[0, 1]$ line. The Coordinate-Wise Median uses two phantom values at 0 and two phantom values at 1. The solid vertical lines denote voters' proposals, the dashed vertical lines denote the phantom values and the rectangles the median at each project.

Median mechanism from [44]. The former distributes the n phantom values uniformly in the $[0, 1]$ line. The latter splits the phantom values to the extremes: if $n + 1$ is even, then half the phantom values are equal to 0, while the rest are equal to 1; otherwise, there exists a phantom value equal to $1/2$, while the remaining phantom values are split equally to 0 and 1. In Figure 1.11 we present two instances of these mechanisms. The Uniform Phantom mechanism appears under different contexts in various works [26, 86, 87, 9], and we analyze it further in Chapter 5.

It is not hard to see that the generalized median rules cannot be used for more than two projects, as they do not guarantee feasible outcomes. However, the literature provides us with various incentive compatible mechanisms which are useful for any number of projects. A first incentive compatible budget aggregation mechanism is the *Fractional Knapsack Voting* mechanism, proposed by Goel et al. [48]. The same mechanism was proposed also in an unpublished paper by Lindner et al. [68][18]. This mechanism works as follows: There exists a given budget B, which is split into k *dollars*, of denomination ϵ, such that $k \cdot \epsilon = B$. Each voter is assigned k dollar. For each project, there exist k ballot boxes, ordered from 1 to k. Each voter assigns dollars to the bins, starting from bin 1 to meet her preferred allocation for that project. Then the mechanism selects the dollar bins with the maximum number of tokens so that the budget is not exceeded. In the case of a tie, the mechanism uses a lexicographic tie-breaking rule.

[18] This work is analyzed mostly in Lindner's Ph.D. thesis [67], written in German language. For a summary of its results for budget aggregation under ℓ_1 preferences see the discussion in [44].

	10	20	30	40	50	60
Project 1	1,2,3,4	1,2,4	1,2,4	1,2,4	1,4	4
Project 2	1,2,3,4	1,3,4	1,3,4	1,3,4	1	
Project 3	2,3	2,3	2,3	2,3	2	

Figure 1.12: An example of the Fractional Knapsack Voting. Four voters propose divisions for three projects. The budget is 100 pounds. The first voter splits the budget equally among projects 1 and 2, hence it assigns 5 dollars (with denomination 10) to each. The second voter, proposes the division $(0.4, 0.1, 0.5)$, hence she assigns her dollars accordingly. Similarly, the third voter proposes $(0.1, 0.4, 0.5)$. The last voter proposes the division $(0.6, 0.4, 0)$. The above table shows how the dollars are split into ballots. The final allocation is decided by the ballot boxes with the maximum number of dollars: the ballots with 4 and 3 dollars cover 80 per cent of the budget, and up to now, 0.4 is assigned to both projects 1 and 2. The remaining budget is decided by the ballot boxes with 2 dollars each. There exists 5 such ballot boxes and the mechanism can choose only 2. Using a lexicographic tie breaking rule, it assigns an extra 0.1 portion to the first project and a 0.1 portion to the third project. Eventually, the outcome is $(0.5, 0.4, 0.1)$.

.

The Fractional Knapsack Voting mechanism is easier to be described using the stylized example in Figure 1.12. Consider a budget of 100 pounds, four voters, and three projects. We use 10 dollars of denomination 10. The four voters propose the divisions $(0.5, 0.50, 0)$, $(0.4, 0.1, 0.5)$, $(0.1, 0.4, 0.5)$ and $(0.6, 0.4, 0)$ respectively, and the outcome of the mechanism is the division $(0.5, 0.4, 0.1)$ over the whole budget.

The Fractional Knapsack Voting mechanism is incentive compatible when the agents are *single-peaked* with ℓ_1 disutilities. Under this type of preferences, a voter has a single most preferred division $\mathbf{v}^*$, her *peak*, and for each other division $\mathbf{x}$ she suffers a disutility equal to the ℓ_1 distance between the division $\mathbf{v}^*$ and $\mathbf{x}$. We refer to these utility preferences as ℓ_1 *preferences*, following the work of Freeman et al. [44].

Apart from being strategyproof, the Fractional Knapsack Voting mechanism maximizes the *utilitarian social welfare*. This is a function that sums the total utility of the voters, and it is a well-known efficiency measure in computational social choice. For ℓ_1 preferences, maximizing the social welfare is achieved by minimizing the sum of disutilities for the voters. To build some intuition on social welfare maximizing for ℓ_1 preferences, we present the following toy example: Consider two voters with ℓ_1 preferences, with peaks at

$(1,0)$ and $(2/3,1/3)$ respectively. Any aggregated division $(x,1-x)$ achieves a social welfare of $-2/3$, for $x \in [2/3,1]$ and $-10/3-4x < -2/3$ for $x \in [0,2/3)$; hence the social welfare is maximized for any $x \geq 2/3$.

Freeman et al. [44] moved a step further and proposed a wide family of strategyproof mechanisms for ℓ_1 preferences which they called *moving phantom mechanisms*. This family of mechanisms is inspired by the *generalized median rules* from [77]. In a nutshell, moving phantom mechanisms with n voters use $n+1$ *phantom* values and returns for each project the median between the n voters' reports and the $n+1$ phantom values. The phantom values are selected from a carefully designed pool: $n+1$ non-decreasing, continuous functions over the $[0,1]$ line, which at each point of the $[0,1]$ line return a weakly ordered set of phantom values. The properties of these functions ensure that the mechanism is both strategyproof and returns a feasible solution (i.e. an aggregated division which sums to 1). Two important mechanisms are depicted in Figure 1.13.

Moving phantom mechanisms are anonymous, neutral, and continuous. This family of mechanisms, defined formally in Definition 10, is very general and it is still an open problem if this family can characterize all strategyproof mechanisms that are anonymous, neutral, and continuous for more than 2 projects. For the case of 2 projects Freeman et al. [44] answered this question affirmatively.

The Fractional Knapsack Voting mechanism is not a moving phantom mechanism, since this mechanism is not neutral; when the Fractional Knapsack Voting has more than 1 candidate divisions maximizing the social welfare, it returns a division based on lexicographical order. A mechanism that maximizes the utilitarian social welfare, can be implemented as a moving phantom mechanism, with a neutral tie-breaking rule as shown in [44]. We refer to this mechanism as the Utilitarian Moving Phantom mechanism, which is a generalization of the Coordinate-Wise Median mechanism. This mechanism has the special feature that at least n of its phantoms have a value of either 0 or 1 and at most, 1 phantom can have a value in $(0,1)$. By using this feature, Freeman et al. [44] show that this is the unique *Pareto optimal* moving phantom mechanism. In a Pareto optimal mechanism, whenever any of the voters improves on her personal loss, some other voter will suffer more. In other words, a Pareto optimal mechanism never returns a division where there is a better division for every voter. In Figures 1.13a and 1.13c we present an example of this mechanism, on the same input as the example of Figure 1.12. Observe the difference in the tie-breaking rule between this mechanism and the Fractional Knapsack Voting mechanism in these examples.

A drawback of social welfare maximizing mechanisms is that they tend to return divisions which are unfair in some situations. An intriguing example is the following: consider an instance with two projects, and let that $k+1$ voters propose $(1,0)$ while k voters propose $(0,1)$. Notice that any social welfare maximizer can only return the aggregate division $(1,0)$[19]. This is a scenario where the minority is submitting to the will of the majority, which is not uncommon in the economics literature (see for example [7, 86] for problems with similar dilemmas). Clearly, under the participatory budgeting paradigm, when a budget is to be split over two projects, we would expect a division closer to $(1/2, 1/2)$ with an input like this.

As a solution to this problem, Freeman et al. [44] proposed a natural fairness property, called *proportionality*. Proportionality demands a mechanism to return the proportional division when the input includes only single-minded voters. As an example, if 4 voters propose $(1,0,0)$, 2 voters propose $(0,1,0)$ and 2 voters propose $(0,0,1)$ a proportional mechanism should return the division $(0.5, 0.25, 0.25)$. For the case of two projects, they have shown that the Uniform Phantom mechanism, is the unique incentive compatible and proportional budget aggregation mechanism, under the mild assumptions of anonymity and continuity.

Freeman et al. [44] proposed also a proportional moving phantom mechanism, called the *Independent Markets* mechanism, which generalized the Uniform Phantom mechanism. This mechanism uses $n+1$ phantom values, which are arranged uniformly in the interval $[0, x]$, where the value $x \in [0,1]$ is selected in such a way to provide a feasible solution. When the input consists solely of single-minded voters, then the Independent Markets mechanism simulates the Uniform Phantom mechanism, by setting $x = 1$.

The "Independent Markets" name arises from an alternative interpretation of the mechanism. The aggregated division decided by this mechanism is equivalent to a market system. More precisely, each project corresponds to a market, where x units of a divisible good are sold. For all markets, the quantity x is the same. Each voter has a value for the good of each market, equal to its preference for the corresponding project, and has one pound to spend. By increasing the amount x until the market clearing prices sum to 1, the aggregate division selected by the Independent Markets mechanism is the same as the market-clearing

[19] To see this, assume that there exists any division $(x, 1-x)$ for $x \in (0,1)$ which maximizes the social welfare. The disutility for the $k+1$ voters proposing $(1,0)$ is $-(|x-1| + |1-x-0|) = -2(1-x)$. Similarly for the other k voter the social is $-(|x-0| + |1-x-1|) = -2x$. The social welfare is $= -2 - 2k + 2x$ which is maximized for $x = 1-$ a contradiction.

prices computed.

Independent Markets satisfies also a natural generalization of proportionality, called *k-proportionality*. Under this property, a mechanism returns the proportional division for any input where all voters are k-uniform. A k-uniform voter proposes $1/k$ for k out of m projects, and 0 to the rest. A mechanism is k-proportional whenever it returns the proportional division when the input consist only of k-uniform voters. For example, when 2 voters propose $(0.5, 0, 0.5)$, one voter proposes $(0, 0.5, 0.5)$ and another voter proposes $(0.5, 0.5, 0)$, the outcome of a k-proportional mechanism should be $(0.375, 0.25, 0.375)$. As Freeman et al. [44] noticed, the obvious generalization to k_i-proportionality, i.e. a property which demands from a mechanism to return the proportional division even when each voter is k_i-uniform is incompatible with strategyproofness.

1.3.3 Further Related Work

The current literature on participatory budgeting is quite broad, involving various models which differ in various directions: whether the projects can be partially implemented or not, according to the assumptions on voters' preferences, the various possible ballot designs, and more. We refer to the recent survey by Aziz and Shah [5] for a thorough presentation of the various models and approaches in the area. Here we present some prominent examples. The interested reader can also find real-world examples of participatory budgeting processes in the digital voting platform (`https://pbstanford.org/`), which is connected to the work of Goel et al. [48]. Among other examples, it presents a case[20] very similar to the divisible case we study here.

A large part of the literature concerning participatory budgeting covers a model where projects cannot be funded partially but are either fully funded or not funded. Part of the work of [48] is dedicated to this model. Other notable examples include [13, 6, 70].

In Chapter 5 we propose a new fairness criterion for budget aggregation mechanisms, where we compare strategyproof mechanisms with a mechanism that assigns the budget proportionally (according to the proposals of the voters). The approval-based setting sometimes referred to as dichotomous preferences, has seen considerable work in defining various notions of fairness. This setting was introduced by Bogomolnaia et al. [17], and the voters either approve or disapprove of the funding of each project. The utility of a voter in this setting is equal to the fraction of the budget spent on approved projects. A

[20] `https://pbstanford.org/boston16internal/knapsack`.

simple fairness criterion, *individual fair share* in that setting requires that every agent (which approves at least one project) have a utility of at least $1/n$. More complex fairness criteria have also been introduced, which require particular groups of voters to gain enough utility [7, 42, 37]. An approach similar to ours is used by Michorzewski et al. [75]. Therein, an approximation measure is proposed, where various voting rules are evaluated against the utilitarian social welfare. A similar approach is taken by Tang et al. [99], where they evaluate various voting rules against the egalitarian social welfare.

Recently, Laraki and Varloot [61] proposed a notion of strategyproofness for a quite related problem: Various experts propose probability distributions for certain phenomena, e.g., they may predict the danger of extremely high levels of temperature in the following days. The prediction would be a probability distribution over various danger levels. Like the model we explore here, these distributions should be aggregated to a single distribution. The crucial difference is that the danger levels are not independent, but they induce a Cumulative Probability Function.

In another recent work, Puppe and Rollmann [85] conducted an experimental study between a *normalized median rule* and the aggregated mean, which are both non-truthful. Their findings show that under the normalized-median rule, people were frequently sincere, while under the aggregated median rule, the voters' behavior was mainly polarized towards the extremes.

Finally, the budget aggregation problem is related to the well-studied problem of *facility location*. For a recent survey concerning mechanism design for facility location problems, see Chan et al. [29].

1.4 Contribution

In this section, we provide a brief summary of our results. In Chapters 3 and 4 we focus on the problem of impartial selection. In Chapter 5, we focus on the on the problem of truthful budget aggregation.

In Chapter 3 we propose and analyze impartial mechanisms, slightly deviating from the current literature trends. Motivated by the fact that (a) the approximation ratio fails to classify deterministic impartial mechanisms and (b) that near-optimal mechanisms exist when the in-degree of a nomination graph is large, we propose the *additive approximation* measure, as an alternative optimization objective in the design of impartial mechanisms. Additive approximation is defined as the *difference* between the maximum in-degree of a nomination graph and the in-degree of the node selected by the mechanism. Under this metric, the cases where the maximum in-degree is small, have a minimal impact.

On the positive side, in Section 3.3 we propose two randomized mechanisms with *sub-linear* (with respect to the number of nodes) additive approximation guarantees. These mechanisms are simple. The first mechanism, which we call *Sample And Vote* is defined for the single-nomination model and guarantees an additive approximation of $\mathcal{O}(\sqrt{n})$. Our second mechanism, *Sample And Poll* is defined for the more general, multiple-nomination model, and guarantees an additive approximation of $\mathcal{O}(n^{2/3} \log^{1/3} n)$.

In Section 3.4, we complement our positive results with lower bounds. At first, we propose the class of *Sample* mechanisms, a wide class that includes all mechanisms which for a given input, firstly select a sample set of voters and this set defines all the eligible winners. No restrictions are set on how the sample set is decided. We analyze a special subset of Sample mechanisms, which we call *Strong Sample* mechanisms. A strong sample mechanism is a mechanism where the sample set cannot change by a single change of a vote in the sample. In a sense, we impose impartiality on the winners and on the members of the sample set. For this class of mechanisms and the single-nomination model, we provide a characterization, showing that the sample set should be a fixed set of nodes, independent of the nomination graph. This implies that any deterministic and impartial strong sample mechanism yields an additive approximation of at least $n-2$. The same characterization reveals an $\Omega(\sqrt{n})$ lower bound for the case of randomized impartial strong sample mechanisms. We note here that both Sample And Vote and Sample And Poll mechanisms belong to the class of strong sample mechanisms. The former one is asymptotically best possible, among all strong sample mechanisms for the single-nomination model. For the multiple-nomination model, we obtain a lower bound of 3 for the additive approximation of any deterministic impartial mechanism. We note here that the best known upper bound, for deterministic impartial mechanisms in the multiple-nomination model is $n-1$, which leaves an enormous gap to be closed. Closing this gap is the most interesting open problem imposed by this thesis. Finally, in Section 3.5 we present additive approximation bounds for various known impartial mechanisms.

In Chapter 4 we remain on the problem of impartial selection, exploring now a setting where the mechanism designer has some information on the nodes' preferences. Under this model, the nodes receive their preferences from probability distributions over the set of the remaining nodes, a piece of information that is publicly available to everyone. The realizations (i.e. the edges of the nomination graph) are still private knowledge of each node. We assume also *voter independence*, in the sense that the nodes get their preferences from independent probability distributions. We build a hierarchy of information models: In the higher level, we have the *Opinion poll* model, where each directed edge (i, j) in a graph exists with probability p_{ij}. Then we propose the *a-priory popularity* model, where each node j has a popularity index $p_j \in [0, 1]$, and each directed edge towards j exists with probability p_j. Finally, we propose the uniform model, where each edge appears with probability $p \in [0, 1]$, and which we use for lower bound construction.

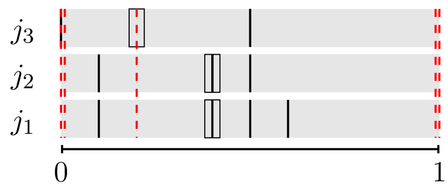
Section 4.3.1 works as a warm-up and there we analyze the extremely simple *Fixed Winner* mechanism, which always assigns the prize to the node with the (a-priori) maximum in-degree. We show that under this model even this naive mechanism achieves an additive approximation of $\mathcal{O}(\sqrt{n \log n})$ for the Opinion Poll model. In Section 4.3.2, which follows, we propose the *Approval Voting with Default* (AVD) mechanism, getting inspiration from the work of Holzman and Moulin [57]. This yields a substantial improvement when applied to inputs from the a-priory popularity model, guaranteeing an additive approximation of $\mathcal{O}(\log^2 n)$. En route to showing this upper bound we provide an upper bound for the *hazard rate* of the binomial distribution, of independent interest. Finally, in Section 4.4.2 we complement our upper bound for AVD with a nearly matching bound of $\Omega(\ln n)$.

In Chapter 5 we turn our attention to the problem of truthful budget aggregation. Under this model, a set of strategic voters propose budget allocations for projects, and an *aggregation* mechanism needs to come up with a joint budget allocation. Following previous work, we emphasize on the class of *moving phantom mechanisms*, which are known to be truthful, and particularly on the *Independent Markets* mechanism. This mechanism, apart from being truthful bears a useful *proportionality* guarantee, which implies that for inputs where all voters are supporting single projects only, the budget will be allocated proportionally. We build on these ideas and we propose a quantitative measure for the loss of the proportionality due to truthfulness, for any possible input. We call our measure the ℓ_1-loss, and it is defined as the ℓ_1 distance between the aggregated budget and the proportional division.

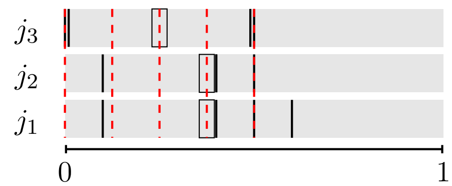
In Section 5.3.1 we show that the *Uniform Phantom* mechanism of [44] guarantees an ℓ_1-loss of at most 1/2 in the worst-case. As we show in Section 5.4.1, this is best possible among any truthful mechanisms. In Section 5.3.2 we propose a new moving phantom mechanism, which we call the *Piecewise Uniform* mechanism. Our mechanism guarantees an ℓ_1-loss no larger than 2/3. As we show in Section 5.4.2 this is best possible among the family of moving phantom mechanisms. Our result is an improvement compared to the INDEPENDENT MARKETS mechanism, for which we show that the ℓ_1-loss can be as large as 0.6862 in some instances. We complement our results with lower bounds for an arbitrary number of projects. In Section 5.4.4, we provide a lower bound of $2 - \Theta(m^{-1/3})$, for both the PIECEWISE UNIFORM and the INDEPENDENT MARKETS mechanisms. This lower bound is asymptotically close to the worst possible value for the ℓ_1-loss. Finally, in the same section, we show that any *utilitarian* social welfare maximizing mechanism yields a $2 - \Theta(m^{-1})$ ℓ_1-loss. This bound is quite large even for a modest number of projects.

1.5 Roadmap

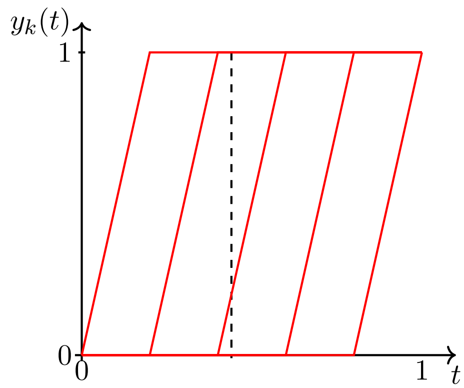
In the following chapters, we discuss our results in detail. Chapter 2 introduces some basic notation and some general formal definitions for our models. It also includes formal definitions of the mechanisms and axioms discussed in the previous sections. Chapter 3 discusses the problem of impartial selection and introduces the concept of additive approximation guarantees. Chapter 4 discusses the problem of impartial selection when prior information is available. Chapter 5 is devoted to the problem of truthful budget aggregation. Finally, in Chapter 6 we discuss some open problems and future directions.



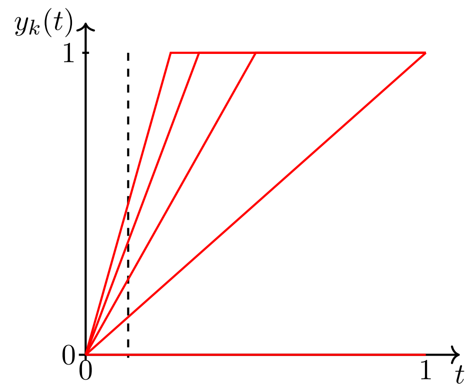
(a) Utilitarian- Example



(b) Independent Markets - Example



(c) Utilitarian - Phantom pool



(d) Independent Markets - Phantom Pool

Figure 1.13: Examples for the Utilitarian and the Independent Markets mechanisms, which are both moving phantom mechanisms. The top figures show an example with four voters, proposing the divisions $(0.5, 0.5, 0)$, $(0.4, 0.1, 0.5)$, $(0.1, 0.4, 0.5)$ and $(0.6, 0, 0.4)$, for both mechanisms. The Utilitarian Moving Phantom mechanism returns the division $(0.4, 0.4, 0.2)$, while the Independent Markets mechanisms returns the division $(0.375, 0.375, 0.25)$. The lower figures show the phantom pool: 5 continuous functions -with solid lines- from which the phantom values are selected. For each value $t \in [0, 1]$, the values $(y_1(t), ..., y_5(t))$ are selected as phantoms values. The vertical dashed line shows the selected phantom values for the respective examples. These phantom values yield feasible outcomes.

Chapter 2

Preliminaries

In this chapter, we introduce basic notation and preliminary definitions that are used in multiple chapters. More specifically, we formally present here core elements of our model regarding the impartial selection problem, discussed in detail in Chapters 3 and 4. In both these chapters, we introduce more specialized definitions and specific notation when needed. We also provide formal descriptions of some mechanisms discussed in the previous chapter. The formal model for Chapter 5, regarding the problem of truthful budget aggregation, is introduced in that chapter.

2.1 Basic Notation

In this section, we explain part of the notation we use throughout this thesis. While the notation we use is rather standard, we explain some symbols, for the sake of clarity and completeness.

Natural numbers are denoted by $\mathbb{N} = \{1, 2, 3, ..., \}$. We use the notation $\mathbb{N}_{\geq z}$ to denote the subset of $\mathbb{N}$, which includes every natural number at least equal to $z \in \mathbb{N}$. Real numbers are denoted by $\mathbb{R}$.

We denote sets using capital Latin letters. Usually, the set $N = \{1, ..., n\}$ is reserved to denote the set of strategic individuals, for some $n \in \mathbb{N}_{\geq 2}$. When convenient, we use the abbreviation $[k] = \{1..k\}$ for any $k \in \mathbb{N}$. We use standard notation for operations on sets, e.g., we use $|A|$ to denote the cardinality of set A and the powerset of a set A is denoted as 2^A. Usually, we denote the difference between two sets A and B as $A \setminus B$. In some cases, we use the abbreviation A_B interchangeably to $A \setminus B$.

We use $[a, b]$ to denote an interval in the real line where the endpoints are included in the interval, i.e. the $\{x : a \leq x \leq b\}$. We use (a, b) to denote intervals where both endpoints are excluded, i.e. the set $\{x : a < x < b\}$. We use the notations notations $(a, b]$ and $[a, b)$ accordingly.

We use $\mathbb{1}\{\chi\}$ to denote the indicator function which gets the value 1 when the logical preposition χ is true and 0 otherwise.

We use bold variables to denote tuples. For the tuple $\mathbf{x}$, x_i denotes the i-th element of $\mathbf{x}$. We use the notation $\mathbf{x}_{-S}$ to denote all elements of tuple $\mathbf{x}$, except the elements in S. For simplicity, when $|S| = \{i\}$, we use $\mathbf{x}_{-i}$ instead. We use $\text{med}(\mathbf{x})$ to denote the median of the tuple $\mathbf{x}$.

Finally, we state mechanism names with Small Capitals.

2.2 Impartial Selection

In this section we give the basic notation and assumptions for the problem of impartial selection, discussed in detail in Chapters 3 and 4. More detailed definitions will be given in the respective chapters when needed.

Let N be a set of individuals (or agents). We usually use n to denote the number of strategic individuals. When convenient, we use N_S as an abbreviation of $N \setminus S$, for any $S \subset N$ and write $N_{i,\ldots,j}$ instead of $N_{\{i,\ldots,j\}}$ for simplicity.

A *nomination profile* $G = (N, E)$ is a directed graph without self-loops and parallel edges that has the agents of N as nodes. Each directed edge $(i, j) \in E$ represents a nomination from agent i to agent j. We use $\mathcal{G}$ to denote the set of all nomination profiles over the agents of N. We refer to this family of profiles as the *multiple-nomination* model. Occasionally, we refer to the outgoing edges as *nominations* or *votes*, and we use the term *nomination graph* to refer to a nomination profile. We define as $x_i = \{(i, j) \in E : j \in N\}$ the set of outgoing edges from node $i \in N$ and use the tuple $\mathbf{x} = (x_1, ..., x_{|N|})$ as an alternative representation for G. We use $\mathbf{x}_{-i}$ to denote the graph $(N, E \setminus ((\{i\} \times N))$.

A special case of nomination profiles consists of all the graphs with an out-degree of exactly 1. We denote this family with $\mathcal{G}^1$. Since the set of outgoing edges from node $i \in N$ is a singleton, we use the notation $x_i = j$ for some $j \in N \setminus \{i\}$ to denote the single note voted by node i.

We use $d_j(S, \mathbf{x})$ to denote the in-degree of node $j \in N$, taking into account only the incoming edges from nodes of set S, given the profile $\mathbf{x}$, i.e.,

$$d_j(S, \mathbf{x}) = |\{i \in S : (i,j) \in E\}| . \tag{2.1}$$

We use the simplified notations $d_j(\mathbf{x})$ when $S \supseteq N \setminus \{j\}$ and $d_j(S)$ when the nomination profile is clearly identified by the context. $\Delta(\mathbf{x})$ denotes the maximum in-degree of the profile $\mathbf{x}$, i.e., $\Delta(\mathbf{x}) = \max_{j \in N} d_j(\mathbf{x})$.

A *selection mechanism* for the family of profiles $\mathcal{G}' \subset \mathcal{G}$ maps each nomination profile $\mathbf{x} \in \mathcal{G}'$ to a probability distribution over the agents in N and, possibly over a special agent (denoted as agent 0) which represents the case where no one wins.

Definition 1 (selection mechanism)**.** For a family of graphs $\mathcal{G}' \subseteq \mathcal{G}$, a *selection mechanism* f for $\mathcal{G}'$ is a function $\mathcal{G}' \to [0,1]^{N \cup \{0\}}$, such that $\sum_{i \in N \cup \{0\}} (f(\mathbf{x}))_i = 1$.

In the following we write $f_i(\mathbf{x})$ instead $(f(\mathbf{x}))_i$ to denote the probability that node i is selected by mechanism f under the nomination profile $\mathbf{x}$, fot the sake of simplicity.

Under Definition 1, we refer to the selection mechanisms as *randomized*, in contrast to *deterministic* selection mechanism, which refers to the special case where $(f(\mathbf{x}))_i \in \{0,1\}$ for all $i \in N \cup \{0\}$. Due to the prevalence of deterministic mechanisms in the following chapters, we use also the following simpler definition:

Definition 2 (deterministic selection mechanism)**.** For a family of graphs $\mathcal{G}' \subseteq \mathcal{G}$, a *deterministic selection mechanism* is a function $f : \mathcal{G}' \to N \cup \{0\}$.

In other words, a deterministic selection mechanism is simply a function which maps each nomination profile to a single node (the *winner*) or does not return a winner at all.

A selection mechanism is *impartial* on the set of profiles $\mathcal{G}'$, when the winning probability of each node is not dependent on its outgoing edges. The following definition states this formally.

Definition 3 (impartial mechanism)**.** A selection mechanism over the set of graphs $\mathcal{G}' \subseteq \mathcal{G}$ is *impartial* when for any node $i \in N$, any nomination graph $\mathbf{x} \in \mathcal{G}'$ and any set x'_i of outgoing edges from node i, it is

$$f_i(\mathbf{x}) = f_i(x'_i, \mathbf{x}_{-i}).$$

For simplicity, we also present the alternative definition for impartial mechanisms:

Definition 4 (deterministic impartial mechanism)**.** A deterministic selection mechanism f is *impartial* when for any node $i \in N$, any nomination graph $\mathbf{x} \in \mathcal{G}'$ and any set x'_i of outgoing edges from node i, it is

$$f(\mathbf{x}) = i \iff f(x'_i, \mathbf{x}_{-i}) = i.$$

We formally present the *approximation ratio*, defined by [2] and for which we have discussed in Section 1.2.2.

Let $\mathbb{E}\left[\, d_{f(\mathbf{x})} \,\right](\mathbf{x})$ denote the expected in-degree of the selected node, when mechanism f is used for the selection, i.e. $\mathbb{E}\left[\, d_{f(\mathbf{x})}(\mathbf{x}) \,\right] = \sum_{u \in N} f_u(\mathbf{x}) \cdot d_u(\mathbf{x})$.

Definition 5. A selection mechanism f, provides an approximation ratio of α on a family of instances $\mathcal{G}' \subseteq \mathcal{G}$ if, for every nomination graph in $\mathcal{G}'$, the expected in-degree of the selected node by f, is at most α times smaller than the maximum in-degree, i.e., when

$$\sup_{\mathbf{x} \in \mathcal{G}} \frac{\Delta(\mathbf{x})}{\mathbb{E}\left[\, d_{f(\mathbf{x})}(\mathbf{x}) \,\right]} \leq \alpha. \tag{2.2}$$

2.2.1 Formal Descriptions of Known Mechanisms

In this section, we formally present known impartial mechanisms, which we have informally presented in Section 1.2.1.

Deterministic Mechanisms

We begin with the Majority With Default mechanism. We recall that this mechanism selects arbitrarily a node as the default winner. If any node receives an absolute majority of the votes, it is the winner. Otherwise, the default node wins. Algorithm 1 presents this mechanism in pseudo-code.

We continue with Plurality with Default, which awards any node that receives a unique maximum in-degree. Otherwise, the winner is a pre-selected default node. The unique maximum in-degree is determined as a node who beats pairwise all other nodes, ignoring the outgoing edges between the nodes under consideration and the outgoing edges of the default node. We examine a similar mechanism for a special sub-class of the multiple-nomination model in Chapter 4. There we use the more appropriate name Approval Voting with Default. Algorithm 2 presents this mechanism in pseudo-code.

Algorithm 1: Majority with Default

Input: Nomination Graph $\mathbf{x} = (N, E)$
: default node $t \in N$
Output: Node w in N

foreach $v \in N \setminus \{t\}$ **do**
  if $d_v(N \setminus \{t\}, \mathbf{x}) \geq \lceil n/2 \rceil$ **then**
    return v;
return t;

Algorithm 2: Plurality with Default

Input: Nomination Graph $\mathbf{x} = (N, E)$
: default node $t \in N$
Output: Node w in N

$W \leftarrow N$; ▷ A set of eligible winners
foreach $v \in N \setminus \{t\}$ **do**
  foreach $u \in N \setminus \{v, t\}$ **do**
    if $d_v(N \setminus \{u, t\}, \mathbf{x}) \leq d_u(N \setminus \{v, t\}, \mathbf{x})$ **then**
      $W \leftarrow W \setminus \{v\}$; ▷ v is no longer eligible to win
if $W \neq \emptyset$ **then**
  return $w \in W$; ▷ if W is not empty includes a single node
else
  return t

Before continuing, we note that in the single-nomination model, any mechanism defined with a default winner can be defined as a mechanism with a default-maker, i.e. a node which selects the winner. This way we can define both the Majority with Default-Maker and Plurality with default-Maker mechanisms. We omit to present the uninformative pseudo-code in these cases.

We next present the Partition Methods mechanisms. This is a family of mechanisms, parameterized by a set of $K \geq 2$ districts which partition the set of nodes N. Each partition performs a local election, and all nodes which receive an absolute majority of nominations in their district is named a *local winner*. All local winners qualify for a second round, where all nodes which are not local winners act as voters and the local winners as candidates. The local winner with the maximum in-degree for the remaining nodes is the winner, with ties broken arbitrarily. If no local winners exist, then the default node (by convention located in the first district) is the winner. Algorithm 3 describes the process in detail.

We note that this mechanism is defined for the single-nomination model. With a small change, it can be made suitable for the multiple-nomination model: The local winner can be selected with pairwise comparisons, similar to the Plurality with Default mechanism.

The Cross-Partition methods, also proposed in [57] is another family of impartial mechanisms for the single-nomination model. This mechanism is much more involved compared to the previous. This mechanism partitions the nodes in K districts of nearly equal size and uses another district for the default winner. The K districts are partitioned further into two orthogonal components: District $k \in [K]$ is partitioned to the sets $(\Pi_k, \bar{\Pi}_k)$ and the sets $(\Lambda_k, \bar{\Lambda}_k)$. These sets have a priority on determining a set of eligible winners called the *outer heroes*. An outer hero in district k is one which is nominated by all nodes in Π_l or $\bar{\Pi}_l$ for all l other than $k-1$ and k and from Λ_{k-1} or $\bar{\Lambda}_{k-1}$ (where subscriptions are taken modulo k). If no outer heroes exist, then in each district a local election happens; If any node receives all the votes from nodes in its district, it's called an *inner hero*. If eligible winners exist, they qualify for the second phase, where the remaining voters determine the winner, with ties- broken arbitrarily.

Impartiality of this mechanism is guaranteed when we notice that there can be at most only two outer heroes, which should be located in the same district. Suppose that an outer hero exists in district S_k. This node is voted by all nodes in either Λ_{k-1} or $\bar{\Lambda}_{k-1}$. Suppose that this is set Λ_{k-1}. The orthogonality of the partition implies that $\Lambda_{k-1} \cap \Pi_{k-1} \neq \emptyset$ and $\Lambda_{k-1} \cap \bar{\Pi}_{k-1} \neq \emptyset$. Since the mechanism is defined for the single nomination model, any other outer hero can only be in district S_k. Also, with the same argument, we can see

Algorithm 3: Partition Methods

Input: Nomination Graph $\mathbf{x} = (N, E)$
: default node $t \in N$
Output: Node w in N
Parameters: number of districts K; list $[S_1, .., S_K]$ with districts.
Requirements: $n \geq 7$, $K \geq 2$, $|S_1| \geq 4$, $|S_j| \geq 3$ for $j \in \{2, .., K\}$.

```
W ← ∅;                                         ▷ The set of local winners
foreach v ∈ N \ {t} do
    if d_v(S_1 \ {t}, x) ≥ ⌈|S_1|/2⌉ then
        W ← W ∪ {v};                           ▷ Node v becomes a local winner
for k = 2 to K do
    foreach v ∈ N \ {t} do
        if d_v(S_k, x) ≥ ⌈|S_k|/2⌉ then
            W ← W ∪ {v};                       ▷ Node v becomes a local winner
if W = ∅ then
    return t ;
else if |W| = 1 then
    return w ∈ W;
else
    return argmax_{v∈W} d_v(N \ W);            ▷ Ties broken arbitrarily
```

that any inner-hero can only be located in district S_k. With this in mind, we can verify impartiality. An outer-hero does not influence whether is an outer-hero or not, and cannot prevent another voter to become such. Also, no voter which is an inner-hero can remove an outer-hero, since they are in the same district. finally, the default node does not influence the first step of the mechanism at all, which is the only step where this node can become a winner. A detailed description of the mechanism appears in Algorithm 5.

Randomized mechanisms

We discuss now in more detail the randomized impartial mechanism for the selection of a single node we have presented in Section 1.2.1. We start with the 2-Partition mechanism. This is the simple mechanism where the nodes are partitioned into two sets S and W uniformly at random. The winner is the node in set W with maximum in-degree from S. Algorithm 4 present this mechanism in detail. Impartiality can be verified easily. The nodes

Algorithm 4: 2-Partition

Input: Nomination Graph $\mathbf{x} = (N, E) \in \mathcal{G}$
: default node $t \in N$
Output: Node w in N

Assign each node $v \in N$ to one of the two sets S and W, uniformly at random ($\Pr[v \in S] = \Pr(v \in W) = 1/2$), such that $S \cap W = \emptyset$ and $S \cup W = N$. **if** $W = \emptyset$ **then**
  return a node uniformly at random from N;
return $\text{argmax}_{v \in W}(d_v(S))$; ▷ `Break ties arbitrarily`

cannot influence the partitioning process, and the ones partitioned nodes in S cannot win and nodes in W cannot influence the selection process. We recall that despite its simplicity, this mechanism yields a reasonable approximation ratio of 4.

A careful generalization of the 2-Partition mechanism lead to k-Partition mechanism. In this case, the nodes are partitioned, uniformly at random, in $k \geq 2$ sets. These sets are examined in a linear order, maintaining and updating a *provisional winner* node. When the j-th partition is examined, the provisional winner is updated when a node with maximum in-degree from the nodes in $\bigcup_{i=1}^{j-1} S_i$ *excluding the outgoing edges of the provisional winner*, is at least equal the in-degree of the current provisional winner. The final provisional winner is selected by the mechanism. Algorithm 6 present this process formally. The k-Partition mechanism is impartial. To win the mechanism, a node must become a provisional winner at some point. This is not dependent on its outgoing edges. Since the current provisional winner's outgoing edges are not counted on the updating process, it cannot maintain this status on its own. This mechanism cannot have an approximation ratio higher than $\frac{2k}{k-1}$.

The Permutation mechanism can be seen as an extreme version of the k-Partition mechanism, where the n nodes are partitioned in n sets. This mechanism also maintains a provisional winner, which can be updated in every step. The mechanism examines the nodes according to random permutation, and at each step compares the in-degree of the node under examination *from nodes that have already been examined, excluding the outgoing edges of the provisional winner* with the in-degree of the provisional winner. If the node under examination has at least equal in-degree to the in-degree of the provisional winner, this node becomes a new provisional winner. The final provisional winner is selected by the mechanism.

We can verify that this mechanism is impartial using similar arguments with those we

Algorithm 5: CROSS-PARTITION METHODS

Input: Nomination Graph $\mathbf{x} = (N, E) \in \mathcal{G}^1$
: default node $t \in N$
Output: Node w in N
Parameters: number of districts K; list $[S_1, .., S_K]$ with districts; list of sets $[\Pi_1, ..., \Pi_K]$ and $[\Lambda_1, ..., \Lambda_K]$; list of sets $[\bar{\Pi}_1, ..., \bar{\Pi}_K]$ and $[\bar{\Lambda}_1, ..., \bar{\Lambda}_K]$.
Requirements: $n \geq 13$, $K \geq 3$, $|S_k| - |S_j| \leq 1$ for any distinct pair $j, k \in [K]$; $\Pi_k \cap \bar{\Pi}_k = \emptyset$, $\Lambda_k \cap \bar{\Lambda}_k = \emptyset$ for $k \in [K]$; $\Pi_k \cap \Lambda_k \neq \emptyset$, $\bar{\Pi}_k \cap \Lambda_k \neq \emptyset$, $\Pi_k \cap \bar{\Lambda}_k \neq \emptyset$, $\bar{\Pi}_k \cap \bar{\Lambda}_k \neq \emptyset$ for $k \in [K]$.

▷ `Default node` t `does not belong to any district`
$O = N \setminus \{t\}$; ▷ `Initial set of outer heroes`
$I = \emptyset$; ▷ `Set of inner heroes`
▷ `Search for outer heroes`
for $k = 1$ **to** K **do**
    $p \leftarrow k \mod K$; ▷ `Previous index modulo` K
    foreach $v \in S_k$ **do**
        for $j \in [K] \setminus \{p, k\}$ **do**
            if $d_v(\Pi_k) < |\Pi_k|$ **and** $d_v(\bar{\Pi}_k) < |\bar{\Pi}_k|$ **then**
                $O \leftarrow O \setminus \{v\}$; ▷ v `is not eligible for outer hero`
        if $d_v(\Lambda_k) < |\Lambda_k|$ **and** $d_v(\bar{\Lambda}_k) < |\bar{\Lambda}_k|$ **then**
            $O \leftarrow O \setminus \{v\}$;
▷ `Search for inner heroes`
for $k = 1$ **to** K **do**
    foreach $v \in S_k$ **do**
        if $d_v(S_k \setminus \{v\}) = |S_k \setminus \{v\}|$ **then**
            $I \leftarrow I \cup \{v\}$; ▷ v `is an inner winner`
if $O \neq \emptyset$ **then**
    $W \leftarrow O$; ▷ `Outer heroes are the only eligible winners`
else if $|I| \neq \emptyset$ **then**
    $W \leftarrow I$; ▷ `Inner heroes are the only eligible winners`
else
    $W \leftarrow \{t\}$; ▷ `Default nodes is the winner`
return $\mathrm{argmax}_{v \in W}\, d_v(N \setminus W)$; ▷ `Break ties arbitrarily`

Algorithm 6: k-Partition

Input: Nomination Graph $\mathbf{x} = (N, E) \in \mathcal{G}$
: default node $t \in N$
Output: Node w in N

Assign each node $v \in N$ to one of the two sets $S_1, .., S_k$, uniformly at random (i.r. $\Pr[v \in S_j] = 1/k$ for all $j \in [k]$), such that $\bigcap_{j=1}^{k} S_j = \emptyset$ and $S_j \cap S_l = \emptyset$ for all $j, l \in [k]$ with $j \neq l$. $\{u^*\} \leftarrow \emptyset$ $d^* \leftarrow 0$; ▷ `Initialize provisional winner's in-degree`
$X \leftarrow \emptyset$;
for $j \leftarrow 1$ **to** k **do**
  if $\max_{v \in S_j} d_v(X \setminus \{u^*\}) \geq d^*$ **then**
    $u^* \leftarrow u' \in \text{argmax}_{v \in S_j} d_v(X)$; ▷ `Break ties arbitrarily`
    $d^* \leftarrow d_v(X \setminus \{u^*\})$;
    $X \leftarrow X \cup S_j$;

have used for k-Partition mechanism: Any node should become a provisional winner to be selected by the mechanism, but none can force itself as a provisional winner. Any node that is a provisional winner cannot affect the mechanism anymore.

We finish this presentation with the Slicing mechanism. This mechanism follows ideas from the previous mechanism and yields a near-optimal approximation ratio *given that the maximum in-degree is large enough.* More precisely, for each ϵ there exists some constant Δ_ϵ, for which, in any nomination graph $\mathbf{x} \in \mathcal{G}$ with in-degree $\Delta(\mathbf{x})$ at least equal to Δ_ϵ, the Slicing mechanisms approximation ratio is at most $1 + \epsilon$.

The Slicing mechanism starts with the *sampling* phase, where it selects a sample on nodes X. A node belongs in X with probability $\epsilon \in [0, 1]$. Then, using the nodes in the sample, the mechanism calculates *estimated in-degrees*, as the number of nominations from the sample times $1/\epsilon$, the expected size of the sample. In the *slicing phase* which follows, the mechanism partitions the unsampled vertices into *slices*, which are sets of nearly equal estimated in-degree. The final, *election* phase works similar to *Permutation* and k-Partition mechanisms, and selects the eventual winner. Impartiality follows with the same arguments.

Algorithm 7: PERMUTATION

Input: Nomination Graph $\mathbf{x} = (N, E) \in \mathcal{G}$
: default node $t \in N$
Output: Node w in N
Choose a permutation $\mathbf{p} = (p_1, ..., p_n)$ of N uniformly at random;
$u^* \leftarrow p_1$; ▷ `Set first node as the provisional winner`
$d^* \leftarrow 0$; ▷ `Initialize provisional winner's in-degree`
$S = \{p_1\}$; **for** $j = 2$ **to** n **do**
  $S \leftarrow S \cup p_j$;
  if $d_{p_j}(S \setminus \{u^*\}, \mathbf{x}) \geq d^*$ **then**
    $u^* \leftarrow p_j$; ▷ `provisional winner is updated`
    $d^* \leftarrow d_{p_j}(S, \mathbf{x})$;
return u^*;

2.2.2 Formal Descriptions of Axioms

In this section, we present formally the axioms defined by Holzman and Moulin [57], which we have already discussed in Section 1.2.2. We recall that all the axioms are defined for the single nomination model. All axioms refer to a deterministic selection mechanism f, unless we state otherwise.

Axiom 1 (No Dummy)**.** for all $i \in N$, there exist $x_i, x'_i \in N \setminus \{i\}$ and $\mathbf{x} \in \mathcal{G}^1$ such that:

$$f(x_i, \mathbf{x}_{-i}) = f(x'_i, \mathbf{x}_{-i}). \tag{2.3}$$

No Dummy implies that any node can influence the outcome at least in one nomination profile.

Axiom 2 (No Exclusion)**.** for any $i \in N$, there exists a nomination graph $\mathbf{x} \in \mathcal{G}^1$ such that $f(\mathbf{x}) = i$.

No Exclusion implies that any node is a winner at least in one nomination profile.

Axiom 3 (Monotonicity)**.** for any $\mathbf{x} \in \mathcal{G}^1$ and any $i, j \in N$, x_j, $x'_j \in N \setminus \{j\}$

$$f(\mathbf{x}) = i \text{ and } x'_j = i \implies f(x'_j, \mathbf{x}_{-j}) = i. \tag{2.4}$$

Monotonicity implies that a node does not turn from a winner to a loser by receiving a nomination.

Axiom 4 (Negative Unanimity)**.** for any $\mathbf{x} \in \mathcal{G}^1$ and any $i \in N$,

$$d_i(\mathbf{x}) = 0 \implies f(\mathbf{x}) \neq i \tag{2.5}$$

Negative Unanimity implies that no node with 0 in-degree can be the winner.

Axiom 5 (Positive Unanimity)**.** for any $\mathbf{x} \in \mathcal{G}^1$ any $i \in N$,

$$d_i(\mathbf{x}) = n - 1 \implies f(\mathbf{x}) = i \tag{2.6}$$

Positive Unanimity implies that any node with in-degree $n - 1$ is the winner.

We give now some more information on the two decision power axioms we have presented before, Full Influence and Full Pivots. To show this, we need to define the *pivotal* property. A node i pivotal for the pair of nodes $j, j' \in N \setminus \{j, j'\}$, if for some profile $x \in \mathcal{G}^1$ there exists x_i, $x_i' \in N \setminus \{i\}$ such that $f(x_i, \mathbf{x}_{-i}) = j$ and $f(x_i, \mathbf{x}_{-i}) = j'$, for some deterministic selection mechanism f. In other words, the pivotal node i can change the winner from j to j'.

Axiom 6 (Full Pivots)**.** For all distinct $i, j, j' \in N$, i is pivotal for i, j.

The full Pivots property demands that every node can be decisive at some point. In a weaker notion of Full Pivots, one can say that i *influences* j, if there exists some j' such that i is pivotal for j, j'.

Axiom 7 (Full Influence)**.** For all distinct $i, j, j' \in N$, i is influences j.

Finally, we will present two axioms regarding symmetries. Let $\Pi(N)$ be the set of permutations of N. Let $\mathbf{x} \in \mathcal{G}^1$ be any nomination profile. Then, for any $\pi \in \Pi(N)$, $\mathbf{x}^\pi$ is the profile where whenever i nominates j in $\mathbf{x}$, $\pi(i)$ nominates $\pi(j)$ in $\mathbf{x}^\pi$. This is simply a renaming of the nodes of $\mathbf{x}$ using π.

Axiom 8 (Symmetry)**.** For all $\pi \in \Pi(N)$, $i \in N$ and $\mathbf{x} \in \mathcal{G}^1$ and a selection mechanism f:

$$f_{\pi(i)}(\mathbf{x}^\pi) = f_i(\mathbf{x}) \tag{2.7}$$

Symmetry can be used for the analysis of randomized mechanisms without loss of generality, as stated by Holzman and Moulin [57].

Related with Symmetry is the axiom of Anonymous Ballots, which treat nodes symmetrically as voters. We first need to define a *scoring profile* as $s(\mathbf{x}) = (d_i(\mathbf{x}))_{i \in N}$, i.e. a tuple indicating the in-degrees of each node. Then, the anonymous ballots axiom demands that two nomination graphs with the same scoring profiles to admit the same outcome.

Axiom 9 (Anonymous Ballots)**.** For all $\mathbf{x}, \mathbf{x}' \in \mathcal{G}'$,

$$s(\mathbf{x}) = s(\mathbf{x}') \implies f(\mathbf{x}) = f(\mathbf{x}'). \tag{2.8}$$

Algorithm 8: Slicing

Input: Nomination Graph $\mathbf{x} = (N, E) \in \mathcal{G}$
: default node $t \in N$
: parameter $\epsilon \in [0, 1]$
Output: Node w in N

▷ `Sampling Phase`
Draw a random sample X, where each node is sampled with probability ϵ;
foreach $v \in X$ **do**
  $\hat{d}_v \leftarrow \frac{1}{\epsilon} \cdot d_v(X)$; ▷ `compute estimated in-degrees from` X

▷ `Slicing Phase`
Create $\tau = \lceil \epsilon^{-2} \rceil$ sets $[S_1, ..., S_\tau]$, initialized as empty;
$\hat{\Delta} \leftarrow \max_{v \in N \setminus X}(\hat{d}_v)$; ▷ `find maximum estimated in-degree`
foreach $v \in N \setminus X$ **do**
  for $i = 1$ **to** τ **do**
    if $(i-1) \cdot \epsilon^2 \cdot \hat{\Delta} \leq \hat{d}_v \leq i \cdot \epsilon^2 \cdot \hat{\Delta}$ **then**
      $S_i \leftarrow S_i \cup \{v\}$

▷ `Election phase`
$R \leftarrow X$; ▷ `Revealed nodes set`
$u^* \leftarrow \text{argmax}_{v \in N \setminus R}(d_v(R))$; ▷ `set the first provisional winner`
for $i = 1$ **to** τ **do**
  foreach $v \in S_i \setminus \{u^*\}$ **do**
    $R \leftarrow R \cup \{v\}$ with probability $(1 - \epsilon)$;
  ▷ `Select a temporary provisional winner`
  $u' \leftarrow \text{argmax}_{v \in N \setminus R}(d_v(R))$; ▷ `Break ties arbitrarily`
  $R \leftarrow (R \cup \bigcup_{j \leq i} S_j) \setminus \{u'\}$;
  ▷ `Select a new provisional winner`
  $u \leftarrow \text{argmax}_{v \in N \setminus R}(d_v(R))$; ▷ `Break ties arbitrarily`
  $R \leftarrow (R \cup \bigcup_{j \leq i} S_j) \setminus \{u^*\}$;
return u^*.

Chapter 3

Impartial Selection with Additive Approximation Guarantees

This chapter is based on joint work with George Christodoulou and Ioannis Caragiannis. Part of this chapter was published in [27]. The main problem is introduced in Section 1.2 and formal definitions are given in Section 2.2.

3.1 Our Results

In this chapter, we introduce our first results on the problem of impartial selection. Recall that in this problem, a set of strategic individuals should select the most popular of them as a *winner*, and the input can be modeled as a directed graph, where the nodes act as the individuals and the directed edges act as nominations. We call such graphs *nomination graphs*. A selection mechanism takes as an input a nomination graph and returns a single node as the winner. This process would be trivial in the absence of strategic behavior, and the most popular candidate would be selected. However, we are interested in a case where all nodes act as strategic individuals, and they will misreport their true preferences if this would benefit them. A way to overcome this strategic behavior is the use of an *impartial* mechanism. An impartial mechanism prevents the individual from being strategic, given the crucial assumption that the nodes would only manipulate the mechanism to become winners themselves; instead, it provides appropriate incentives to the nodes to report their preferences truthfully, although not without cost, in the quality of the solution.

The problem was introduced independently by Holzman and Moulin [57] and Alon et al.

[2], which analyzed the problem through different perspectives. The former consider an axiomatic approach (see Section 1.2.2 for details), while the latter quantified the efficiency loss as the worst-case ratio between the maximum in-degree in graph versus the in-degree of the winner, quantity known as the *approximation ratio* (see Section 1.2.2 for more information). This line of research concluded with the work of Fischer and Klimm [43], who proposed optimal impartial selection mechanism with respect to approximation ratio.

It was pointed out [2, 43] that the most challenging nomination profiles for both deterministic and randomized mechanisms are those with small in-degrees. In the case of deterministic mechanisms, the situation is quite extreme as all deterministic mechanisms can be easily seen to have an unbounded approximation ratio on inputs with a maximum in-degree of 1 for a single vertex and 0 for all others; see 1.2.2 for a concrete example. Further to that, Bousquet et al. [19] have shown that if the maximum in-degree is large enough, randomized mechanisms that return a near-optimal impartial winner do exist. As a result, the approximation ratio does not seem to be an appropriate measure to classify deterministic selection mechanisms.

3.1.1 Contribution

We deviate from previous work and instead propose to use *additive approximation* as a measure of the quality of impartial selection mechanisms. This measure is defined as the *difference* between the maximum in-degree and the in-degree of the winner returned by the mechanism in a given graph. Our goal is to design mechanisms with additive approximation as small as possible.

We distinguish between two models already considered in the literature. The *single-nomination* model considered first by Holzman and Moulin [57] consists of all graphs with all nodes having an out-degree of 1. The *multiple-nomination* model, considered first by Alon et al. [2], is more general and allows for arbitrary out-degrees (i.e. multiple nominations and abstentions).

As positive results, we present two randomized impartial mechanisms which have additive approximation guarantees of $\Theta(\sqrt{n})$ and $\mathcal{O}(n^{2/3} \ln^{1/3} n)$ for the single-nomination and multiple-nomination models, respectively. Notice that both these additive approximation guarantees are $o(n)$ functions of the number n of vertices. We remark that an $o(n)$-additive approximation guarantee can be translated to an $1-\epsilon$ multiplicative guarantee for graphs with sufficiently large maximum in-degree, similar to the results of [19]. Conversely, the multiplicative guarantees of [19] can be translated to an $O(n^{8/9})$-additive guarantee[1]. This analysis further demonstrates that additive guarantees allow for a more smooth classification of mechanisms that achieve good multiplicative approximation in the limit.

Our mechanisms first select a small sample of nodes and then select the winner among the vertices nominated by the sample vertices. These mechanisms are randomized variants of a class of mechanisms which we define and call *strong sample mechanisms*. Strong sample mechanisms are impartial mechanisms that select the winner among the vertices nominated by a sample set of vertices. In addition, they have the characteristic that the sample set does not change with changes in the nominations of the nodes belonging to it. For the single nomination model, we provide a characterization, and we show that all deterministic strong sample mechanisms should use a fixed sample set that does not depend on the nomination profile. This yields a $n-2$ lower bound on the additive approximation guarantee of any deterministic strong sample mechanism. For their randomized variants, where the sample set is selected randomly, we present an $\Omega(\sqrt{n})$ lower bound, which shows that our first randomized impartial mechanism is the best possible among all randomized variants of strong sample mechanisms. Also, for the most general, multiple-nomination model, we present a lower bound of 3 for all deterministic mechanisms. Finally, we conclude our work by providing additive approximation bounds for various known mechanisms.

3.1.2 Roadmap

Section 3.2 introduces formally the additive approximation guarantees. Section 3.3 is dedicated to our positive results. Section 3.3.1 introduces the Sample and Vote mechanism, a randomized mechanism tailored to the single-nomination model. Section 3.3.2 introduces the Sample and Poll mechanism, a randomized mechanism for the multiple nomination model. Section 3.4.1, introduces the class of Sample and Strong Sample mechanisms and illustrates a characterization for the latter. Section 3.4.2 show an additive approximation lower bound for any deterministic impartial mechanism. Finally, Section 3.5 presents additive approximation guaranties for various known mechanisms.

[1] The authors in [19] do not provide additive guarantees; hence we based our calculations on their provided bounds on the multiplicative guarantee $1-\epsilon$. It is important to note, however, that they claim that they have not optimized their parameters, so it is possible that a tighter analysis may further reduce this guarantee.

3.2 Preliminaries

In this chapter we follow the notation and definitions proposed in Section 2.2. We evaluate selection mechanisms using additive approximation bounds. We use $\mathbb{E}\left[\, d_{f(\mathbf{x})}(\mathbf{x}) \,\right]$ to denote the expected in-degree of the selected node from the nomination graph $\mathbf{x}$, when f is used for the selection, i.e. $\mathbb{E}\left[\, d_{f(\mathbf{x})}(\mathbf{x}) \,\right] = \sum_{u \in N} f_u(\mathbf{x}) \cdot d_u(\mathbf{x})$. We note here that the randomization is only due to the mechanism. Also, let $\mathcal{G}_n$ be the family of all directed graphs without self-loops and parallel edges, with *exactly* n nodes. Using these notions, we can define additive approximation as:

Definition 6. A selection mechanism f is called $\alpha(n)$-additive in a family of nomination profiles with n nodes $\mathcal{G}'_n \leq \mathcal{G}_n$ if

$$\max_{\mathbf{x} \in \mathcal{G}_n} \left\{ \Delta(\mathbf{x}) - \mathbb{E}\left[\, d_{f(\mathbf{x})}(\mathbf{x}) \,\right] \right\} \leq \alpha(n),$$

for every $n \in \mathbb{N}_{\geq 2}$.

3.3 Upper Bounds

In this section we provide randomized selection mechanisms for the two best studied models in the literature. First, in Section 3.3.1 we propose a mechanism for the single nomination model of Holzman and Moulin [57], where nomination profiles consist only of graphs with all vertices having an out-degree of 1. Then, in Section 3.3.2 we provide a mechanism for the more general model studied by Alon et al. [2], which allows for multiple nominations and abstentions.

3.3.1 The Sample and Vote Mechanism

Our first mechanism, which we call Sample and Vote, forms a sample S of nodes by repeating k times the selection of a node uniformly at random with replacement[2]. Any node that is selected at least once belongs to the sample S. Let $W := \{u \in N \setminus S : d_u(S, \mathbf{x}) \geq 1\}$ be the set of nodes outside S that are nominated by the nodes of S. If $W = \emptyset$, no winner is returned. Otherwise, the winner is a node in $\operatorname{argmax}_{u \in W}\ d_u(N \setminus W, \mathbf{x})$. We note here the

[2]Sampling uniformly at random with replacement allows for a simple analysis of the mechanism. In section 3.4.1, we show that this is indeed a good choice, as no other sampling method yields better additive approximation.

crucial fact that the selection of the sample set S is independent of the nomination profile $\mathbf{x}$. Algorithm 9 displays this mechanism in more detail.

Algorithm 9: SAMPLE AND VOTE

Input: Nomination Graph $\mathbf{x} = (N, E) \in \mathcal{G}$
: default node $t \in N$

Draw a random sample S, where each node is sampled uniformly with replacement;
$W \leftarrow 0$;
foreach $v \in N \setminus S$ **do**
    if $d_v(S) > 0$ **then**
        $W \leftarrow W \cup \{v\}$
return $\text{argmax}_{v \in W}(d_v(N \setminus W))$;   ▷ `Break ties arbitrarily`

Impartiality follows since a node that does not belong to W (no matter if it belongs to S or not) cannot become the winner and the nominations of nodes in W are not taken into account for deciding the winner among them. We now argue that, for a carefully selected k, this mechanism also achieves a good additive guarantee.

Theorem 3.1. *For $k = \Theta(\sqrt{n})$, the* SAMPLE AND VOTE *mechanism is impartial and $\Theta(\sqrt{n})$-additive in the single nomination model.*

Proof. Consider a nomination graph and let u^* be a node of maximum in-degree Δ. In our proof of the approximation guarantee, we will use the following two technical lemmas.

Lemma 3.2. *If $u^* \in W$, then the winner has in-degree at least $\Delta - k$.*

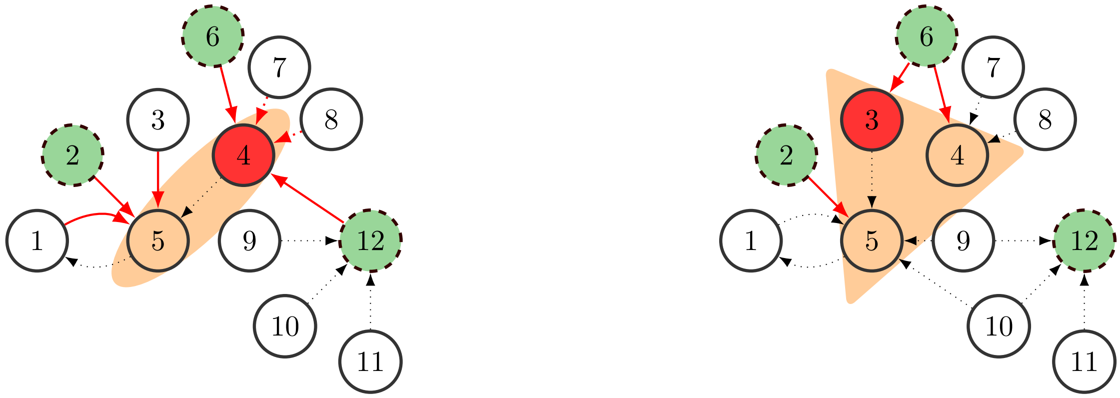
Proof. This is clearly true if the winner returned by SAMPLE AND VOTE is u^*. Otherwise, the winner w satisfies

$$d_w(N, \mathbf{x}) \geq d_w(N \setminus W, \mathbf{x}) \geq d_{u^*}(N \setminus W, \mathbf{x}) = d_{u^*}(N, \mathbf{x}) - d_{u^*}(W, \mathbf{x}) \geq \Delta - k.$$

The first inequality is trivial. The second inequality follows by the definition of the winner w. The third inequality follows since W is created by nominations of nodes in S, taking into account that each node has an out-degree exactly 1. Hence, $d_{u^*}(W, \mathbf{x}) \leq |W| \leq |S| \leq k$. □

Lemma 3.3. *The probability that u^* belongs to the nominated set W is*

$$\Pr[\, u^* \in W \,] = \left(1 - \left(1 - \frac{\Delta}{n-1}\right)^k\right)\left(1 - \frac{1}{n}\right)^k.$$



(a) Sample and Vote (b) Sample and Poll

Figure 3.1: Examples for Sample and Vote and Sample and Poll mechanisms, with sample size $k = 3$ and $n = 12$. In both cases we use the same sample set $S = \{2, 6, 12\}$. For Sample and Vote, the nodes in S define the set $W = \{4, 5\}$ of possible winners. The winner is then the node with maximum in-degree from the votes from $N \setminus W$ to W (the solid drawn edges in the figure). For Sample and Poll, the sample set S immediately declares the winner, as one of the maximum in-degree nodes from edges starting in S, while the edges from nodes in $N \setminus S$ are completely ignored. In both cases, the dark colored node is the winner and the light colored, dashed-lined nodes belong to the sample set S. Also, all edges drawn with a dotted line are ignored by the mechanism. The shaded areas indicate which nodes belong in set W.

Proof. Indeed, u^* belongs to W if it does not belong to the sample S and instead some of the Δ nodes that nominate u^* is picked in some of the k node selections. The probability that u^* is not in the sample is

$$\Pr[\, u^* \notin S \,] = \left(1 - \frac{1}{n}\right)^k, \tag{3.1}$$

i.e., the probability that node u^* is not picked in some of the k node selections. Observe that the probability that some of the Δ nodes that nominate u^* is picked in a node selection step, assuming that u^* is never selected, is $\frac{\Delta}{n-1}$. Hence, the probability that some of the Δ nodes nominating u^* is in the sample, assuming that $u^* \notin S$, is

$$\Pr[\, d_{u^*}(S, \mathbf{x}) \geq 1 | u^* \notin S \,] = 1 - \left(1 - \frac{\Delta}{n-1}\right)^k. \tag{3.2}$$

The lemma follows by the chain rule

$$\begin{aligned}\Pr[\,u^* \in W\,] &= \Pr[\,u^* \notin S \wedge d_{u^*}(S,\mathbf{x}) \geq 1\,]\\ &= \Pr[\,d_{u^*}(S,\mathbf{x}) \geq 1 | u^* \notin S\,] \cdot \Pr[\,u^* \notin S\,]\end{aligned}$$

and equations (3.1) and (3.2). □

By Lemmas 3.2 and 3.3, we have that the expected degree of the winner returned by mechanism Sample and Vote is

$$\begin{aligned}\mathbb{E}[\,d_w(N,\mathbf{x})\,] &\geq \Pr[\,u^* \in W\,] \cdot (\Delta - k)\\ &= \left(1 - \left(1 - \frac{\Delta}{n-1}\right)^k\right)\left(1 - \frac{1}{n}\right)^k (\Delta - k)\\ &\geq \left(1 - \left(1 - \frac{\Delta}{n-1}\right)^k\right)\left(1 - \frac{k}{n}\right)(\Delta - k)\\ &> \left(1 - \left(1 - \frac{\Delta}{n-1}\right)^k\right)(\Delta - 2k)\\ &= \Delta - 2k - \left(1 - \frac{\Delta}{n-1}\right)^k (\Delta - 2k)\end{aligned}$$

The second inequality follows by Bernoulli's inequality $(1+x)^r \geq 1 + rx$ for every real $x \geq -1$ and $r \geq 0$ and the third one since $n > \Delta$. Now, the quantity $\left(1 - \frac{\Delta}{n-1}\right)^k (\Delta - 2k)$ is maximized for $\Delta = \frac{n-1+2k^2}{k+1}$ to a value that is at most $\frac{n+1}{k+1} - 2$. Hence,

$$\mathbb{E}[\,d_w(N,\mathbf{x})\,] \geq \Delta - 2(k-1) - \frac{n+1}{k+1}.$$

By setting $k \in \Theta(\sqrt{n})$, we obtain that $\mathbb{E}[\,d_w(N,\mathbf{x})\,] \geq \Delta - \Theta(\sqrt{n})$, as desired. □

3.3.2 The Sample and Poll Mechanism

For the multiple nominations model, we propose the randomized mechanism Sample and Poll, which is even simpler than Sample and Vote. Sample and Poll forms a sample S of nodes by repeating k times the selection of a node uniformly at random with replacement. The winner (if any) is a node w in $\text{argmax}_{u \in N \setminus S}\, d_u(S,\mathbf{x})$. We remark that, for technical reasons, we allow S to be a multi-set if the same node is selected more than once. Then, edge multiplicities are counted in $d_u(S, \mathbf{x})$. Clearly, SAMPLE AND POLL is impartial. The winner is decided by the nodes in S, which in turn have no chance to become winners. Our approximation guarantee is slightly weaker now.

Algorithm 10: SAMPLE AND POLL

```
Input: Nomination Graph x = (N, E) ∈ 𝒢
     : default node t ∈ N
Draw a random sample S, where each node is sampled uniformly with replacement;
W ← 0;
foreach v ∈ N \ S do
    if d_v(S) > 0 then
        W ← W ∪ {v}
return argmax_{v∈W}(d_v(S));                    ▷ Break ties arbitrarily
```

Theorem 3.4. *For* $k = \left\lceil 4^{1/3} n^{2/3} \ln^{1/3} n \right\rceil$*, the* SAMPLE AND POLL *mechanism is impartial and* $\Theta(n^{2/3} \ln^{1/3} n)$*-additive.*

Proof. Let u^* be a node of maximum in-degree Δ. If $\Delta \leq k$, SAMPLE AND POLL is clearly $\Theta(n^{2/3} \ln^{1/3} n)$-additive. So, in the following, we assume that $\Delta > k$. Let C be the set of nodes of in-degree at most $\Delta - k - 1$. We first show that the probability $\Pr[\delta(w, \mathbf{x}) \leq \Delta - k - 1]$ that some node of C is returned as the winner by SAMPLE AND POLL is small.

Notice that if one of the nodes of C is the winner, then either node u^* belongs to the sample set S or it does not belong to S but it gets the same or fewer nominations compared to some node u of C. Hence,

$$\begin{aligned}
&\Pr[d_w(N, \mathbf{x}) \leq \Delta - k - 1] \\
&\leq \Pr[u^* \in S] + \Pr[u^* \notin S \wedge d_{u^*}(S, \mathbf{x}) \leq d_u(S, \mathbf{x}) \text{ for some } u \in C \text{ s.t. } u \notin S] \\
&\leq \Pr[u^* \in S] + \sum_{u \in C} \Pr[u^* \notin S \wedge u \notin S \wedge d_{u^*}(S, \mathbf{x}) \leq d_u(S, \mathbf{x})] \\
&= \Pr[u^* \in S] + \sum_{u \in C} \Pr[u^*, u \notin S] \Pr[d_{u^*}(S, \mathbf{x}) \leq d_u(S, \mathbf{x}) | u^*, u \notin S] \qquad (3.3)
\end{aligned}$$

We will now bound the rightmost probability in (3.3).

Claim 3.4.1. *For every $u \in C$,* $\Pr\left[\, d_{u^*}(S,\mathbf{x}) \le d_u(S,\mathbf{x}) | u^* \notin S, u \notin S \,\right] \le \exp\left(-\frac{k^3}{2n^2}\right)$.

Proof. Assuming that both u^* and u do not belong to the sample set S, we will express the difference $d_{u^*}(S,\mathbf{x}) - d_u(S,\mathbf{x})$ as the sum of independent random variables Y_i for $i = 1, ..., k$. Variable Y_i indicates the contribution of the i-th node selection to the difference $d_{u^*}(S,\mathbf{x}) - d_u(S,\mathbf{x})$. In particular, Y_i is equal to 1, -1, and 0 if the outgoing edges of the i-th node selected in the sample set points to node u^* but not to node u, to node u but not to node u^*, and either to none or to both of them, respectively. Hence, $d_{u^*}(S,\mathbf{x}) - d_u(S,\mathbf{x}) = \sum_{i=1}^{k} Y_i$ with $Y_i \in \{-1, 0, 1\}$ and

$$\begin{aligned}\mathbb{E}\left[\, d_{u^*}(S,\mathbf{x}) - d_u(S,\mathbf{x}) | u^*, u \notin S \,\right] &= (\Delta - d_{u^*}(\{u\},\mathbf{x}) - d_u(N,\mathbf{x}) + d_u(\{u^*\},\mathbf{x}))\frac{k}{n-2}\\ &\ge \frac{k^2}{n}.\end{aligned}$$

Notice that for the computation of the expectation, we have used the facts that $\Delta - d_{u^*}(\{u\},\mathbf{x})$ nodes besides u have outgoing edges pointing to u^*, $d_u(N,\mathbf{x}) - d_u(\{u^*\},\mathbf{x})$ nodes besides u^* have outgoing edges pointing to u, and each of them is included in the sample set with probability $\frac{k}{n-2}$. The inequality follows since $d_u(N,\mathbf{x}) \le \Delta - k - 1$ and $d_{u^*}(\{u\},\mathbf{x}), d_u(\{u^*\},\mathbf{x}) \in \{0,1\}$.
We will now apply Hoeffding's bound, which is stated as follows.

Lemma 3.5 (Hoeffding [56])**.** *Let $X_1, X_2, ..., X_t$ be independent random variables so that* $\Pr\left[\, a_j \le X_j \le b_j \,\right] = 1$. *Then, the expectation of the random variable $X = \sum_{j=1}^{t} X_j$ is* $\mathbb{E}[X] = \sum_{j=1}^{t} \mathbb{E}[X_j]$ *and, furthermore, for every $\nu \ge 0$,*

$$\Pr\left[\, X \le \mathbb{E}[X] - \nu \,\right] \le \exp\left(-\frac{2\nu^2}{\sum_{j=1}^{t}(b_j - a_j)^2}\right).$$

In particular, we apply Lemma 3.5 on the random variable $X = d_{u^*}(S,\mathbf{x}) - d_u(S,\mathbf{x})$ (assuming that $u^*, u \notin S$). Note that $t = k$, $a_j = -1$ and $b_j = 1$, and recall that $\mathbb{E}\left[\, X \,\right] \ge \frac{k^2}{n}$. We obtain

$$\begin{aligned}\Pr\left[\, d_{u^*}(S,\mathbf{x}) - d_u(S,\mathbf{x}) \le 0 | u^*, u \notin S \,\right] &= \Pr\left[\, X \le 0 \,\right]\\ &\le \left(-\frac{\mathbb{E}\left[\, X \,\right]^2}{2k}\right) \le \exp\left(-\frac{k^3}{2n^2}\right),\end{aligned}$$

as desired. □

Using the definition of $\mathbb{E}[\,\delta(w,\mathbf{x})\,]$, inequality (3.3), and Claim 3.4.1, we obtain

$$\begin{aligned}
\mathbb{E}[\,d_w(N,\mathbf{x})\,] &\geq (\Delta-k)\cdot(1-\Pr[\,d_w(N,\mathbf{x})\leq \Delta-k-1\,]) \\
&\geq (\Delta-k)\Pr[\,u^*\notin S\,] \\
&\quad -(\Delta-k)\left(\sum_{u\in C}\Pr[\,u^*,u\notin S\,]\cdot\Pr[\,d_{u^*}(S,\mathbf{x})\leq d_u(S,\mathbf{x})|u^*,u\notin S\,]\right) \\
&\geq (\Delta-k)\left(1-\frac{1}{n}\right)^k-(\Delta-k)\left(\sum_{u\in C}\left(1-\frac{2}{n}\right)^k\cdot\exp\left(-\frac{k^3}{2n^2}\right)\right) \\
&\geq (\Delta-k)\left(1-\frac{k}{n}\right)-(\Delta-k)\cdot n\cdot\exp\left(-\frac{k^3}{2n^2}\right) \\
&\geq \Delta-2k-n^2\cdot\exp\left(-\frac{k^3}{2n^2}\right).
\end{aligned} \tag{3.4}$$

The last inequality follows since $n\geq\Delta$. By setting $k=\left\lceil 4^{1/3}n^{2/3}\ln^{1/3}n\right\rceil$, inequality (3.4) yields $\mathbb{E}[\,d_w(N,\mathbf{x})\,]\geq\Delta-\Theta\left(n^{2/3}\ln^{1/3}n\right)$, as desired.

□

3.4 Lower Bounds

In this section we complement our positive results by providing impossibility results. First, in Section 3.4.1, we provide lower bounds for a class of mechanisms which we call strong sample mechanisms, in the single-nomination model of Holzman and Moulin [57].

Then, in Section 3.4.2, we provide a lower bound for the most general, multiple-nomination model of Alon et al. [2], which applies to any deterministic mechanism.

3.4.1 Strong Sample Mechanisms

In this section, we give a characterization theorem for a class of impartial mechanisms satisfy a special property we call *strong sample*. We then use this characterization to provide lower bounds on the additive approximation of deterministic and randomized impartial mechanisms satisfying this property. Our results suggest that the Sample and Vote mechanism from Section 3.3.1 is essentially the best possible randomized mechanism in this class.

For a graph $G \in \mathcal{G}$ and a subset of nodes S, let $W := W_S(G)$ be the set of nodes outside S nominated by S, i.e. $W = \{w \in N \setminus S : (v, w) \in E, v \in S\}$. Then, a deterministic *sample* mechanism[3] (g, f) firstly selects a subset S using some *sample function* $g : \mathcal{G} \to 2^N$, and then applies a (possibly randomized) selection mechanism f by restricting its range on nodes in W; notice that if $W = \emptyset$, f does not select any node.

This definition allows for a large class of impartial mechanisms. For example, the special case of sample mechanisms with $|S| = 1$ (in which, the winner has in-degree at least 1), coincides with all negative unanimous mechanisms defined by Holzman and Moulin [57]. Indeed, when $|S| = 1$, the set W in never empty and the winner has in-degree at least 1. This is not however the case for $|S| > 1$, where W could be empty when all nodes in S have outgoing edges destined for nodes in S and no winner can be declared. Characterizing all sample mechanisms is an outstanding open problem. We are able to provide a first step, by providing a characterization for the more restricted class of strong sample mechanisms. Informally, in such mechanisms, nodes cannot affect their chance of being selected in the sample set S.

Definition 7. (Deterministic strong sample mechanisms) We call a deterministic sample mechanism (g, f) with sample function $g : \mathcal{G} \to 2^N$ *strong* if $g(x'_u, \mathbf{x}_{-u}) = g(\mathbf{x})$ for all $u \in g(\mathbf{x})$, $x'_u \in N \setminus \{u\}$ and $\mathbf{x} \in \mathcal{G}$.

The reader may observe the similarity of this definition with impartiality (function g of a strong sample mechanism satisfies similar properties with function f of an impartial selection mechanism). The following lemma describes a straightforward, yet useful, consequence of the above definition.

Lemma 3.6. *Let (g, f) be a deterministic strong sample mechanism and let $S \subseteq N$. For any nomination profiles $\mathbf{x}, \mathbf{x}'$ with $\mathbf{x}'_{-S} = \mathbf{x}_{-S}$, if $S \setminus g(\mathbf{x}) \neq \emptyset$ then $S \setminus g(\mathbf{x}') \neq \emptyset$.*

Proof. For the sake of contradiction, let us assume that $S \setminus g(\mathbf{x}') = \emptyset$, i.e., the sample nodes in $\mathbf{x}'$ are disjoint from S. Then, by Definition 7, $g(\mathbf{x})$ remains the same as outgoing edges from nodes in S should not affect the sample set. But then, $S \setminus g(\mathbf{x}) = \emptyset$, which is a contradiction. □

[3] For simplicity we use the notation (g, f) rather the more precise $(g, f(g))$.

In the next theorem, we provide a characterization for the sample function of deterministic impartial strong sample mechanisms in the single nomination model. The theorem essentially states that the only possible way to choose the sample set must be independent of the graph.

Theorem 3.7. *In the single nomination model, any impartial deterministic strong sample mechanism* (g, f) *selects the sample set independently of the nomination profile, i.e., for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{G}^1$, $g(\mathbf{x}) = g(\mathbf{x}') = S$.

Proof. Since we are in the single nomination model, without loss of generality, we can use the simplified notation $x_u = v$ instead of $x_u = \{(u, v)\}$ for any graph $\mathbf{x} \in \mathcal{G}^1$. Consider any sample mechanism (g, f) and any nomination profile $\mathbf{x} \in \mathcal{G}^1$. It suffices to show that for any node u, and any alternative vote x'_u, the sample set must remain the same, i.e., $g(x'_u, \mathbf{x}_{-u},) = g(\mathbf{x})$. If $u \in g(\mathbf{x})$, this immediately follows by Definition 7. In the following, we prove two claims showing that this holds also when $u \notin g(\mathbf{x})$; Claim 3.7.1 treats the case where u is a winner of a profile, while Claim 3.7.2 treats the case where u is a not a winner.

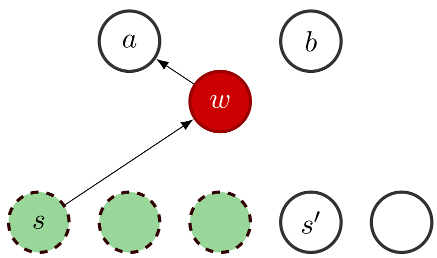
Claim 3.7.1. *Let* (g, f) *be an impartial deterministic strong sample mechanism and let* $\mathbf{x}$ *be any nomination profile in* $\mathcal{G}^1$*. Then the sample set must remain the same for any other vote of the winner, i.e.,* $g(\mathbf{x}) = g(x'_{f(\mathbf{x})}, \mathbf{x}_{-f(\mathbf{x})})$ *for any* $x'_{f(\mathbf{x})} \in N \setminus \{f(\mathbf{x})\}$.

Proof. Let $w = f(\mathbf{x})$ be the winner, for some nomination profile $\mathbf{x}$. We will prove the claim by induction on the in-degree of the winner, denoted as $d_w(\mathbf{x})$. Note that $d_w(\mathbf{x}) > 0$ for any $\mathbf{x} \in \mathcal{G}^1$, for any Sample mechanism.

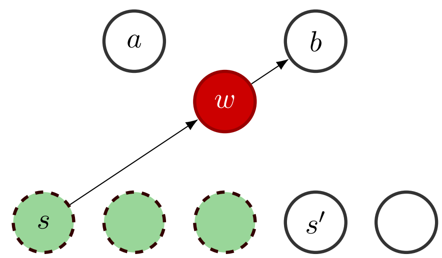
(Base case: $d_w(\mathbf{x}) = 1$**)** Let $S = g(\mathbf{x})$ be the sample set for profile $\mathbf{x}$. Assume for the sake of contradiction that when w changes its vote to x'_w, the sample for profile $\mathbf{x}' = (x'_w, \mathbf{x}_{-w})$ changes, i.e., $g(\mathbf{x}') = S' \neq S$. We first note that impartiality of f implies that $w = f(\mathbf{x}')$. Next, observe that the node voting for w in S must be also in S'; otherwise, w becomes a winner without getting any vote from the sample set, which contradicts our definition of sample mechanisms. We will show that this must be the case for all nodes in S.

To do this, we will expand two parallel branches, creating a sequence of nomination profiles starting from $\mathbf{x}$ and $\mathbf{x}'$ which will eventually lead to a contradiction. Figure 3.2 depicts the situation for $\mathbf{x}$ and $\mathbf{x}'$.

We start with the profile $\mathbf{x}'$. Consider a node $s' \in S' \setminus S$. We create a profile $\mathbf{z}'$ in which all nodes in $S' \setminus s'$ vote for s' (i.e., $z_v = s'$, for each $v \in S' \setminus s'$), node s' votes for w (i.e., $z_v = w$), while the rest of the nodes vote as in $\mathbf{x}'$ (i.e., $z_v = x_v$, for each $v \notin S'$). For illustration, see Figures 3.3a and 3.3b. By the definition of a strong sample mechanism,



(a) profile $\mathbf{x}$



(b) profile $\mathbf{x}'$

Figure 3.2: The starting profiles $\mathbf{x}$ and $\mathbf{x}'$ in Claim 3.7.1. The dark node is the winner, while the light, dashed-lined nodes are the members of the sets S and S', respectively.

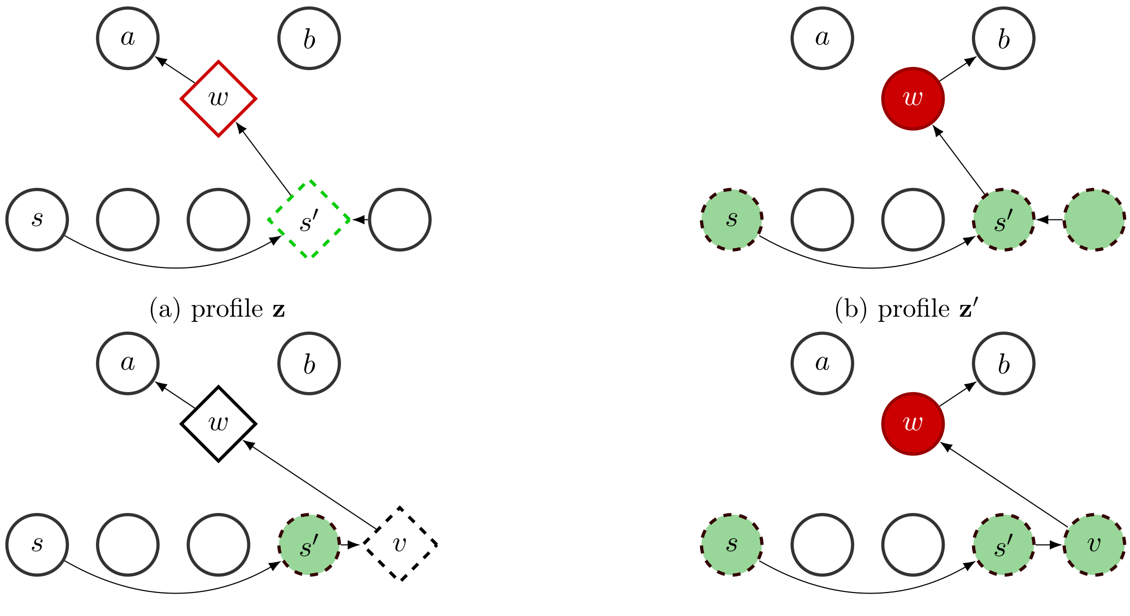
we obtain $g(\mathbf{z}') = g(\mathbf{x}')$, since only votes of nodes in S' have changed. Notice also that $f(\mathbf{z}') = w$, as this is the only node outside S' that receives votes from S'. We now move to profile $\mathbf{x}$ and apply the same sequence of vote changes, involving all the nodes in S'. These changes lead to the profile $\mathbf{z}$, which differs from $\mathbf{z}'$ only in the outgoing edge of node w.

By Lemma 3.6, there is a node $v \in S'$ such that $v \notin g(\mathbf{z})$. If $v = s'$, then we end up in a contradiction. This is because $f(\mathbf{z}) \neq w$, since s' is the only node voting for w in $\mathbf{z}'$ and s' is not in the sample, while $f(\mathbf{z}') = w$, as stated by the other branch and since, when w change its vote to x'_w, the created profile is $(x'_w, \mathbf{z}_{-w}) = \mathbf{z}'$ contradicting impartiality (see Figures 3.3a and 3.3b).

We are now left with the case where $s' \in g(\mathbf{z})$ and $v \neq s'$. Starting from $\mathbf{z}$ and $\mathbf{z}'$, we will create profiles $\mathbf{y}$ and $\mathbf{y}'$ (see Figures 3.3c and 3.3d) as follows: we construct $\mathbf{y}$ by letting s' vote towards v (i.e., $y_{s'} = v$), v vote towards w (i.e., $y_v = w$) and $y_i = z_i$ for all other nodes $i \neq v, s'$. By the strong sample property, when s' votes towards v the sample set is preserved, i.e., v cannot get in the sample. Also, when v votes, v cannot get in the sample (by a trivial application of Lemma 3.6); therefore, $v \notin g(\mathbf{y})$. Hence, w cannot be the winner as its only incoming vote is from v, a node that does not belong to the sample set $g(\mathbf{y})$.

Starting from $\mathbf{z}'$, we create similarly $\mathbf{y}'$ by letting s' vote towards v ($y'_{s'} = v$), v to vote towards w ($y'_v = w$) and $y'_i = z_i$ for all other nodes $i \neq v, s'$. In this case, S' will be preserved as sample set in profile $\mathbf{y}'$ (i.e. $g(\mathbf{y}') = S'$). Therefore, w is the only node voted by the sample set and must be the winner, leading to a contradiction (see Figures 3.3c and 3.3d).

(Induction step) Assume as induction hypothesis that, for all profiles $\mathbf{x} \in \mathcal{G}^1$, it holds $g(\mathbf{x}) = g(x'_w, \mathbf{x}_{-w})$ when $d_w(\mathbf{x}) \leq \lambda$, for some $\lambda \geq 1$. Now, consider any profile



(a) profile $\mathbf{z}$ (b) profile $\mathbf{z}'$ (c) profile $\mathbf{y}$ (d) profile $\mathbf{y}'$

Figure 3.3: Profiles $\mathbf{z}$ and $\mathbf{z}'$ in the base case of the proof of Claim 3.7.1: if $v = s'$ then $s' \notin g(\mathbf{z})$ and since this is the only node voting for w, w cannot win in $\mathbf{z}$, while it must be the winner in $\mathbf{z}'$ —a contradiction. Profiles $\mathbf{y}$ and $\mathbf{y}'$: if $s' \in g(\mathbf{y})$, we let s' vote for v and v for w, making w the winner in $\mathbf{y}'$ but not in $\mathbf{y}$. A dark circle denotes the winner, while light, dashed-lined circles denote the members of the sample sets S and S'. A solid-lined diamond denotes a node that cannot be the winner and a dashed-lined diamond denotes a node that cannot be in the sample set.

$\mathbf{x}$ where $f(\mathbf{x}) = w$ and $d_w(\mathbf{x}) = \lambda + 1$ and assume for the sake of contradiction that there is some graph $\mathbf{x}' = (x'_w, \mathbf{x}_{-w})$ where $g(\mathbf{x}') = S' \neq S$. Without loss of generality, let $d_w(S, \mathbf{x}) \leq d_w(S', \mathbf{x}')$.

Starting from $\mathbf{x}'$, we create profile $\mathbf{z}'$, by letting all nodes in S' vote for some $s' \in S'$ and s' vote for w, i.e., $z'_v = s'$ for each node $v \in S' \setminus \{s'\}$ and $z'_{s'} = w$. The strong sample property implies that $g(\mathbf{z}') = S'$ and $f(\mathbf{z}') = w$. We focus now on profile $\mathbf{x}$, and create the profile $\mathbf{z}$, by performing the same series of changes, i.e., by letting all nodes in $S' \setminus \{s'\}$ vote for s' and s' vote for w. Note here that $\mathbf{z}$ differs from $\mathbf{z}'$ only in the outgoing edge of w. Like before, Lemma 3.6 establishes that there will be some node $v \in S'$ such that $v \notin g(\mathbf{z})$, i.e., $g(\mathbf{z}) \neq S'$. Turning our attention back to $\mathbf{z}'$, we let w change its vote to x_w, creating profile $(x_w, \mathbf{z}'_{-w})$. Observe that $(x_w, \mathbf{z}'_{-w}) = \mathbf{z}$. When $d_w(\mathbf{z}') < d_w(\mathbf{x})$, by the induction hypothesis we have $g(\mathbf{z}) = g(\mathbf{z}') = S'$, a contradiction.

We are left with the case $d_w(\mathbf{z}') = d_w(\mathbf{x})$. We will use a series of careful steps to decrease the in-degree of w, without changing the sample set. This will allow us to use the induction hypothesis to finalize our proof.

Let L denote the set of nodes which vote for w in profile $\mathbf{x}$. We note here that, we may end up in the case $d_w(\mathbf{z}') = d_w(\mathbf{x})$ because only one node vote for w in $\mathbf{z}'$, i.e. $|g(\mathbf{z}') \cap L| = 1$, and due to the strong sample property, $|g(\mathbf{x}') \cap L| = 1$; otherwise we could decrease the in-degree of w without changing the sample set and directly use the induction hypothesis to prove the claim. Node s' is the single node in $g(\mathbf{z}') \cap L$. Note here that there exists at least one node in $g(\mathbf{x}) \cap L$. Say this is node s. We first show how we handle the case $d_s(\mathbf{z}') \leq \lambda - 1$. Then we provide a generalization of this idea to handle the case $d_s(\mathbf{z}') > \lambda - 1$.

When $d_s(\mathbf{z}') \leq \lambda - 1$, we can create the profile $\mathbf{y}'$, where s' votes for s (i.e. $y'_{s'} = s$ and $\mathbf{y}' = (y'_{s'}, \mathbf{z}'_{-s'})$), hence $f(\mathbf{y}') = s$ and $g(\mathbf{y}') = g(\mathbf{z}') = S'$ (recall that s' is the single node in the sample set $g(\mathbf{z}')$ voting outside of $g(\mathbf{z}')$, hence determining the winner). We can create now the profile $\mathbf{q}'$ where node s votes for node s' (i.e. $q'_s = s'$) and the other nodes vote like in $\mathbf{y}'$, i.e. $\mathbf{q}' = (q'_s, \mathbf{y}'_{-s})$. Since $d_s(\mathbf{y}') \leq \lambda$ and $f(\mathbf{y}') = s$, by changing the outgoing edge of the winning node s, the sample set does not change, due to the induction hypothesis, i.e. $g(\mathbf{q}') = g(\mathbf{y}') = S'$. Finally, we create the profile $\mathbf{r}'$, where s' votes for w (i.e. $r'_{s'} = w$) and the other nodes vote like in $\mathbf{q}'$ (i.e. $\mathbf{r}' = (r'_{s'}, \mathbf{q}'_{-s'})$). The strong sample property now implies that $g(\mathbf{r}') = g(\mathbf{q}') = S'$ and $f(\mathbf{r}') = w$. Since $d_w(\mathbf{r}') = \lambda$, we can invoke the induction hypothesis once again: we can create profile $\mathbf{r}$ by letting w vote as in $\mathbf{x}$ (i.e. $\mathbf{r} = (x_w, \mathbf{r}'_{-w})$) and $g(\mathbf{r}) = S'$.

At this point, we reverse our previous moves. First, we create the profile $\mathbf{q}$ by allowing s' change its vote for s (i.e. $q_{s'} = s$ and $\mathbf{q} = (q_{s'}, \mathbf{r}_{-s'})$). Again, the strong sample property implies that $g(\mathbf{q}) = g(\mathbf{r}) = S'$, which results to $f(\mathbf{q}) = s$. Finally, we create profile $\mathbf{y}$ by letting s vote for w (i.e. $y_s = w$ and $\mathbf{y} = (y_s, \mathbf{q}_{-s})$). The induction hypothesis implies now that $g(\mathbf{y}) = S'$. Observe that $\mathbf{y}$ is indeed profile $\mathbf{z}$ (i.e. $\mathbf{y} = \mathbf{z}$), which is identical to $\mathbf{z}'$, except from the vote of w, hence $g(\mathbf{z}) = S'$. Recall however, that $g(\mathbf{z}) \neq S'$, a contradiction. Given that $d_s(\mathbf{z}') \leq \lambda - 1$, the claim follows.

If $d_s(\mathbf{z}') > \lambda - 1$, we generalize the idea of the previous paragraph to decrease the in-degree of s. We will revert enough edges towards s', for the in-degree of s to become equal to $\lambda - 1$. To do this we identify a tree T in $\mathbf{z}'$, with root s: We start from the nodes voting towards s and we select $d_s(\mathbf{z}) - (\lambda - 1)$ of them as children of s. We firstly select nodes with in-degree at most $\lambda - 1$. If we end up with nodes with higher in-degree that

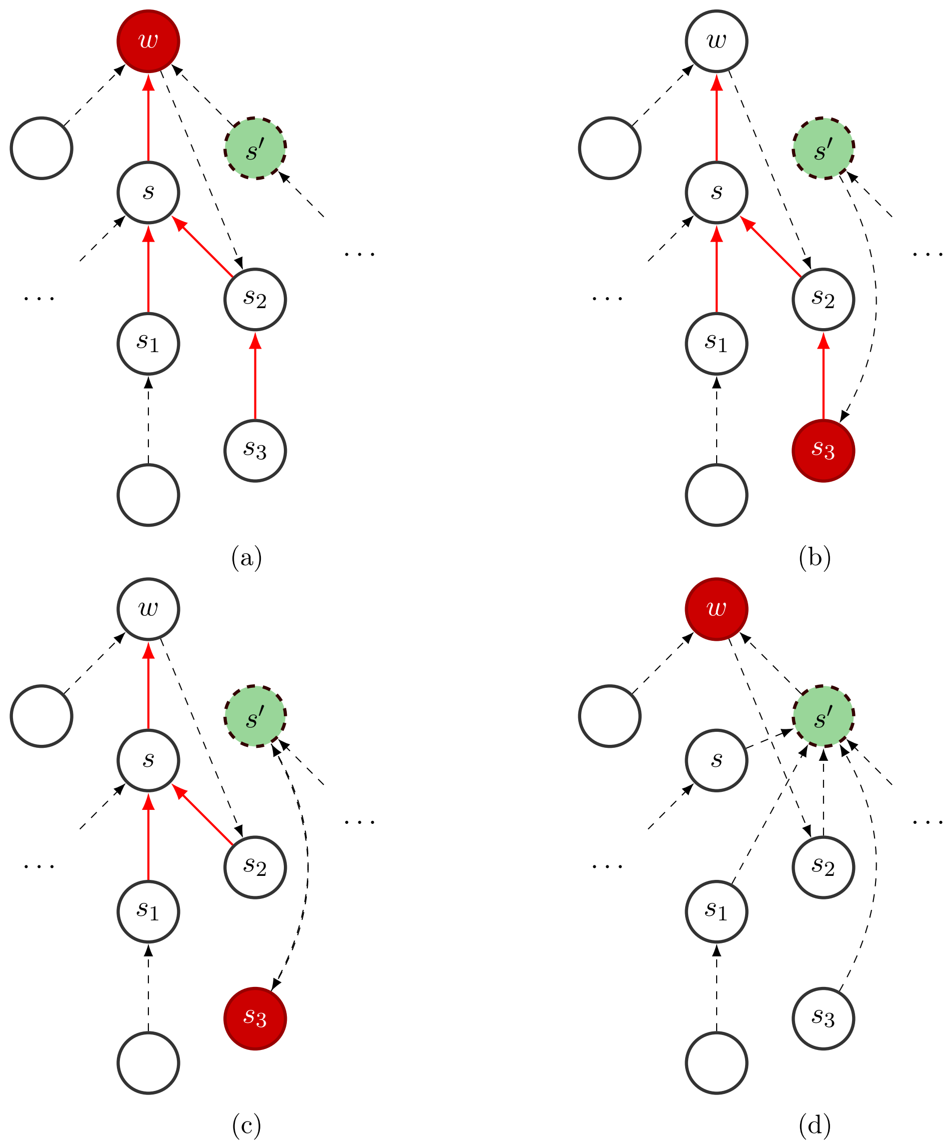


Figure 3.4: Induction step for the proof of Claim 3.7.1. An example with $\lambda = 2$. To use the induction hypothesis, we need to decrease the in-degree of w to 2. In Figure 3.4a, node s' is the single node in the sample, voting for the winner w. The rest of the sample is voting for s'. To decrease the in-degree of w, we identify the tree T, denoted by the solid thick edges. First, we let s' vote for s_3, which is now the winner (Figure 3.4b). Then we let s_3 vote for s'. By impartiality s_3 (with in-degree 1) retain its winner status and the sample set is still the same (Figure 3.4c). Finally, s' vote again for node s_2, turning it to the winner. The same procedure continues until all edges of T are redirected to s' and the in-degree of w decreases to 2, without any change in the sample set (Figure 3.4d).

$\lambda - 1$, we repeat the process for each child. Since we are in the single nomination model, there exists at most one directed path which passes from node s, which might not start from a node with degree 0.

Let k be the number of vertices in the tree T and let $\mathbf{r}^{(0)'} = \mathbf{z}'$; hence $g(\mathbf{r}^{(0)'}) = S'$. Starting from an arbitrary leaf on T, let v_i denote the i-th vertex we visit. For each $i \in \{1, ..., k\}$, we create three profiles: $\mathbf{y}^{(i)'}, \mathbf{q}^{(i)'}$ and $\mathbf{r}^{(i)'}$. First, we create profile $\mathbf{y}^{(i)'}$ by letting s' vote for v_i (i.e. $y_{s'}^{(i)'} = v_i$) and the remaining vertices vote like in $\mathbf{r}^{(i-1)'}$, i.e. $\mathbf{y}^{(i)'} = (y_{s'}^{(i)'}, \mathbf{r}_{-v_i}^{(i-1)'})$. Due to the strong sample property, $g(\mathbf{y}^{(i)'}) = g(\mathbf{r}^{(i-1)'})$. Also, $f(\mathbf{y}^{(i)'}) = v_i$ since v_i is the only vertex voted by the sample. Then we create the profile $\mathbf{q}^{(i)'}$, where $q_{v_i}^{(i)'} = s'$ and the other vertices vote like $\mathbf{y}^{(i)'}$, i.e. $\mathbf{q}^{(i)'} = (q_{v_i}^{(i)'}, \mathbf{y}_{-i}^{(i)'})$. Due to impartiality $f(\mathbf{q}^{(i)'}) = v_i$. If we are traversing the tree T from the leaves to the root, each vertex v_i has in-degree at most λ and by the induction hypothesis $g(\mathbf{q}^{(i)'}) = g(\mathbf{y}^{(i)'}) = S'$. Finally, we create the profile $\mathbf{r}^{(i)'}$ by letting s' vote for w, i.e $r_{s'}^{(i)'} = w$ and $\mathbf{r}^{(i)'} = (r_{s'}^{(i)'}, \mathbf{y}_{-s'}^{(i)'})$. We traverse the vertices starting from a leaf, and after visiting all vertices in the same level, we pass to the next level and we keep the order of the vertices visited. Note that in all these changes the *sample set does not change* and each vertex in T (including vertex s, which is traversed last, i.e. $v_k = s$) has in-degree at most λ. An example of this process is depicted in Figure 3.4.

At this point, we start a reverse procedure. We first create the profile $\mathbf{r}^{(k)} = (x_w, \mathbf{r}_{-w}^{(k)'})$, where we let vertex w to vote like in profile $\mathbf{x}$. By the induction hypothesis, $g(\mathbf{r}^{(k)}) = g(\mathbf{r}^{(k)'}) = S'$, since $f(\mathbf{r}^{(k)'}) = w$ and $d_w(\mathbf{r}^{(k)'}) \leq \lambda$. We then start to traverse the vertices in tree T on the opposite direction, i.e. $v_k, v_{k-1}, ..., v_1$. For each $i \in \{1, ..., k\}$ We create a similar series of profiles, where the sample set will remain invariant. Starting from $\mathbf{r}^{(i)}$ we create the profile $\mathbf{q}^{(i)}$, where s' votes towards $q_{s'}^{(i)'}$, i.e. $\mathbf{q}^{(i)} = (q_{s'}^{(i)'}, \mathbf{r}_{-s'}^{(i)})$. Due to the strong sample property $g(\mathbf{q}^{(i)}) = g(\mathbf{r}^{(i)}) = S'$ and $f(\mathbf{q}^{(i)}) = q_{s'}^{(i)'}$. Observe that $q^{(i)}$ and $q^{(i)'}$ differ only in the outgoing edge of w. As a result $d_{q_{s'}^{(i)'}}(\mathbf{q}^{(i)}) \leq \lambda$. We create now the profile $\mathbf{y}^{(i)}$ where $q_{s'}^{(i)'}$, the winning node in $\mathbf{q}^{(i)}$, votes towards $y_{q_{s'}^{(i)'}}^{(i)'}$, i.e. $y_{q_{s'}^{(i)'}}^{(i)} = y_{q_{s'}^{(i)'}}^{(i)'}$ and $\mathbf{y}^{(i)} = (y_{q_{s'}^{(i)'}}^{(i)}, \mathbf{q}_{-q_{s'}^{(i)'}}^{(i)})$. Again $\mathbf{y}^{(i)}$ and $\mathbf{y}^{(i)'}$ defer only in the vote of w. Because of the induction hypothesis $g(\mathbf{y}^{(i)}) = g(\mathbf{q}^{(i)}) = S'$. Finally, we revert s' towards w and create the profile $\mathbf{r}^{(i-1)}$ such that $r_{s'}^{(i-1)} = w$ and $\mathbf{r}^{(i-1)} = (r_{s'}^{(i-1)}, \mathbf{y}_{-s'}^{(i)})$. Again $g(\mathbf{r}^{(i-1)}) = g(\mathbf{y}^{(i)}) = S'$.

After this series of changes, we end up in profile $\mathbf{r}^{(0)}$, which differs from $\mathbf{r}^{(0)'}$ only in the outgoing edge of w. Since in all changes described above the sample set remains invariant, then $g(\mathbf{r}^{(0)}) = S'$. Observe now that $\mathbf{r}^{(0)} = \mathbf{z}$, for which we know that $g(\mathbf{z}) \neq S'$, a contradiction. This concludes the proof.

□

The next claim establishes the remaining case, that no node $u \notin g(\mathbf{x}), u \neq f(\mathbf{x})$ can change the sample set.

Claim 3.7.2. *Let (g, f) be an impartial deterministic strong sample mechanism, $\mathbf{x}$ be a nomination profile in $\mathcal{G}^1$ and u a node such that $u \notin g(\mathbf{x}), u \neq f(\mathbf{x})$. Then $g(\mathbf{x}) = g(x'_u, \mathbf{x}_{-u})$ for any other vote $x'_u \in N \setminus \{u\}$.*

Proof. For the sake of contradiction, consider any profile $\mathbf{x} \in \mathcal{G}^1$ and assume that there exists some nomination profile $\mathbf{x}' = (x'_u, \mathbf{x}_{-u})$ such that $g(\mathbf{x}') = S' \neq g(\mathbf{x})$. Starting from $\mathbf{x}'$, we define a profile $\mathbf{z}'$ in which all nodes in S' vote for u, and the rest vote as in $\mathbf{x}'$. That is, $z'_v = u$, for all $v \in S'$ and $z'_v = x'_v$ otherwise. Clearly $f(\mathbf{z}') = u$, as all the sample nodes vote for u. By Claim 3.7.1, we know that $g(x_u, \mathbf{z}'_{-u}) = g(\mathbf{z}') = S'$.

Starting from $\mathbf{x}$, we define a profile $\mathbf{z}$ in which all nodes in S' vote for u, and the rest vote as in $\mathbf{x}$. Since $S' \neq g(\mathbf{x})$, by Lemma 3.6, we get $g(\mathbf{z}) \neq S'$. Observe that $\mathbf{z} = (x_u, \mathbf{z}'_{-u})$, which leads to a contradiction. □

This completes the proof of Theorem 3.7. □

We next use Theorem 3.7 to obtain lower bounds on the additive approximation guarantee obtained by any deterministic strong sample mechanisms.

Corollary 3.8. *There is no impartial deterministic strong sample mechanism with additive approximation better than $n - 2$ in the single nomination model.*

Proof. Let S be the sample set which, by Theorem 3.7, must be selected independently of $\mathbf{x}$, and let $v \in S$. Define $\mathbf{x}$ so that all nodes in $N \setminus \{v\}$ vote for v and all other nodes have in-degree either 0 or 1. Then, $\Delta(\mathbf{x}) = n - 1$, but the mechanism selects a node of in-degree exactly 1. □

We remark that the strong sample mechanism that uses a specific node as singleton sample achieves this additive approximation guarantee.

Our next step, is to extend the notion of sample mechanisms to randomized variants and provide a lower bound on their additive approximation guarantee, which shows that

Sample and Vote (with $k = \Theta(\sqrt{n})$; see Section 3.3.1) is an optimal mechanism from this class. We next define the family of randomized strong sample mechanisms.

Definition 8. (Randomized strong sample mechanisms) A randomized strong sample mechanism (g, f) is a probability distribution over a family $\{(g_i, f_i) : i \in \mathbb{N}\}$ of strong sample mechanisms.

Note that Sample and Vote and Sample and Poll are both randomized strong sample mechanisms: For a given k, each of the possible sample sets define a deterministic sample mechanism, and the winner (if any) belongs in the set W. This is however not the case for more complex mechanisms like those appearing in [19] and [43].

Corollary 3.9. *There is no impartial randomized strong sample mechanism with additive approximation better than $\Omega(\sqrt{n})$ in the single nomination model.*

Proof. By Theorem 3.7, in any deterministic strong sample mechanism, the sample set is the same for any input graph $\mathbf{x}$. Hence, in a randomized strong sample mechanism, the probability that a node u belongs in the sample set, is affected only by the sample functions used by the mechanism. As such, it is independent of the input graph. Then for any such mechanism we can construct graphs which yield additive approximation $\Omega(\sqrt{n})$.

First, if there exists any node $v \in N$ with $\Pr v \in S > 1/\sqrt{n}$, then consider a nomination profile consisting of node v having maximum in-degree $\Delta = n - 1$ (i.e., all other nodes are pointing to it), with all other nodes having in-degree either 1 or 0. Since u^* belongs to the sample (and, hence, cannot be the winner) with probability at least $1/\sqrt{n}$, the expected degree of the winner is at most $1 + (n-1)(1 - 1/\sqrt{n}) = \Delta - \Theta(\sqrt{n})$.

Otherwise, assume that every node $v \in N$ has probability at most $1/\sqrt{n}$ of being selected in the sample set. Consider a nomination profile with a node $u^* \in N$ having maximum degree $\Delta = \sqrt{n}/2$ and all other nodes having in-degree either 0 or 1. Consider a node u pointing to node u^*. The probability that u belongs to the sample is at most $1/\sqrt{n}$. Hence, by the union bound, the probability that some of the $\sqrt{n}/2$ nodes pointing to u^* is selected in the sample set is at most $1/2$. Hence, the probability that u^* is returned as the winner is not higher than $1/2$ and the expected in-degree of the winner is at most $1 + \sqrt{n}/2 \cdot 1/2 = \Delta - \Theta(\sqrt{n})$. □

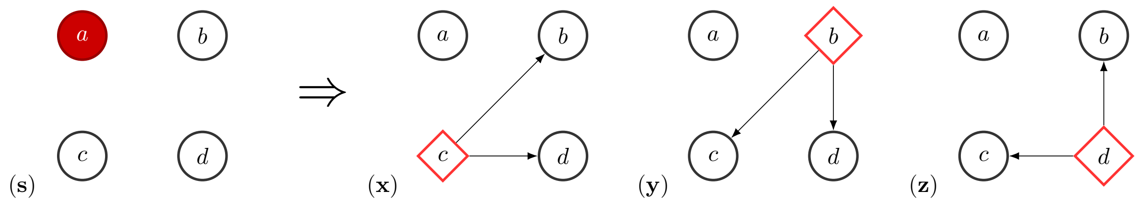


Figure 3.5: Node a is the winner in profile $\mathbf{s}$. This leads to the three profiles $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, where each diamond-shaped node cannot be the winner.

3.4.2 General Lower Bound

Our last result is a lower bound for all deterministic impartial mechanisms in the most general model of Alon et al. [2], where each agent can nominate multiple other agents or even abstain. We remark that our current proof applies to mechanisms that always select a winner.

Theorem 3.10. *There is no impartial deterministic α-additive mechanism for $\alpha < 3$.*

Proof. Let f be a deterministic impartial mechanism and, for the sake of contradiction, assume that it achieves additive approximation at most equal to 2. We will show that there is a profile with four nodes (denoted by a, b, c and d), in which the winner has in-degree 0, while the maximum in-degree is 3, which leads to a contradiction.

We first consider the profile with no edges, say $\mathbf{s}$, and let us assume, without loss of generality, that the winner is a (see Figure 3.5). Now consider the three profiles $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ produced when each of the other three nodes c, b and d vote for the other two of them, respectively (c votes for d and b, b votes for c and d, and d votes for b and c, as shown in Figure 3.5). In all these profiles, the voter (the node which changes its outgoing edges, compared to profile $\mathbf{s}$) cannot be the winner since this would break impartiality. Focus for example on the profile $\mathbf{x}$. Since c cannot be the winner, it must be either a, b or d. There are essentially two cases, which we treat separately.

Case 1: Node a is the winner for at least one of $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Consider the profile $\mathbf{x}$, where node c votes for both b and d and assume that $a = f(\mathbf{x})$ (see Figure 3.6). We let a vote for both b and d, to get the profile $\mathbf{w} = (\{(a, b)(a, d)\}, \mathbf{x}_{-a})$. Impartiality implies that $a = f(\mathbf{w})$.

On the one hand, if b votes for d (profile $\mathbf{w}' = (\{(b, d)\}, \mathbf{w}_{-b})$), impartiality implies that $b \neq f(\mathbf{w}')$ and approximation allows only $f(\mathbf{w}') = d$. On the other hand, if d votes for b

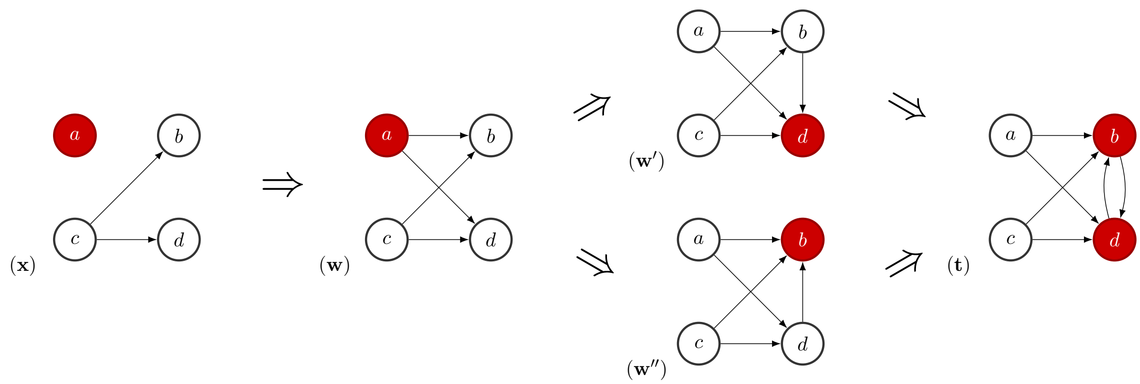


Figure 3.6: Case 1 in the proof of Theorem 3.10. We assume that a is the winner in profile $\mathbf{x}$. When a adds votes, the profile $\mathbf{w}$ is created. Then, in the upper profile $\mathbf{w}'$, node d must win (for an additive approximation guarantee strictly less than 3) and in the lower profile $\mathbf{w}''$, node b must win. This however leads to the final profile $\mathbf{t}$, where both d and b must win, due to impartiality —a contradiction.

(profile $\mathbf{w}'' = (\{(d,b)\}, \mathbf{w}_{-d})$), by similar arguments we have $f(\mathbf{w}'') = b$ (see Figure 3.6). Now, consider the profile $\mathbf{t}$ where both b and d vote for each other, i.e., $\mathbf{t} = (\{(d,b)\}, \mathbf{w}'_{-d})$ and, at the same time, $\mathbf{t} = (\{(b,d)\}, \mathbf{w}''_{-b})$. Impartiality (applied to $\mathbf{w}'$ and $\mathbf{w}''$, respectively) implies that both b and d must be winners which is absurd and leads to a contradiction. Similar arguments would apply for the other cases, establishing that a cannot be the winner in any of the profiles $\mathbf{x}, \mathbf{y}$ or $\mathbf{z}$.

Case 2: Node a is not the winner for any $\mathbf{x}, \mathbf{y}, \mathbf{z}$. In this case, due to impartiality, only nodes with in-degree 1 are possible winners. Hence, we are left only with two sub-cases; either two of these profiles share the same winner or all of them have a different winner.

In the first sub-case, consider (without loss of generality) the scenario where $f(\mathbf{x}) = f(\mathbf{y})$. Impartiality, plus the fact that a is not the winner in $\mathbf{x}$, imply that $f(\mathbf{x}) = d$. Assume that $f(\mathbf{z}) = b$ (illustrated in Figure 3.7). The alternative case $f(\mathbf{z}) = c$ follows through similar arguments. In profile $\mathbf{y}$, we let d add 2 votes and create profile $\mathbf{t}' = (\{(d,b)(d,c)\}, \mathbf{y}_{-d})$. By impartiality, $f(\mathbf{t}') = d$. In profile $\mathbf{z}$, we let b add 2 votes and create again the profile $(\{(b,c)(b,d)\}, \mathbf{z}_{-b}) = \mathbf{t}'$. Note that these graphs are, indeed, the same. By impartiality $f(\mathbf{t}') = b$, hence the impartial mechanism f at profile $\mathbf{t}'$ must award two nodes, a contradiction. Similar arguments hold in all the cases where two of the profiles $\mathbf{x}, \mathbf{y}, \mathbf{z}$

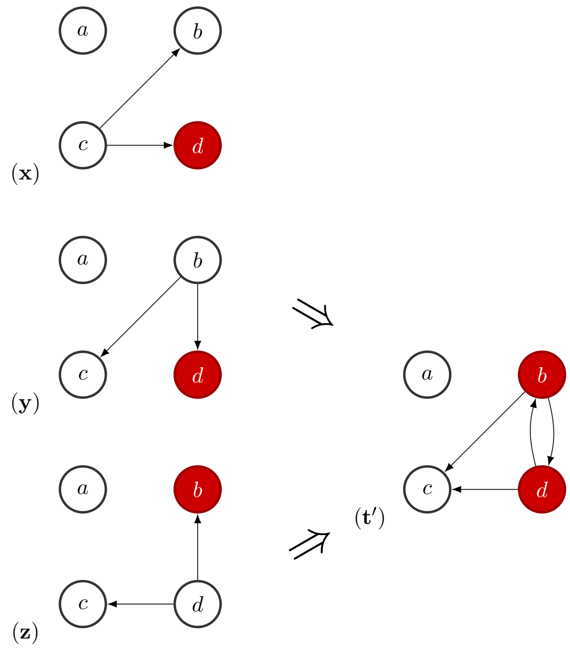


Figure 3.7: Case 2 in the proof of Theorem 3.10. Node a is not the winner in any of the profiles $\mathbf{x}, \mathbf{y}, \mathbf{z}$. We assume that two of the winners in $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are the same, and this leads to a mechanism with two winners —a contradiction.

share the same winner.

We are left now with the case where all these profiles, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ have different winners, where none of them is a. There are 2 possible such scenarios: $f(\mathbf{x}) = d$, $f(\mathbf{y}) = c$ and $f(\mathbf{z}) = b$ (see Figure 3.8), or $f(\mathbf{x}) = b$, $f(\mathbf{y}) = d$ and $f(\mathbf{z}) = c$. Consider the first one (similar arguments hold also for the second). From these profiles $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ we reach the profiles $\mathbf{x}' = (\{(d,b)(d,c)\}, \mathbf{x}_{-d})$, $\mathbf{y}' = (\{(c,b)(c,d)\}, \mathbf{y}_{-c})$ and $\mathbf{z}' = (\{(b,c)(b,d)\}, \mathbf{z}_{-d})$, by letting the respective winners to add edges. Because of impartiality, all the winners are preserved, i.e., $f(\mathbf{x}) = f(\mathbf{x}')$, $f(\mathbf{y}) = f(\mathbf{y}')$ and $f(\mathbf{z}) = f(\mathbf{z}')$. Let us now focus on profile $\mathbf{x}'$. By letting node b add edges (b,c) and (b,d), we create the profile $\mathbf{t}'' = (\{(b,c)(b,d)\}, \mathbf{x}'_{-b})$: a directed clique on the nodes b, c and d and the node a with no incoming nor outgoing edges (see Figure 3.8). Focusing now on profile $\mathbf{y}'$, we reach the profile $(\{(d,b)(d,c)\}, \mathbf{y}'_{-d}) = \mathbf{t}''$ by a deviation of node d; the same profile as before. In a similar fashion, on profile $\mathbf{z}'$ we reach the profile $(\{(c,b)(c,d)\}, \mathbf{z}'_{-c}) = \mathbf{t}''$ by a deviation of c. By impartiality, $f(\mathbf{t}'') \notin \{b,c,d\}$, which implies that $f(\mathbf{t}'') = a$. Now, if a votes for at least any other node, impartiality implies that a must remain the winner, while the nominees of a will have in-degree 3, contradicting the

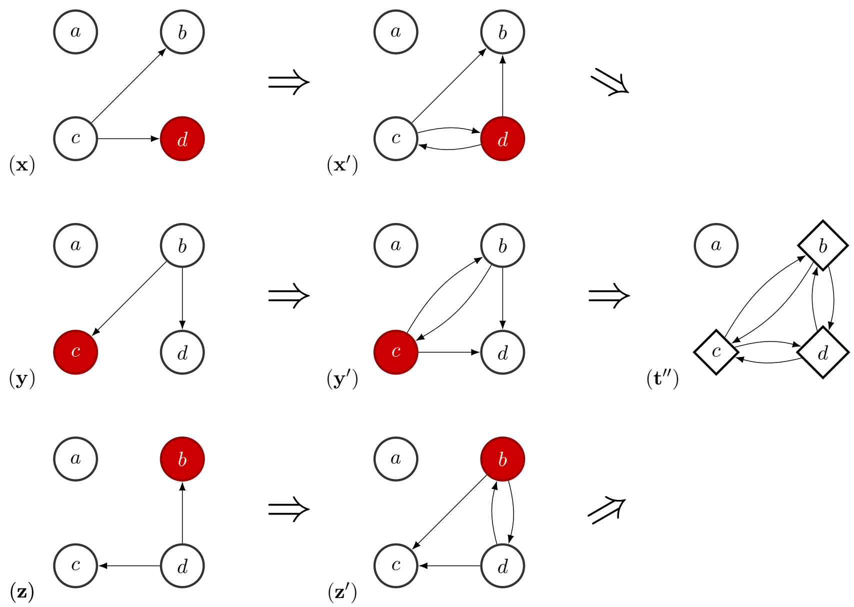


Figure 3.8: Case 2 in the proof of Theorem 3.10. Node a is not the winner in any of the profiles $\mathbf{x}, \mathbf{y}, \mathbf{z}$. we assume that no two profiles among $\mathbf{x}, \mathbf{y}, \mathbf{z}$ have the same winner. Impartiality implies that in the rightmost profile $\mathbf{t}''$ where three nodes have in-degree 2, the winner must be a, with in-degree 0. When a votes towards any other node, a should remain a winner, while there is some node with in-degree 3 —a contradiction.

approximation guarantee of f. □

3.5 Additive Approximation Bounds for Known Mechanisms

In this section, we provide additive approximation bounds for known mechanisms. We have introduced these mechanisms in Section 1.2.1.

3.5.1 Deterministic Mechanisms

In this subsection, we present additive approximation bounds for the mechanisms presented in [57]. All these mechanisms are defined for the single-nomination model. Table 3.1 summaries the additive approximation bounds. In contrast to the approximation ratio, additive approximation guarantees can be used effectively to compare deterministic mechanisms.

For the multiple-nomination model, no deterministic mechanism has been explicitly proposed. However, Plurality with Default, the Fixed Winner and version of the

Mechanism	Lower Bound	Upper Bound
MAJORITY WITH DEFAULT	$n/2$	$n/2$
PLURALITY WITH DEFAULT	$\lceil n/2 \rceil$	$\lceil n/2 \rceil$
PARTITION METHODS WITH K DISTRICTS	$n - \frac{\lambda}{2} + 1$	$n - \frac{n}{2K} - 1$
CROSS-PARTITION METHODS	$n - 4$	$n - 1$
Strong Sample mechanisms	$n - 2$	$n - 1$

Table 3.1: Additive approximation bounds for various deterministic mechanisms. The bounds are computed for the case of the single nomination model. In the guarantee for the PARTITION METHODS, λ denotes the minimum district size. The upper bound on CROSS-PARTITION Methods is trivial.

Mechanism	Lower Bound	Upper Bound
SAMPLE AND POLL	$\Omega(\sqrt{n})$	$\mathcal{O}(n^{2/3} \log^{2/3} n)$
PERMUTATION	$\Omega(n)$	$\mathcal{O}(n)$
SLICING	$\Omega(\sqrt{n})$	$\mathcal{O}(n^{8/9})$

Table 3.2: Additive approximation bounds for various randomized mechanisms. The upper bound for the SLICING mechanism is derived from [19]. The upper bound on the Permutation mechanism is trivial. The rest bounds are from this work.

PARTITION METHODS mechanisms (using plurality instead of majority for identifying local winners) naturally extend to the multiple-nomination model. They all however have an additive approximation of $n - 1$, which can be proven almost trivially. A nomination profile that combines a clique between the $n - 1$ node excluding the default winner and $n - 1$ from the default winner to the other nodes will do the trick. The randomized mechanisms we have presented in Section 1.2.1 have deterministic counterparts (for example, a deterministic version of the permutation mechanism will examine the nodes according to a pre-determined order). In all these mechanisms, there is at least one node that can never win (i.e. No Exclusion fails), hence they are all at least $(n - 1)$-additive in both single-nomination and multiple-nomination models. We will discuss these mechanisms in more detail later on.

Theorem 3.11. *The* MAJORITY WITH DEFAULT *mechanism is at most $n/2$-additive in the single nomination model.*

Proof. Consider any nomination graph $\mathbf{x} \in \mathcal{G}^1$. Let u^* be the node with maximum in-degree and let $t \in N$ be the default agent. Clearly, when u^* is supported by the majority, i.e. $\Delta(\mathbf{x}) = d_{u^*}(\mathbf{x}) \geq d_{u^*}(N \setminus \{t\}) \geq \lceil n/2 \rceil$, then u^* is the winner and the additive approximation is 0. Hence the additive approximation cannot be larger than the maximum in-degree, when no majority exists, i.e. when $\Delta(\mathbf{x}) \leq \lceil n/2 \rceil - 1 \leq n/2$ and the MAJORITY WITH DEFAULT mechanism is at most $n/2$-additive. □

The following example shows that the MAJORITY WITH DEFAULT mechanism has additive approximation exactly $n/2$.

Example 4. Consider an instance $\mathbf{x} \in \mathcal{G}^1$ with n nodes, where n is even. Exactly $n/2$ nodes vote for u^*. One of these voters is the default agent v. The remaining nodes vote for each other, in a directed circle. Finally, u^* votes for some node different than v. Since $d_{u^*}(N \setminus \{v\}, \mathbf{x}) < n/2$, u^* is not the winner. The winner is the default node with in-degree 0, and the additive approximation is at least $n/2$.

The PLURALITY WITH DEFAULT mechanism yields similar additive approximation guarantees.

Theorem 3.12. *The* PLURALITY WITH DEFAULT *mechanism is at most* $\lceil n/2 \rceil$*-additive in the single nomination model.*

Proof. We will show that the mechanism returns the maximum in-degree node when the in-degree is strictly higher than $\lceil n/2 \rceil$. Consider any nomination graph $\mathbf{x} \in \mathcal{G}^1$. Let $u^* \in N$ be a node with maximum in-degree and assume for the sake of contradiction that $\Delta(\mathbf{x}) = \delta_{u^*}(\mathbf{x}) \geq \lceil n/2 \rceil + 1$ while u^* is not the winner. Hence the winner is some different node $v \in N \setminus \{u^*\}$.

Assume first that v is not the default node, i.e. the default node is $t \in N \setminus \{v\}$[4]. Node v wins when $d_v(N \setminus \{t, v'\}, \mathbf{x}) > d_{v'}(N \setminus \{t, v\}, \mathbf{x})$ for any $v' \in N \setminus \{v, t\}$. Clearly $d_{u^*}(N \setminus \{t, v\}, \mathbf{x}) \geq \lceil n/2 \rceil - 1$ and for v to win it must be that $d_v(N \setminus \{t, u^*\}, \mathbf{x}) > \lceil n/2 \rceil - 1$. There exists exactly n edges, hence the remaining edges are $\lfloor n/2 \rfloor - 1$, and the second highest in-degree is at most $\lfloor n/2 \rfloor - 2$ (since that node need to vote, too). Since $\lceil n/2 \rceil + 1 > \lfloor n/2 \rfloor - 2$ for any positive integer n, v cannot be the winner. A contradiction.

Assume now that the winner node v is the default node. Then there exists at least one node with in-degree at least $\lfloor n/2 \rfloor - 1$. otherwise, u^* would be the winner. As we have seen in the previous paragraph, no other node can have an in-degree higher than $\lfloor n/2 \rfloor - 1$. A contradiction. □

The following example shows that this bound is also tight.

[4]Note that u^* can be the default node.

Example 5. Consider a nomination graph $\mathbf{x}$ with $n \in \mathbb{N}_{\geq 2}$ nodes for n odd. Let that node $u \in N$ has in-degree $\lfloor n/2 \rfloor$ while node u^* has in-degree $\lceil n/2 \rceil$. The default node is node $t \in N \setminus \{u, u^*\}$. Nodes u and u^* vote for each other and the rest nodes have in-degree 0. The default agent votes for u^*. Hence $d_{u^*}(N \setminus \{u,t\}) = \lceil n/2 \rceil - 2 = \lfloor n/2 \rfloor - 1 = d_u(N \setminus \{u^*, t\})$, since n is odd. The default node t is the winner, with in-degree 0.

The Partition Methods is a family of mechanisms proposed in [57]. A detailed implementation of the mechanism appears in Algorithm 3 in Chapter 2 and an example of this mechanism is shown in Figure 1.4 in Chapter 1. These mechanisms partition the set of nodes into $K \geq 2$ districts. If a node receives an absolute majority in each set is declared a *local winner*, and then an election between the local winners (with them as candidates and the rest nodes as voters) decides the winner. If no local winner exists, a default node (which belongs by convention in the first district) becomes the winner.

In the following, we first show that these mechanisms are at most $(n - \frac{\lambda}{2} + 1)$-additive, where λ is the size of the minimum size district, and then we will show an almost tight lower bound, by presenting a family of examples where the default node is the winner for any number of districts.

Theorem 3.13. *A Partition Method mechanism with minimum district size λ is at most $(n - \frac{\lambda}{2} + 1)$-additive.*

Proof. Consider any nomination graph $\mathbf{x} \in \mathcal{G}^1$. Let $S_1, ..., S_K$ be the districts, for $K \geq 2$ and let that $|S_j| \geq |S_{j+1}|$ for $j \in \{1..K-1\}$.

For simplicity, we assume that the default node t can be in any district and that a local winner $v \in S_j$ exists when $d_v(S_j, \mathbf{x}) \geq \lceil (|S_j| + 1)/2 \rceil$, counting even the default node's vote. Note that this cannot affect the impartiality of the mechanism and it can only increase the approximation.

Let $z = \frac{|S_K|}{2} - 1$ and $d_{u^*}(N, \mathbf{x}) \geq n - z$. Since K is the minimum size district, $n - z \geq n/2$ and u^* is a node with maximum in-degree. Hence, there exists exactly z nodes not voting for u^* (including u^*). Suppose now that u^* is located in a district which include the other $z - 1$ nodes not voting for u^*. Then

$$d_{u^*}(S_k, \mathbf{x}) \geq |S_K| - z = \frac{|S_K|}{2} + 1 \geq \left\lceil \frac{|S_K| + 1}{2} \right\rceil \tag{3.5}$$

and u^* is the winner. The theorem follows. □

The following example shows that our upper bound is tight.

Example 6. Without loss of generality, let K be the district with smaller size. We can construct a graph $\mathbf{x} \in \mathcal{G}^1$ where all nodes in $\cup_{k=1}^{K} S_k$ vote of some node $v \in S_K$. Node v is also voted by almost half the voters in S_K, i.e. $d_v(S_K, \mathbf{x}) = \lfloor S_K/2 \rfloor < \lceil (S_K+1)/2 \rceil$. The default node t vote is one of the voters of v. All nodes not voting for v form a directed cycle. Hence, there exists no local winner, and t is the winner, with in-degree 0. Hence the additive approximation is exactly $n - \left\lfloor \frac{S_K}{2} \right\rfloor - 1 \geq n - \frac{S_K}{2} - 1 \geq n - \frac{n}{2K} - 1$. The inequalities are due to $\lfloor z/2 \rfloor \leq z/2$ for any $z \in \mathbb{N}$ and $S_K \leq n/K$.

Cross-Partition methods Cross-Partition methods is a quite involved procedure, proposed by [57]. An example is depicted in Figure 1.5 and a detailed algorithm for the selection of the winner is given in Algorithm 5. To show additive approximation bounds, we roughly present the mechanism here, hiding some unnecessary details. This method partitions the nodes into $K \geq 3$ districts, $S_1, ..., S_K$ of nearly equal size (any two districts have a difference in size at most 1), as well as another district including only the default winner. Each district S_k is also partitioned into two orthogonal partitions $(\Pi_k, \bar{\Pi}_k)$ and $(\Lambda_k, \bar{\Lambda}_k)$. As an example, each node is assigned a color, dark or light (resp. the node belongs in Π_k or $\bar{\Pi}_k$), and a shape, circle or diamond (resp. the node belongs in Λ_k or $\bar{\Lambda}_k$). Each district contains all 4 possible combinations of shapes and colors: (dark, diamond), (light, diamond), (light, circle), and (dark, circle).

The mechanism works in two steps. A node in district S_k is called an outer-hero if it gets the votes of all nodes in Π_l or $\bar{\Pi}_l$ for every l par $(k-1 \mod K)$ and $(k \mod K)$, and all nodes in $\Lambda_{(k-1 \mod K)}$ or $\bar{\Lambda}_{(k-1 \mod K)}$. If outer-heroes exist, they are the only eligible winners. If no outer heroes exist, the mechanism searches for inner-heroes. These are nodes that are voted by everyone in their own district. If such nodes exist, they are the only eligible winners. Otherwise, the default node wins.

Given a Cross-Partition method, described by all the partition sets, we can easily construct an input where the default node must win, with in-degree 0, while another node has quite a large in-degree.

Example 7. Let that $u^* \in S_1$ be a node which is voted by almost all nodes in S_1 (just one other node in S_1 is not voting for u^*). Node u^* is also voted by all other nodes, par a node belonging in Π_2 and a node belonging in $\bar{\Pi}_2$. Hence u^* is no outer-hero. Similarly, u^* is not an inner-hero, since someone in its own district is not voting for it. Notice also that no other outer hero exists: only 3 nodes are not voting for u^*, which are not enough to declare an outer-hero. Similarly, no other inner-hero exists, since in each district at least one node votes for u^*. Eventually, the default node is the winner with in-degree 0 (the 3 nodes not voting for u^* are also not voting for the default node). The additive approximation of these nomination profiles is $n-4$.

Finally, we discuss the case of deterministic mechanisms, inspired by the randomized mechanisms we have presented. First, note that a deterministic version of 2 PARTITION is a strong sample mechanism, hence it yields an $n-1$ lower bound, even for the single-nomination model. PERMUTATION, k-PARTITION and SLICING are not strong sample mechanisms, however, their deterministic versions preclude at least one node from being a winner. To make this clear we note that the deterministic version of the permutation mechanism will follow one-by-one the steps of the randomized version, following a deterministic order of the nodes: as a result, the first node can never win. Similarly, the k-PARTITION mechanism will use pre-determined partitions, and any node in the first partition can never win. SLICING will use a pre-determined starting sample. In all these cases, a star graph where the node which cannot win has in-degree $n-1$ leads to an $n-1$ lower bound for the multiple-nomination model, and $n-2$ for the single-nomination model.

3.5.2 Randomized Mechanisms

In this section, we present additive approximation lower bounds for randomized mechanisms. Table 3.2 summarizes these results. We use The 2-PARTITION mechanism as a warm-up. It's not hard to see that this mechanism yields at least linear additive approximation. Recall that this mechanism partitions the nodes in two sets, S and W and each node belongs to any of these sets with probability 1/2 and the winner is a node in W with maximum in-degree from S. Consider a nomination graph where a single node is nominated by everyone. This node is assigned in set S with probability 1/2 and the winner is a node with in-degree 0. Hence 2-PARTITION is at least $\frac{n-1}{2}$ optimal.

The SAMPLE AND POLL mechanism is essentially a modified 2-PARTITION mechanism which yields a much more improved additive approximation bound. We note here that the construction used in Corollary 3.9 implies a $\Omega(\sqrt{n})$ lower bound for this mechanism,

We continue with a lower bound on the PERMUTATION mechanism from [43]. This mechanism is randomized, and guarantees that the expected degree of the winner is at least half of the optimal winner's degree, which is best possible. We will show that the additive approximation of this mechanism is linear in the number of nodes in some instances.

The PERMUTATION mechanism considers a random permutation $\mathbf{p}$ of the nodes and examines every node sequentially according to this permutation, from left to right. At first, the mechanism sets the node p_1 as the provisional winner. When examining node in position $i \in \{1, ..., n\}$, the mechanism compares the in-degree of node p_i from nodes in positions $\{1, ..., i-1\}$ with the in-degree of the provisional winner. We call this value the *score* of p_i. Crucially, for impartiality, the outgoing edges from the provisional winner are ignored. If the score of p_i is equal to the provisional winner's score, p_i becomes a new provisional winner. The final provisional winner is the node selected by the mechanism. We formally describe the mechanism in Algorithm 7. In the following theorem, we show that the PERMUTATION mechanism yields additive approximation at least linear in the number of nodes.

Theorem 3.14. *The* PERMUTATION *mechanism is at least* $\Omega(n)$*-additive.*

Proof. Consider an nomination graph $\mathbf{x} \in \mathcal{G}$ where an node $u^* \in N$ is the maximum in-degree node with in-degree $2n/3$, i.e $\Delta(\mathbf{x}) = d_{u^*}(\mathbf{x}) = 2n/3$. Node $v \in N \setminus \{u^*\}$ is the second maximum in-degree node and $d_v(\mathbf{x}) = n/3$. All the other nodes have in-degree 0. We assume that n is divisible by 3. It is sufficient to show that the mechanism fails to declare as node u^* as a winner with constant probability. When the probability $\Pr[u^* \text{ is not the winner}] \geq c$, for some $c \in [0, 1]$ independent of the number of agents n, then $\mathbb{E}\left[\, d_w \,\right] \leq \Delta(\mathbf{x}) \cdot (1-c)$ and $\Delta(\mathbf{x}) - \mathbb{E}\left[\, d_w \,\right] \geq c \cdot \Delta(\mathbf{x}) = \Omega(n)$. We will refer to the $2n/3$ nodes voting for u^* as red nodes and the $n/3$ nodes voting for u as blue nodes.

Let $\mathbf{p} = (p_1, ..., p_n)$ denote the permutation by which the nodes are examined. Consider now the event where node u is examined after node u^*, and between them there is a gap of at least $2n/3$ nodes, i.e. $u^* = p_i$ and $u = p_j$ for some $i \leq 2n/3 + j$. In this event node u is the winner. To see this, suppose that $p_i - 1$ red nodes are located before u^*. When this is the case, then $2n/3 - p_i + 1$ red nodes will be arranged in various positions after u^*. If all these positions are before u, then there are at least $2n/3 - (2n/3 - p_i + 1) = p_i - 1$ blue nodes before node u. Thus u is declared the winner. Note that this is the worst case for u. If any of the red nodes, now located in a position with index lower than i, is moved to a position after u, then the score of u can only increase. If the red voters before u^* are less than $p_i - 1$ (as we supposed before) then the score of u^* will decrease, while the score of u can only increase.

We count now all those permutations. Clearly u^* can only be located between position 1 and $n/3$. If u^* is fixed at position 1 then u can be anywhere between positions $\{\lceil 2n/3 \rceil, ..., n\}$.

When u^* is located at position 2, u can be anywhere in $\{\lceil 2n/3 \rceil + 1, ..., n\}$. In a similar manner, for each fixed position $l \in \{1..\lfloor n/3 \rfloor\}$ of u^*, there exists $n/3 - l$ ways to arrange u and $(n-2)!$ ways to arrange all other nodes. Thus a total of $(n-2)! \sum_{l=2n/3}^{n}(n-l) = \frac{1}{18}n(n+3)(n-2)!$ permutations have a gap of at least $2n/3$ position between u^* and u and u^* is before u. Dividing over all possible permutations we get the probability of this happening:

$$\begin{aligned}\Pr[\ u \text{ wins }] &\geq \Pr[\ u^* \text{ is located before } u \textbf{ and } (j-i \geq 2n/3)] \\ &\geq \frac{1}{18} \cdot \frac{n(n+3)(n-2)!}{n!} \\ &\geq \frac{1}{18}, \end{aligned} \tag{3.6}$$

for every $n \in \mathbb{N}_{\geq 2}$. The additive approximation for this nomination profile is at least:

$$(d_{u^*}(\mathbf{x}) - d_u(\mathbf{x})) \cdot \Pr[\ u \text{ wins }] \geq \frac{n}{3} \cdot \frac{1}{18} = \Omega(n). \tag{3.7}$$

□

Finally, we provide a lower bound for the Slicing Mechanism, proposed in [19]. The mechanism selects first a sample set on nodes, and uses the outgoing edges to partition the unsampled nodes in slices, and then uses an election process where a provisional winner is updated after examining each slice. In Algorithm 8 we present this mechanism in detail. Our lower bound is simple and shows that of any selection of the parameter ϵ, there exists a graph that implies an additive approximation of $\sqrt{n}$. An interesting open question is whether this lower bound can be pushed further up, or the Slicing mechanism yields an additive approximation of $\mathcal{O}(\sqrt{n})$.

Theorem 3.15. *The* Slicing *mechanism is at least* $\Omega(\sqrt{n})$*-additive.*

Proof. We assume first that $\epsilon \geq \frac{1}{\sqrt{n}}$. Consider the graph $\mathbf{x}^1 \in \mathcal{G}$. Graph $\mathbf{x}^1$ is a directed star, where all nodes par u^*, vote for u^*. We focus now on the Slicing mechanism. Each node is selected in the sample with probability $\epsilon \geq \frac{1}{\sqrt{n}}$. The Slicing mechanism will return a node with in-degree zero if node u^* is selected in the sample. Otherwise, it will return node u^* as the winner. Hence the expected additive approximation is equal to $n - n\left(1 - \frac{1}{\sqrt{n}}\right) = \sqrt{n}$ in this case.

We assume now that $\epsilon < \frac{1}{\sqrt{n}}$, and we consider graph $\mathbf{x}^2 \in \mathcal{G}$, which can be described with three sets of nodes. Set A where all nodes vote for each other in a directed clique, set B, where all nodes vote for a single node $u^* \in A$, and set C. In set C all nodes are voting for each other in pairs (we assume that $|C|$ is even) and are all voted by all nodes in A. Let that $|A| = n/2$ and $|B| = \sqrt{n}$ and $|C| = n - |A| - |B|$. We examine the SLICING mechanism starting from the sampling phase. We are interested in the event where no node from B is selected in the sample set X. This happens with probability $\Pr[X \cap B = \emptyset] = (1-\epsilon)^{|B|} \geq \left(1 - \frac{1}{\sqrt{n}}\right)^{\sqrt{n}}$. Given that this happens, the sample set has only nodes from sets A and C. This implies that all unsampled nodes have expected in-degrees equal to $\epsilon^{-1} \cdot |X \cap A|$ or $\epsilon^{-1} \cdot (|X \cap A| + 1)$. The first one concerns all nodes in set $A \setminus X$, all nodes in B and some nodes in $C \setminus X$. The second concerns nodes in $C \setminus X$, who are not selected in the sample set but their paired note is not selected in X.

In the slicing phase, we note that there exist at least n slices, hence each slice will include only nodes with exactly the same in-degree. Given that at least one node from C belongs in the sample and its paired note does not belong in the sample there exists at least one node with expected in-degree equal to $\epsilon^{-1} \cdot (|X \cap A| + 1)$. This can happen with probability at least $1 - (1 - \epsilon(1-\epsilon))^{|C|/2}$. At this point, u^* is not the provisional winner, since there exists at least one node with expected in-degree strictly higher than u^*. Node u^* belongs in the penultimate slice, with all other nodes in $A \setminus X$ and the nodes in B. The mechanism gives u^* an ϵ chance to become the provisional winner. If this does not happen u^* loses, and a node belonging in C, with in-degree $n/2 + 1$ is the winner. The probability of u^* losing, when is not selected in the sample X, is at least

$$\begin{aligned}
\Pr[u^* \text{ loses}] \geq\ & (1-\epsilon)^{|B|} \cdot (1 - (1 - \epsilon(1-\epsilon))^{|C|/2}) \cdot (1-\epsilon)^2 \\
\geq\ & \left(1 - \frac{1}{\sqrt{n}}\right)^{\sqrt{n}} \cdot \left(1 - \left(1 - \frac{1}{2\sqrt{n}}\right)^{n/2+\sqrt{n}}\right) \cdot \left(1 - \frac{1}{\sqrt{n}}\right)^2 \\
\geq\ & e^{-2}
\end{aligned} \tag{3.8}$$

for $n \geq 8$. In the first inequality we have used $\epsilon(1-\epsilon) \geq \epsilon/2$ for $\epsilon \in [0, 1/2]$ and $\epsilon < \frac{1}{\sqrt{n}} < \frac{1}{2}$.

This implies that the additive approximation for this nomination profile is at least $\frac{\sqrt{n}-1}{e^2} = \Omega(\sqrt{n})$. □

Chapter 4

Impartial Selection with Prior Information

This chapter is based on joint work with George Christodoulou and Ioannis Caragiannis. The problem in its more general form is introduced in Section 1.2 and some formal definitions for this work are given in Section 2.2.

4.1 Our Results

In this chapter, we continue our work on the problem of impartial selection, taking now a different perspective. We get inspiration from the successful design of auctions and posted prices mechanism for social welfare and profit maximization, and we consider models where some prior information is available, both to the mechanism designer and all participants. A selection mechanism takes a graph as input and returns a single node as the winner. Recall that the nodes of the graph represent strategic individuals, and they might attempt to manipulate a selection mechanism if this would increase their chance of winning. An impartial mechanism is robust to such strategic behavior, by ensuring that the chance of a node becoming the winner, does not depend on its outgoing edges.

One of the most natural and well-studied selection rule is the *approval voting rule* (AV), which has received much attention in social choice theory [62]. In our context, AV always returns the node with the highest in-degree. Unfortunately, as already mentioned in Chapters 1 and 3, this mechanism is not impartial. The reason is that in case of a tie at the maximum degree, some of the nodes involved in the tie may have an incentive to vote non-truthfully, as we have already seen in the previous chapters. Fortunately, there is a simple fix to this deficiency, which is inspired by the *plurality with default* mechanism by Holzman and Moulin [57], already discussed in Section 1.2.1; *in case of a tie, select as winner a predetermined/default node.* We refer to this modified version of AV as *approval voting with default* (AVD). Although (a careful implementation of) this tweak re-establishes impartiality, this modification comes at a cost, as this preselection should be independent of the input graph. Imagine a scenario, where there is a tie between two nodes with the maximum degree. In the unfortunate situation where the default node receives only a small number of votes, this might lead to a poor additive approximation, linear in the number of nodes.

Most of the previous work, including the work presented in Chapter 3, consider randomized mechanisms, hence the efficiency is measured in expectation. However, in the design of selection mechanisms, determinism is arguably more desirable. Unfortunately, all the known deterministic mechanisms have very poor, linear additive approximation, and it is wide open whether substantially better mechanisms exist. The approximation ratio criterion is even worse for these mechanisms, as no deterministic mechanism yields a non-infinite approximation ratio.

In this work, we take a different route: getting our motivation from the successful application of prior information to auction and posted pricing mechanisms (see an excellent survey of related work in [55]), we study the extent to which *prior information* on the preferences of the voters could allow the design of deterministic impartial selection mechanisms with good additive approximation guarantees.

Our focus is on the analysis of AVD, for which our design choice boils down to an effective choice of the default node, with the help of the prior information. We assume that the preferences are drawn from a probability distribution that is known to the mechanism. We assume throughout *voter independence*,[1] that is, the random choice of preferences for each voter is independent of those of the others.

We build on the *multiple-nomination* model (see Section 1.2), and we propose different information models that capture several aspects of the problem. In the *opinion poll* model, we assume that the prior information concerns information about the preferences of different (types of) voters. The designer has access to the probability $p_i(S)$, with which voter i (or all voters of type i) would approve a subset S of candidates. The *a priori popularity* model

[1] We should note that, with correlated distributions there is not much one can achieve (see Example 8 in Section 4.5).

assumes that the designer has prior information about the *popularity* of each candidate j, which is summarized by a scalar p_j. We assume that each candidate j receives independently a vote fro each voter with probability p_j. As a special case, we also study the uniform model, in which every candidate j has the same popularity $p_j = p$.

Note that these models capture different information scenarios; the former assumes that the designer has access to opinion poll statistics for each (type of) voter, while the latter assumes that the designer has access only to aggregate information about the popularity of a candidate. This aggregation is over the whole population of voters, as the actual information may be sanitized to preserve anonymity of those (types) who participated in the poll. Note that popularity may measure other forms of biases over specific individuals. For example, consider the situation in which a PC wants to decide the best paper award; then, the a priori popularity of a paper could be a function of the authors' esteem, affiliation, etc.

4.1.1 Contribution and Techniques

Our main focus is the analysis of the AVD mechanism. We begin with the opinion poll model and, as a warm-up, in Section 4.3.1, we present a simple mechanism that ignores the edges of the graph and selects as winner a pre-selected node of *maximum expected in-degree*. We call this mechanism the *Fixed Winner mechanism*, and show that it is $\Theta(\sqrt{n \ln n})$-additive (Theorems 4.5 and 4.13).

The AVD mechanism that selects as default the node of highest expected in-degree, can only perform better than the FIXED WINNER mechanism; this is because it returns the default node (which is the same as for the constant mechanism), but only if the maximum in-degree is not unique. This raises the question of what is the additive guarantee of AVD with the *best choice* for the default. We answer this question negatively by showing that regardless of the choice of the default node, AVD is asymptotically as good as the constant mechanism, that is $\Omega(\sqrt{n \ln(n)})$-additive.

Our main and most technically involved result shows that this version of the AVD is $\mathcal{O}(\ln^2 n)$-additive in the a priori popularity model (Theorem 4.6). We complement this result by showing that our analysis is tight, up to a logarithmic factor: even for uniform inputs where all candidates are a priori equally popular, AVD has additive approximation $\Omega(\ln n)$, *for any* choice of the default node (Theorem 4.14).

The analysis of the FIXED WINNER mechanism serves multiple purposes. It illustrates that when prior information is available, a low expected additive approximation is achievable even by simple deterministic mechanisms, and by the simplest statistic of the prior, that is the expected in-degree. This is in sharp contrast to the no-prior case, where deterministic mechanisms have a very poor performance for both additive (see Chapter 3) and multiplicative approximation [2]. Second, the analysis of the Fixed Winner mechanism is quite simple; e.g., the upper bound follows by a simple application of the Hoeffding bound. However, it introduces some of the techniques (such as tail inequalities and reverse Chernoff bounds) that our strongest results in Sections 4.3.2 and 4.4.2 use. Finally, the upper bound on the expected additive approximation of the Fixed Winner mechanism serves as a benchmark of efficiency for all impartial mechanisms with priors.

The analysis of AVD is considerably more involved than the analysis of the Fixed Winner mechanism. Roughly speaking, the important quantity that affects the additive approximation is the difference between the maximum in-degree and the in-degree of the default node when two or more nodes are tied with the highest in-degree, times the probability of this tie. In the a priori popularity model, the in-degree d of node j is a random variable following the binomial probability distribution with parameters n (the number of trials) and p_j (the probability that a trial is successful). Furthermore, the in-degrees of different nodes are independent. Hence, bounding the probability of a tie at the maximum in-degree is related to (but more demanding than) bounding the probability that two out of many independent binomial random variables take the same maximum value.

Unfortunately, even though problems of this kind have been studied in the literature of applied probability and statistics (e.g., see [21, 38, 39]), the existing results have not been proved useful for our purposes. When the difference between the maximum in-degree and the in-degree of the default node is large, Chernoff bounds can unsurprisingly be used to show that the probability of a tie at maximum is negligible and, hence, the contribution to the expectation of the quantity of interest is negligible as well. The real challenge is when the difference of the two in-degrees is small. In this regime, it turns out that we need sharp bounds on the ratio $\Pr[d = x]/\Pr[d \geq x]$ (also called the *hazard function*) for a binomial random variable d and value x that is close to the expectation μ of d (i.e., so that $\Pr[d \geq x]$ is only polynomially small in terms of n). As we show in Lemma 4.12, en route to proving Theorem 4.6, the ratio $\Pr[d = x]/\Pr[d \geq x]$ is at most $\mathcal{O}\left(\sqrt{\frac{\ln n}{\min\{\mu, n-\mu\}}}\right)$ in this case.

We believe that this technical tool can be of independent interest and could find applications elsewhere. The bound is asymptotically tight; its tightness for $p = 1/2$ is exploited in the proof of our logarithmic lower bound (Theorem 4.14). The exact dependence on the quantity $\sqrt{\min\{\mu, n-\mu\}}$ is very important to achieve a polylogarithmic upper bound on the additive approximation of AVD for every a priori popularity input. In addition to the above crucial idea and Chernoff bounds, our proof in Theorem 4.14 involves an inverse Chernoff bound to bound *from below* the probability that a binomial random variable is far from its expectation. These statements are less popular than Chernoff bounds but rather standard.

We should also note that there is an interesting connection of the techniques needed for the analysis in the a priori popularity model with the literature on random graphs [18, 45] and, in particular, results regarding the multiplicity of the highest in-degree in $G_{n,p}$ graphs. Unfortunately, such results have a focus on asymptotics: for example, en route to proving bounds on the chromatic number, Erdős and Wilson [41] showed that the maximum degree is unique with probability $1 - o(1)$ in $G_{n,1/2}$ graphs. Instead, for proving our approximation guarantees, we need sharp estimations of the hidden $o(1)$ term. So, such results are not directly applicable to our analysis.

4.1.2 Roadmap

The rest of the chapter is structured as follows. We begin with preliminary definitions and tail inequality statements in Section 4.2. Section 4.3.1 is devoted to the analysis of the Fixed Winner mechanism in the opinion poll model. Our polylogarithmic additive approximation for AVD is presented in Section 4.3.2 and the logarithmic lower bound in Section 4.4.2.

4.2 Preliminaries

In this chapter we follow the definitions presented in Section 2.2 from Chapter 2. We should note that we make extensive use of the notation $N_{i...j}$ instead of $N \setminus \{i, ..., j\}$ to keep our proofs neat.

Following our work in Chapter 3, we evaluate the performance of a mechanism f on a nomination profile $\mathbf{x}$ using the additive approximation $\Delta(\mathbf{x}) - d_{f(\mathbf{x})}(\mathbf{x})$, i.e., the difference between the maximum in-degree over all nodes and the in-degree of the winner returned by mechanism f.

We assume that the input is a random nomination profile (among the agents of N), selected according to a probability distribution $\mathbf{P}$ over all such profiles. We assume *voter independence*, which means that the distribution $\mathbf{P}$ is a product $\prod_{i \in N} \mathbf{P}_i$ of independent distributions, where $\mathbf{P}_i$ denotes the distribution according to which node i selects its set of outgoing edges.

We examine a hierarchy of three families of distributions, giving raise to *opinion poll*, *a priori popularity*, and *uniform* instances (or models), respectively:

- In the opinion poll model, each node $i \in N$ selects its set of outgoing edges among all possible edges to nodes of N_i, according to the probability distribution $\mathbf{P}_i$. Due to voter independence, the in-degree $d_j(\mathbf{x})$ of each node j is equal to the sum $\sum_{i \in N_j} x_{ij}$ of independent Bernoulli random variables, each denoting whether the directed edge from node i to node j exists in the nomination profile ($x_{ij} = 1$) of not ($x_{ij} = 0$). For simplicity of exposition, in our proofs, we consider N to have $n + 1$ agents; then, the in-degree of each node is the sum of n independent random variables.
- The a priori popularity model is the special case of opinion poll where each node j has a popularity $p_j \in [0, 1]$ and the directed edge (i, j) exists in the nomination profile with probability p_j, independently on all other edges. In this case, the in-degree of node j follows the binomial distribution $\mathcal{B}(n, p_j)$, where n denotes the number of trials and p_j is the success probability for each trial.
- We call uniform the special case of the a priori popularity model with $p_j = p$ for every agent j.

We assume that *prior information* about the underlying probability distributions is known in advance to everyone. Hence, we examine selection mechanisms that are defined using this information and evaluate them in terms of their expected additive approximation

$$\mathbb{E}_{\mathbf{x} \sim \mathbf{P}}[\Delta(\mathbf{x}) - d_{f(\mathbf{x})}(\mathbf{x})].$$

We use the term α-additive to refer to a selection mechanism with expected additive approximation at most α. Our aim is to design deterministic impartial selection mechanisms that have as low as possible expected additive approximation in any distribution from the above classes. Our positive results apply to opinion poll or to a priori popularity distributions; our proofs of negative results use the simplest uniform ones.

4.2.1 Tail Inequalities

We include some tail bounds here that will be very useful later in our analysis.

Lemma 4.1 (Hoeffding [56]). *Let $X_1, X_2, ..., X_n$ be independent random variables so that $\Pr[\, a_j \le X_j \le b_j \,] = 1$. Then, the expectation of the random variable $X = \sum_{j=1}^n X_j$ is $\mathbb{E}[X] = \sum_{j=1}^n \mathbb{E}[X_j]$ and, furthermore, for every $\nu \ge 0$,*

$$\Pr[\, |X - \mathbb{E}[X]| \ge \nu \,] \le 2 \exp\left(-\frac{2\nu^2}{\sum_{j=1}^n (b_j - a_j)^2} \right).$$

Lemma 4.2 (Chernoff bounds). *Let $B \sim \mathcal{B}(n,p)$ and $\mu = np$. Then, the following inequalities hold*

- *Let $x \ge \mu$. If $\mu \ge n/2$,*

$$\Pr[B \ge x] \le \exp\left(-\frac{(x-\mu)^2 n}{2\mu(n-\mu)} \right). \tag{4.1}$$

  If $\mu < n/2$,

$$\Pr[B \ge x] \le \exp\left(-\frac{(x-\mu)^2}{\mu + x} \right), \tag{4.2}$$

  while if $\mu < n/2$ and, furthermore, $x \le 2\mu$,

$$\Pr[B \ge x] \le \exp\left(-\frac{(x-\mu)^2}{3\mu} \right). \tag{4.3}$$

- *Let $x \le \mu$. If $\mu \le n/2$,*

$$\Pr[B \le x] \le \exp\left(-\frac{(\mu-x)^2 n}{2\mu(n-\mu)} \right), \tag{4.4}$$

  while if $\mu > n/2$ and, furthermore, $x \ge 2\mu - n$,

$$\Pr[B \le x] \le \exp\left(-\frac{(\mu-x)^2}{3(n-\mu)} \right). \tag{4.5}$$

Inequalities (4.2),(4.3) and (4.5) are the standard Chernoff bounds; e.g., see [76]. Inequalities (4.1) and (4.4) are due to Okamoto [82]. The following lemma (see [3], Lemma 4.7.2, page 116]) indicates that Chernoff bounds are asymptotically tight.

Lemma 4.3. *Let $B \sim \mathcal{B}(n,p)$ and $\delta \in [0, 1-p)$. Then,*

$$\Pr[B \geq n(p+\delta)] \geq \frac{1}{\sqrt{8n(p+\delta)(1-p-\delta)}} \cdot \left(\left(\frac{p}{p+\delta}\right)^{p+\delta} \left(\frac{1-p}{1-p-\delta}\right)^{1-p-\delta} \right)^n .$$

In particular, we will utilize the following cleaner statement, which follows easily by Lemma 4.3.

Corollary 4.4 (Inverse Chernoff bound)**.** *Let $B \sim \mathcal{B}(n, 1/2)$ and $\delta \in [0, 1/10]$. Then,*

$$\Pr\left[B \geq n\left(\frac{1}{2}+\delta\right)\right] \geq \frac{1}{\sqrt{2n}} \exp\left(-3\delta^2 n\right)$$

Proof. By applying Lemma 4.3 to the random variable B, we have

$$\begin{aligned}
\Pr\left[B \geq n\left(\frac{1}{2}+\delta\right)\right] &\geq \frac{1}{\sqrt{2n}} \left(\left(\frac{1/2}{1/2+\delta}\right)^{1/2+\delta} \left(\frac{1/2}{1/2-\delta}\right)^{1/2-\delta} \right)^n \\
&= \frac{1}{\sqrt{2n}} \left(\frac{1}{\sqrt{1-4\delta^2}} \left(\frac{1/2-\delta}{1/2+\delta}\right)^{\delta} \right)^n \\
&\geq \frac{1}{\sqrt{2n}} \exp\left(\frac{-2\delta^2 - 4\delta^3}{1-2\delta} n\right) \geq \frac{1}{\sqrt{2n}} \exp\left(-3\delta^2 n\right).
\end{aligned}$$

The first inequality follows by Lemma 4.3 and since $(p+\delta)(1-p-\delta)$ is at most $1/4$. The second inequality follows by the inequality $e^z \geq 1+z$ for $z \in \mathbb{R}$ which implies that $\sqrt{1-4\delta^2} \leq \exp(-2\delta^2)$ and $\frac{1/2+\delta}{1/2-\delta} \leq \exp\left(\frac{4\delta}{1-2\delta}\right)$. The third inequality follows since $\delta \leq 1/10$. □

4.3 Upper Bounds

In this section we present deterministic mechanisms with sub-linear additive approximation bounds. In Section 4.3.1 we examine the Opinion Poll model and a we show that a trivial mechanism can yields additive approximation of $O(\sqrt{n \ln n})$. This mechanism shows the power of prior information and works as an warm-up session for our main result in Section 4.3.2. In Section 4.3.2 we analyze the Approval Voting with Default mechanism for the a-priori popularity model and our analysis shows a that this mechanism guarantees a $\mathcal{O}(\ln^2 n)$ for such type of instances.

4.3.1 Opinion Poll Model

We first consider a simple mechanism, which we call the *Fixed Winner* mechanism. This mechanism ignores all edges and awards a particular preselected node, which we call the *default winner* (or default node). The selection of the default winner depends only on the prior. For example, the criterion that we consider here is to select as default winner a node of maximum expected in-degree, i.e.,

$$f_{\mathrm{FW}} \in \operatorname*{argmax}_{v \in N} \mathbb{E}[d_v(\mathbf{x})].$$

Our first statement is an upper bound on the additive approximation of the Fixed Winner mechanism; its proof follows by a simple application of the Hoeffding bound (Lemma 3.5).

Theorem 4.5. *For opinion poll inputs, the* Fixed Winner *mechanism that uses the maximum expected in-degree node as the default winner has expected additive approximation* $\mathcal{O}\left(\sqrt{n \ln n}\right)$.

Proof. Recall that, in the opinion poll model, the in-degree of node v is the sum of n independent Bernoulli random variables, i.e., $d_v(\mathbf{x}) = \sum_{u \in N_v} x_{uv}$. Then, a simple application of the Hoeffding bound (Lemma 3.5) yields

$$\Pr\left[d_v(\mathbf{x}) \geq \mathbb{E}[d_v(\mathbf{x})] + \sqrt{n \ln n}\right] \leq \Pr\left[|d_v(\mathbf{x}) - \mathbb{E}[d_v(\mathbf{x})]| \geq \sqrt{n \ln n}\right] \leq \frac{2}{n^2}.$$

Hence, the probability that some node has in-degree at least $\mathbb{E}[d_{f_{\mathrm{FW}}}(\mathbf{x})] + \sqrt{n \ln n}$ is at most the probability that some node v has in-degree at least $\mathbb{E}[d_v(\mathbf{x})] + \sqrt{n \ln n}$. By the inequality above and the union bound, this probability is at most $\frac{2}{n^2} \cdot (n+1) \leq \frac{3}{n}$. Thus, the expected maximum in-degree is

$$\mathbb{E}[\Delta(\mathbf{x})] \leq \mathbb{E}[d_{f_{\mathrm{FW}}}(\mathbf{x})] + \sqrt{n \ln n} + n \cdot \frac{3}{n} \leq \mathbb{E}[d_{f_{\mathrm{FW}}}(\mathbf{x})] + 3 + \sqrt{n \ln n},$$

and the expected additive approximation $\mathbb{E}[\Delta(\mathbf{x}) - d_{f_{\mathrm{FW}}}(\mathbf{x})]$ is no more than $3 + \sqrt{n \ln n}$. □

4.3.2 A Priori Popularity and the AVD Mechanism

We devote this section to AVD mechanism and its analysis on a priori popularity instances. AVD uses a preselected node t as the default winner. To give a formal definition of the

Algorithm 11: Approval Voting with Default

Input: Nomination Graph $\mathbf{x} = (N, E)$
: default node $t \in N$
Output: Node w in N

$W = N$; ▷ `A set of eligible winners`
foreach $u \in N$ **do**
  if $d_u(N \setminus \{u, t\}) \leq d_t(N \setminus \{u, t\})$ **then**
    $W \leftarrow W \setminus \{t\}$; ▷ `The default node is beaten`
foreach $v \in N \setminus \{t\}$ **do**
  foreach $u \in N \setminus \{v, t\}$ **do**
    if $d_v(N \setminus \{u, t\}) \leq d_u(N \setminus \{v, t\})$ **then**
      $W \leftarrow W \setminus \{v\}$; ▷ `Node` v `is beaten`
if $W \neq \emptyset$ **then**
  return $w \in W$; ▷ `if` W `is not empty includes a single node`
else
  return t

mechanism, we say that a non-default node k *beats* another non-default node j in the nomination profile $\mathbf{x}$ if $d_k(N_{j,k,t}, \mathbf{x}) > d_j(N_{j,k,t}, \mathbf{x})$, i.e., if node k has higher in-degree than node j when ignoring incoming edges from nodes j, k, and the default node t. Node k beats (respectively, is beaten by) the default node t if $d_k(N_{k,t}, \mathbf{x}) > d_t(N_{k,t}, \mathbf{x})$ (respectively, $d_k(N_{k,t}, \mathbf{x}) < d_t(N_{k,t}, \mathbf{x})$). When applied on the nomination profile $\mathbf{x}$, AVD returns as the winner w the node that beats every other node, or the default node if no node that beats every other node exists.[2] We remark that the default node is not prohibited to win when beating every other node. See Algorithm 11 for a possible implementation of the mechanism.

Notice that, by misreporting its outgoing edges, a node cannot affect the set of other nodes it beats. Hence, AVD is clearly impartial. In addition, the above formal definition allows us to observe that the in-degree of the winner returned by AVD is never lower than the in-degree of the default node t. Indeed, when the default node is not the winner, it is beaten by the winner, who has at least as high in-degree. Hence, the upper bound of $\mathcal{O}(\sqrt{n \ln n})$ on the expected additive approximation of the Fixed Winner mechanism on opinion poll instances carries over to AVD mechanism when the default node is selected to

[2] The case in which no node beats every other node refines the notion of a tie that we informally used in Section 4.1.

be a node of highest expected in-degree. In the following, we present a much stronger result that applies specifically to a priori popularity instances.

Theorem 4.6. *The AVD mechanism that uses the node of highest expected in-degree as the default node has an expected additive approximation of* $\mathcal{O}(\ln^2 n)$ *when applied on a priori popularity instances.*

Again, for simplicity of notation, in our analysis of AVD, we consider profiles with $n+1$ nodes. We assume that the number of nodes is large, e.g., $n \geq 10^6$ (otherwise, Theorem 4.6 holds trivially). Let p_k be the popularity of node k and recall that the in-degree $d_k(\mathbf{x})$ of node k is a random variable taking values between 0 and n following the binomial distribution $\mathcal{B}(n, p_k)$, which is also independent on the in-degree of the other nodes. Also, let $\mu_k = \mathbb{E}[d_k(\mathbf{x})] = p_k n$ and $\xi_k = \min\{\mu_k, n - \mu_k\}$.

We first consider the case of $\xi_t < 8200 \ln n$. This means that the expected in-degree of the default node is either very high, i.e., $\mu_t > n - 8200 \ln n$, or very low, i.e., $\mu_t < 8200 \ln n$. When $\mu_t > n - 8200 \ln n$, the expected degree of the winner (be it the default node or not; recall the argument above that compares AVD with the constant mechanism) is more than $n - 8200 \ln n$. As $d_w \leq n$, the expected additive approximation is less than $8200 \ln n$. In the case $\mu_t < 8200 \ln n$, a simple application of a Chernoff bound (i.e., the tail inequality (4.2) from Lemma 4.2) yields that $\Pr[d_k(\mathbf{x}) \geq 9000 \ln n] \leq n^{-37}$. By applying the union bound, the probability that any node has in-degree higher than $9000 \ln n$ is no larger than n^{-36}. Hence, the expected additive approximation is at most $n^{-36} \cdot n + (1 - n^{-36}) \cdot 9000 \ln n \leq 1 + 9000 \ln n$.

So, in the following, we analyze the AVD mechanism assuming that $\xi_t \geq 8200 \ln n$. Let h be the highest in-degree among all nodes. Denote by A the event that there is no node that beats every other node and the default node is beaten by a non-default node of degree h. The following lemma addresses the simplest case where the event A is not true. Recall that w denotes the winner and $d_w(\mathbf{x})$ its in-degree.

Lemma 4.7. $\mathbb{E}[(h - d_w(\mathbf{x}))\, \mathbb{1}\{\overline{A}\}] \leq 1$.

Proof. If the event A does not hold, there must either be a node that beats every other node or the default node is not beaten by any node of degree h.

So, first, assume that there is a node w that beats every other node. The lemma follows if this node has degree h. Otherwise, let i be a node of degree h. Since w beats i, we have $d_w(\mathbf{x}) \geq d_w(N_{i,w,t}, \mathbf{x}) \geq d_i(N_{i,w,t}, \mathbf{x}) + 1 \geq d_i(\mathbf{x}) - 1 = h - 1$ if $w \neq t$, and $d_w(\mathbf{x}) \geq d_t(N_{i,t}, \mathbf{x}) \geq d_i(N_{i,t}, \mathbf{x}) + 1 \geq d_i(\mathbf{x}) = h$ if $w = t$.

Now, assume that the default node is not beaten by any node of in-degree h and there is no node that beats every other node. In this case, the winner will be the default node t. The lemma clearly follows if t has in-degree h. Otherwise, since t is not beaten by some node i of degree h, we have $d_t(\mathbf{x}) \geq d_t(N_{i,t}, \mathbf{x}) \geq d_i(N_{i,t}, \mathbf{x}) \geq d_i(\mathbf{x}) - 1 = h - 1$. □

We will now bound $\mathbb{E}[(h - d_w(\mathbf{x}))\, \mathbb{1}\{A\}]$; to do so, we will use a structural lemma.

Lemma 4.8. *Assume that A is true and let i be a node of highest in-degree h that beats the default node t. Then, there is a node j, different than i and t, that has degree either h, or $h-1$, or $h-2$.*

Proof. Since node i does not beat every other node, there must be some node j that is not beaten by i (clearly, j is different than t). Then, $d_j(\mathbf{x}) \geq d_j(N_{i,j,t}, \mathbf{x}) \geq d_i(N_{i,j,t}, \mathbf{x}) \geq d_i(\mathbf{x}) - 2 = h - 2$. □

By Lemma 4.8, we can bound $\mathbb{E}[(h - d_w(\mathbf{x}))\, \mathbb{1}\{A\}]$ by the expected value of the difference $h - d_t(\mathbf{x})$ for all possible values of the maximum degree h, all possibilities for an agent $i \neq t$ having degree h and an agent $j \neq i, t$ having degree either h or $h-1$ or $h-2$, with the degree of the default node ranging from 0 to h and the degree of all other nodes ranging from 0 to h as well. We have

$$\begin{aligned}\mathbb{E}[(h - d_w(\mathbf{x}))\, \mathbb{1}\{A\}] \leq \sum_{h=0}^{n} \sum_{g=0}^{h} (h-g) \cdot \Pr[d_t(\mathbf{x}) = g] \sum_{i \in N_t} \Pr[d_i(\mathbf{x}) = h] \\ \cdot \sum_{j \in N_{i,t}} \Pr[\max\{0, h-2\} \leq d_j(\mathbf{x}) \leq h] \prod_{k \in N_{i,j,t}} \Pr[d_k(\mathbf{x}) \leq h]\end{aligned} \tag{4.6}$$

For every $k \in \{1, ..., n+1\}$, define the *comfort zone* Z_k of agent k to be the set of integers $\{L_k, ..., U_k\}$ with the boundaries satisfying $L_k \leq \mu_k \leq U_k$ and being defined as follows. The lower boundary L_k is equal to the highest integer c such that $\Pr[d_k(\mathbf{x}) < c] \leq n^{-5.33}$ or 0 if no such c exists. The upper boundary U_k is equal to the lowest integer c such that $\Pr[d_k(\mathbf{x}) > c] \leq n^{-5.33}$ or n if no such c exists. We use the terms "above Z_k" and "below Z_k" to denote the ranges of integers (if any) $\{0, ..., L_k - 1\}$ and $\{U_k + 1, ..., n\}$.

Now, by simple properties of the binomial distribution and the fact that node t has maximum expected in-degree, we observe that if h lies above the comfort zone of agent t, it also lies above the comfort zone of agent i and, hence, $\Pr[d_i(\mathbf{x}) = h] \leq n^{-5.33}$. Also, if g lies below the comfort zone Z_t, it holds $\Pr[d_t(\mathbf{x}) = g] \leq n^{-5.33}$. Furthermore, if $h-2$ lies above the confort zone Z_j, then $\Pr[\max\{0, h-2\} \leq d_j(\mathbf{x}) \leq h] \leq \Pr[d_j(\mathbf{x}) \geq U_j] < n^{-5.33}$ as well. Since, trivially, $h - d_t(\mathbf{x}) \leq n$, the contribution of the at most n^4 terms of the sum in which either h or g does not belong to the comfort zone Z_t or $h-2$ lies above the comfort zone Z_j is at most $n^4 \cdot n \cdot n^{-5.33} < 1$. Hence, equation (4.6) becomes

$$\mathbb{E}[(h - d_w(\mathbf{x}))\, \mathbb{1}\{A\}] \leq 1 + \sum_{h=L_t}^{U_t} \sum_{g=L_t}^{h} (h-g) \cdot \Pr[d_t(\mathbf{x}) = g] \sum_{i \in N_t : h \in Z_i} \Pr[d_i(\mathbf{x}) = h] \\ \cdot \sum_{j \in N_{i,t} : h-2 \in Z_j} \Pr[\max\{0, h-2\} \leq d_j(\mathbf{x}) \leq h] \prod_{k \in N_{i,j,t}} \Pr[d_k(\mathbf{x}) \leq h] \tag{4.7}$$

Our aim in the following is to evaluate each term in the sum at the RHS of (4.7). To do so, we will need three auxiliary technical lemmas. The proofs of the first two follow easily by applying Chernoff bounds. In the first of these lemmas, we bound the comfort zones of any node with $\xi_k \geq 8200 \ln n$. These include, the default node t.

Lemma 4.9. *For any node $k \in N$ such that $\xi_k \geq 8200 \ln n$, for the boundaries of the comfort zone Z_k we have $U_k \leq \mu_k + 4\sqrt{\xi_k \ln n}$ and $L_k \geq \mu_k - 4\sqrt{\xi_k \ln n}$.*

Proof. If $\mu_k \geq n/2$, by applying the tail inequality (4.1) from Lemma 4.2, we get

$$\Pr[d_k(\mathbf{x}) \geq \mu_k + 4\sqrt{\xi_k \ln n}] \leq \exp\left(-\frac{(4\sqrt{\xi_k \ln n})^2 n}{2\xi_k(n-\xi_k)}\right) \leq n^{-8}.$$

If $\mu_k < n/2$, and by the assumption $\xi_k \geq 8200 \ln n$, observe that $4\sqrt{\xi_k \ln n} \leq \mu_k$. Hence, by applying the tail inequality (4.3) from Lemma 4.2, we get

$$\Pr[d_k(\mathbf{x}) \geq \mu_k + 4\sqrt{\xi_k \ln n}] \leq \exp\left(-\frac{(4\sqrt{\xi_k \ln n})^2}{3\mu_k}\right) \leq n^{-5.33}.$$

The bounds on U_k follows by its definition.

Similarly, if $\mu_k \leq n/2$, by applying the tail inequality (4.4) from Lemma 4.2, we get

$$\Pr[d_k(\mathbf{x}) \leq \mu_k - 4\sqrt{\xi_k \ln n}] \leq \exp\left(-\frac{(4\sqrt{\xi_k \ln n})^2 n}{2\xi_k(n-\xi_k)}\right) \leq n^{-8}.$$

If $\mu_k > n/2$, and by our assumption $\xi_k = n - \mu_k \geq 8200 \ln n$, observe that $4\sqrt{\xi_k \ln n} \leq n - \mu_k$, which implies that $\mu_k - 4\sqrt{\xi_k \ln n} \geq 2\mu_k - n$. Hence, by applying the tail inequality (4.5) from Lemma 4.2, we get

$$\Pr[d_k(\mathbf{x}) \leq \mu_k - 4\sqrt{\xi_k \ln n}] \leq \exp\left(-\frac{(4\sqrt{\xi_k \ln n})^2}{3\mu_k}\right) \leq n^{-5.33}.$$

Again, the bound on L_k follows by its definition. □

Now it remains to focus on nodes with very small or very large expected in-degree, i.e. for each node $k \in N$ such that $\xi_k = \min\{\mu_k, n - \mu_k\} < 8200 \ln n$. Recall however, that we are in the case where $\xi_t \geq 8200 \ln n$, and since $\mu_k \leq \mu_t$, it only remains to treat the case $\mu_k < 8200 \ln n$, which we do in the next lemma.

Lemma 4.10. *For any node $k \in N$ such that $\mu_k < 8200 \ln n$, $U_k \leq \mu_k + 4\sqrt{\xi_t \ln n}$.*

Proof. First note that $\mu_k < 8200 \ln n$ implies that $\mu_k < n/2$, for large enough n. By applying the tail inequality (4.2) of Lemma 4.2, and since $\xi_t \geq 8200 \ln n$ we get

$$\Pr[d_k(\mathbf{x}) \geq \mu_k + 4\sqrt{\xi_t \ln n}] \leq \exp\left(-\frac{(4\sqrt{8200} \ln n)^2}{2\mu_k + 4\sqrt{8200} \ln n}\right) < n^{-7.83}.$$

Both inequalities are due to $\mu_k < 8200 \ln n$. □

At this point we have provided upper and lower bounds for the comfort zone of the default node t, and upper bounds for the comfort zone of any other node.

We will say that the comfort zones Z_k and $Z_{k'}$ *almost intersect* if $L_{k'} - U_k \leq 2$ or $L_k - U_{k'} \leq 2$. For example, since $h \in Z_t$ and $h - 2 \in Z_j$, the two comfort zones Z_t and Z_j almost intersect.

Lemma 4.11. *If the comfort zone of the default node t, Z_t, almost intersects with the comfort zone of a node $k \in N_t$, Z_k, then $\mu_t \leq \frac{4}{3}\mu_k$ and $\frac{16}{25}\xi_k \leq \xi_t \leq \frac{25}{16}\xi_k$.*

Proof. Since t is the default node, by definition $\mu_k \leq \mu_t$ and $L_k \leq U_t$. Then, for Z_t and Z_k to almost intersect, it must be that $L_t - U_k \leq 2$.

Due to Lemmas 4.9 and 4.10, and the facts $\xi_k \leq \mu_k \leq \mu_t$ and $\xi_t \leq \mu_t$, observe that $U_k \leq \mu_k + \sqrt{\mu_t \ln n}$, for all ξ_k. Using this, and the fact $\mu_t \geq \xi_t \geq 8200 \ln n$, as well as, Lemma 4.9, on inequality $L_t - U_k \leq 2$ we obtain

$$\mu_t \leq 2 + \mu_k + 4\sqrt{\xi_t \ln n} + 4\sqrt{\mu_t \ln n} \leq 2 + \mu_k + 8\sqrt{\mu_t \ln n} \leq \mu_k \left(\frac{1}{4100} + 1\right) + \frac{8\mu_t}{\sqrt{8200}},$$

which clearly implies that $\mu_t \leq \frac{4}{3}\mu_k$.

Also, notice that for all ξ_k, $U_k \leq \mu_k + \sqrt{\max\{\xi_t, \xi_k\} \ln n}$. By using this, the fact $\xi_t \geq 8200 \ln n$ and Lemma 4.9, when Z_t and Z_k almost intersect, then

$$\begin{aligned}
\max\{\xi_t, \xi_k\} - \min\{\xi_t, \xi_k\} &\leq \mu_t - \mu_k \leq 2 + 4\sqrt{\max\{\xi_t, \xi_k\} \ln n} + 4\sqrt{\xi_t \ln n} \\
&\leq 2 + 8\sqrt{\max\{\xi_k, \xi_t\} \ln n} \\
&\leq \left(\frac{1}{4100} + \frac{8}{\sqrt{8200}}\right) \max\{\xi_t, \xi_k\},
\end{aligned}$$

which implies that $\max\{\xi_t, \xi_k\} \leq \frac{25}{16} \min\{\xi_t, \xi_k\}$ as desired. □

We turn now to prove the most important technical lemma in our analysis.

Lemma 4.12. *Let $\ell \in \{0, 1, 2\}$ and h be such that $h \in Z_t$ and $h - 2 \in Z_k$ for some agent k. Then,*

$$\Pr[d_k(\mathbf{x}) = h - \ell] \leq 792e\sqrt{\frac{\ln n}{\xi_t}} \cdot \Pr[d_k(\mathbf{x}) > h].$$

Proof. By the definition of the binomial distribution, we have

$$\Pr[d_k(\mathbf{x}) = z] = \binom{n}{z} p_k^z (1 - p_k)^{n-z}$$

for every integer z with $0 \leq z \leq n$. Let x be any positive integer such that $1 \leq x - \mu_k \leq 4\sqrt{\xi_k \ln n}$. For every integer $y > x$, we have

$$\begin{aligned}
\frac{\Pr[d_k(\mathbf{x}) = x]}{\Pr[d_k(\mathbf{x}) = y]} &= \frac{\binom{n}{x} p_k^x (1 - p_k)^{n-x}}{\binom{n}{y} p_k^y (1 - p_k)^{n-y}} \\
&= \frac{(x+1) \cdot (x+2) \cdot \ldots \cdot y}{(n-y+1) \cdot (n-y+2) \cdot \ldots \cdot (n-x)} \cdot \frac{(1-p_k)^{y-x}}{p_k^{y-x}} \\
&= \frac{\left(1 + \frac{x-\mu_k+1}{\mu_k}\right) \cdot \left(1 + \frac{x-\mu_k+2}{\mu_k}\right) \cdot \ldots \cdot \left(1 + \frac{y-\mu_k}{\mu_k}\right)}{\left(1 - \frac{y-\mu_k-1}{n-\mu_k}\right) \cdot \left(1 - \frac{y-\mu_k-2}{n-\mu_k}\right) \cdot \ldots \cdot \left(1 - \frac{x-\mu_k}{n-\mu_k}\right)} \\
&\leq \frac{\left(1 + \frac{y-\mu_k}{\mu_k}\right)^{y-x}}{\left(1 - \frac{y-\mu_k-1}{n-\mu_k}\right)^{y-x}} \\
&\leq \exp\left(\frac{(y-\mu_k)(y-x)}{\mu_k} + \frac{(y-\mu_k+1)(y-x)}{n-y+1}\right)
\end{aligned} \tag{4.8}$$

The first inequality follows since $x < y$. In the second inequality, we have used the properties $1 + z \leq e^z$ for $z \in \mathbb{R}$ and, consequently, $\frac{1}{1-z} = 1 + \frac{z}{1-z} \leq \exp(\frac{z}{1-z})$ for $z \neq 1$, which implies that $1 - z \geq \exp(-\frac{z}{1-z})$.

We now use inequality (4.8) to argue that by selecting y such that $y > h - \ell$ and

$$(y - \mu_k + 1)(y - h + \ell) \leq \frac{3}{11}(\xi_t - y + \mu_k), \tag{4.9}$$

we get

$$\frac{\Pr[d_k(\mathbf{x}) = h - \ell]}{\Pr[d_k(\mathbf{x}) = y]} \leq e. \tag{4.10}$$

Recall that Z_k and Z_t almost intersect. Hence, we have $\mu_k \geq 3\mu_t/4$ (by Lemma 4.11), and (4.9) yields

$$\frac{(y - \mu_k)(y - h + \ell)}{\mu_k} \leq \frac{3}{11} \cdot \frac{\xi_t - y + \mu_k}{\mu_k} \leq \frac{4}{11} \cdot \frac{2\mu_t - y}{\mu_t} \leq \frac{8}{11}. \tag{4.11}$$

Furthermore, using again (4.9), and the inequalities $\mu_k \leq \mu_t$ and $\xi_t \leq n - \mu_t$, we get

$$\frac{(y - \mu_k + 1)(y - h + \ell)}{n - y + 1} \leq \frac{3}{11} \cdot \frac{\xi_t - y + \mu_k}{n - y + 1} \leq \frac{3}{11} \cdot \frac{\xi_t - y + \mu_t}{n - y + 1} \leq \frac{3}{11} \cdot \frac{n - y}{n - y + 1} \leq \frac{3}{11}. \tag{4.12}$$

Inequality (4.10) now follows by inequalities (4.8), (4.11), and (4.12).

Solving inequality (4.9), we get that the range of values for y so that (4.10) is true satisfies

$$h - \ell < y \leq \frac{h - \ell + \mu_k - \frac{14}{11} + \sqrt{(h - \ell - \mu_k)^2 + \frac{16}{11}(h - \ell - \mu_k) + \frac{12}{11}\xi_t + \frac{196}{121}}}{2}.$$

Hence, the number of integer values for y so that $y > h - \ell$ and (4.9) holds is at least

$$\begin{aligned} &\frac{h - \ell + \mu_k - \frac{14}{11} + \sqrt{(h - \ell - \mu_k)^2 + (h - \ell - \mu_k) + \frac{12}{11}\xi_t + \frac{196}{121}}}{2} - h + \ell - 2 \\ &= \frac{\sqrt{(h - \ell - \mu_k)^2 + (h - \ell - \mu_k) + \frac{12}{11}\xi_t + \frac{196}{121}} - (h - \ell - \mu_k + \frac{102}{11})}{2} + 2. \end{aligned} \tag{4.13}$$

The derivative of the quantity at the RHS of (4.13) with respect to h is

$$\frac{2(h-\ell-\mu_k)+1}{4\sqrt{(h-\ell-\mu_k)^2+(h-\ell-\mu_k)+\frac{12}{11}\xi_t+\frac{196}{121}}}-\frac{1}{2}<0,$$

i.e., it is decreasing. Since $h-\ell\in Z_k$ and Z_k and Z_t almost intersect, using Lemmas 4.9, 4.10 and 4.11 we have $h-\ell-\mu_k\leq 4\sqrt{\xi_k\ln n}\leq 5\sqrt{\xi_t\ln n}$ when $\xi_k\geq 8200\ln n$ and $h-\ell-\mu_k\leq 4\sqrt{\xi_t\ln n}$ when $\xi_k<8200\ln n$. In any case, $h-\ell-\mu_k\leq 5\sqrt{\xi_t\ln n}$. Hence, we can bound the RHS of (4.13) as follows:

$$\begin{aligned}
&\frac{\sqrt{(h-\ell-\mu_k)^2+(h-\ell-\mu_k)+\frac{12}{11}\xi_t+\frac{196}{121}}-\left(h-\ell-\mu_k+\frac{102}{11}\right)}{2}+2\\
&=\frac{(h-\ell-\mu_k)^2+(h-\ell-\mu_k)+\frac{12}{11}\xi_t+\frac{196}{121}-\left(h-\ell-\mu_k+\frac{102}{11}\right)^2}{2(\sqrt{(h-\ell-\mu_k)^2+(h-\ell-\mu_k)+\frac{12}{11}\xi_t+\frac{196}{121}}+(h-\ell-\mu_k+\frac{102}{11}))}+2\\
&=\frac{12\xi_t-193(h-\ell-\mu_k)-928}{22\left(\sqrt{(h-\ell-\mu_k)^2+(h-\ell-\mu_k)+\frac{12}{11}\xi_t+\frac{196}{121}}+h-\ell-\mu_k+\frac{102}{11}\right)}+2\\
&\geq\frac{12\xi_t-965\sqrt{\xi_t\ln n}-928}{22\left(\sqrt{25\xi_t\ln n+5\sqrt{\xi_t\ln n}+\frac{12}{11}\xi_t+\frac{196}{121}}+5\sqrt{\xi_t\ln n}+\frac{102}{11}\right)}+2\geq\frac{1}{792}\sqrt{\frac{\xi_t}{\ln n}}+2.
\end{aligned}$$

In the second inequality, we have used $965\sqrt{\xi_t\ln n}\leq(10+2/3)\xi_t$ and $928\leq\xi_t$ to bound the numerator by $2\xi_t/3$ (recall that $\xi_t\geq 8200\ln n$), while the parenthesis in the denominator is clearly at most $12\sqrt{\xi_t\ln n}$.

Now, let $\ell\in\{0,1,2\}$ and $r=\left\lceil\frac{1}{792}\sqrt{\frac{\xi_t}{\ln n}}\right\rceil$. By the discussion above, for $x=h-\ell$ we have

$$\Pr[d_k(\mathbf{x})=h-\ell]\leq e\Pr[d_k(\mathbf{x})=y]$$

for $y=h-\ell+1,h-\ell+2,...,h-\ell+r+2$. By summing these inequalities for $y=h+1,...,h+r$, we get

$$r\Pr[d_k(\mathbf{x})=h-\ell]\leq e\sum_{y=h+1}^{r}\Pr[d_k(\mathbf{x})=y]\leq e\Pr[d_k(\mathbf{x})>h]$$

and, equivalently,

$$\Pr[d_k(\mathbf{x}) = h - \ell] \le \frac{e}{r} \Pr[d_k(\mathbf{x}) > h] \le 792e\sqrt{\frac{\ln n}{\xi_t}} \Pr[d_k(\mathbf{x}) > h].$$

The lemma follows. □

We are ready to complete the proof of Theorem 4.6. For $h \in Z_t$, using Lemma 4.9, we have

$$\sum_{h \in Z_t} \sum_{g \in Z_t} (h - g) \Pr[d_t(\mathbf{x}) = g] \le \sum_{h \in Z_t} 8\sqrt{\xi_t \ln n} \sum_{g \in Z_t} \Pr[d_t(\mathbf{x}) = g] \le 64\xi_t \ln n. \tag{4.14}$$

Furthermore, Lemma 4.12 yields

$$\begin{aligned}
&\sum_{i \in N_t} \sum_{j \in N_{i,t}} \Pr[d_i(\mathbf{x}) = h] \Pr[\max\{0, h-2\} \le d_j(\mathbf{x}) \le h] \prod_{k \in N_{i,j,t}} \Pr[d_k(\mathbf{x}) \le h] \\
&\le 3\left(792 \cdot e\sqrt{\frac{\ln n}{\xi_t}}\right)^2 \sum_{i \in N_t} \sum_{j \in N_{i,t}} \Pr[d_i(\mathbf{x}) > h] \Pr[d_j(\mathbf{x}) > h] \prod_{k \in N_{i,j,t}} \Pr[d_k(\mathbf{x}) \le h] \\
&\le 1881792 \cdot e^2 \cdot \frac{\ln n}{\xi_t}. \qquad (4.15)
\end{aligned}$$

since the last double sum is simply the probability that exactly two agents have degree higher than h (and, hence, has value at most 1). Using (4.14) and (4.15), equation (4.7) yields $\mathbb{E}[(h - d_w(\mathbf{x}))\, \mathbb{1}\{A\}] \le 1 + 10^9 \cdot \ln^2 n$ and the proof of Theorem 4.6 is now complete. □

4.4 Lower Bounds

In this section we complement our results with lower bounds. In Section 4.4.1 we show that our analysis for the Fixed Winner is tight. In Section 4.4.2, we show a lower bound for the Approval Voting with Default of $\Omega(\ln n)$, which nearly matches our upper bound.

4.4.1 A Lower Bound for Fixed Winner Mechanism

The bound in Theorem 4.5 is asymptotically tight. The lower bound instances that we use in the proof of the next statement are the simplest ones: uniform instances with $p = 1/2$.

Consequently, it holds for any selection of the default winner. The proof exploits the reverse Chernoff bound (Corollary 4.4).

Theorem 4.13. *The* Fixed Winner *mechanism has expected additive approximation* $\Omega\left(\sqrt{n \ln n}\right)$*, even when applied to uniform inputs.*

Proof. Consider a uniform prior with $p = 1/2$ over $n+1$ nodes, where n is large, e.g., $n \geq 80$. Then, the in-degree of any node u is a random variable following the binomial probability distribution $\mathcal{B}(n, 1/2)$. Let $u^* = f_{\text{FW}}$ be the node returned by the Fixed Winner mechanism; clearly, $\mathbb{E}[d_{u^*}(\mathbf{x})] = n/2$. Denote by $\mathcal{E}$ the event that some node different than u^* has in-degree at least $\frac{n}{2} + \sqrt{\frac{n \ln n}{6}}$. By applying Corollary 4.4 with $\delta = \sqrt{\frac{\ln n}{6n}}$ (the fact that n is large guarantees that $\delta \leq 1/10$) to the random variable $d_u(\mathbf{x})$, we have

$$\Pr\left[d_u(\mathbf{x}) \geq \frac{n}{2} + \sqrt{\frac{n \ln n}{6}}\right] \geq \frac{1}{n\sqrt{2}},$$

for every node $u \neq u^*$ and, hence,

$$\Pr[\mathcal{E}] \geq 1 - \left(1 - \frac{1}{n\sqrt{2}}\right)^n \geq 1 - e^{-1/\sqrt{2}} \geq \sqrt{2} - 1,$$

where the second inequality follows by the inequality $(1 - r/n)^n \leq e^{-r}$ and the third one by the inequality $e^z \geq 1 + z$ (and, thus, $e^{1/\sqrt{2}} \geq 1 + 1/\sqrt{2}$). We now have

$$\mathbb{E}[\Delta(\mathbf{x})] \geq \mathbb{E}[\max_{u \neq u^*} d_u(\mathbf{x})\, \mathbb{1}\{\mathcal{E}\}] + \mathbb{E}[d_{u^*}(\mathbf{x})\, \mathbb{1}\{\overline{\mathcal{E}}\}] \geq \left(\frac{n}{2} + \sqrt{\frac{n \ln n}{6}}\right) \Pr[\mathcal{E}] + \mathbb{E}[d_{u^*}(\mathbf{x})] \Pr[\overline{\mathcal{E}}]$$
$$= \mathbb{E}[d_{u^*}(\mathbf{x})] + \sqrt{\frac{n \ln n}{6}} \cdot \Pr[\mathcal{E}] \geq \mathbb{E}[d_{u^*}(\mathbf{x})] + \frac{1}{6}\sqrt{n \ln n},$$

and the desired lower bound on the expected additive approximation $\mathbb{E}[\Delta(\mathbf{x}) - d_{u^*}(\mathbf{x})]$ follows. □

4.4.2 A Lower Bound for AVD

In this section, we prove the following lower bound for the uniform domain.

Theorem 4.14. *When applied on uniform instances with* $p = 1/2$*, the AVD mechanism has expected additive approximation* $\Omega(\ln n)$*.*

With uniform instances, the in-degree of each node follows the binomial distribution. In the proof of Theorem 4.14, we use the random variables B and B' following the distributions $\mathcal{B}(n, 1/2)$ and $\mathcal{B}(n-1, 1/2)$, respectively. We also assume that n is sufficiently large.

Let U be the lowest integer c such that $\Pr[B > c] \leq \frac{1}{3e^2 n\sqrt{6}}$. Similarly, let L be the lowest integer c such that $\Pr[B > c] < \frac{1}{n\sqrt{2}}$. Consider the following event D:

- The default node has degree at most $n/2$,
- two non-default nodes (called the *potential winners*) have the same in-degree $d \in [L+1, U]$, without counting the edges between them,
- the remaining non-default nodes (called the *losers*) have in-degree at most $d-1$.

Then, AVD returns the default node as a winner and the additive approximation is at least $L - n/2$. We will show that $\mathbb{E}[(L-n/2)\,\mathbb{1}\{D\}]$ is $\Omega(\ln n)$, proving the lemma. In particular, we will use the inequality

$$\mathbb{E}[(L-n/2)\,\mathbb{1}\{D\}] \geq \frac{1}{2}(L-n/2)\sum_{d=L+1}^{U}\binom{n}{2}\Pr[B'=d]^2 \cdot \Pr[B \leq d-1]^{n-2}. \tag{4.16}$$

The RHS of equation (4.16) is the product of the lower bound of the additive approximation $L - n/2$ when event D happens, with $1/2$ which is (a lower bound on) the probability that the default node has degree at most $n/2$, and with the probability $\Pr[B' = d]^2$ that the two potential winners have degree d (ignoring the edges between them) and the probability $\Pr[B \leq d-1]^{n-2}$ that the losers have degree at most $d-1$, for all the $\binom{n}{2}$ selections of the two potential winners.

We will make use of a series of lemmas to bound the several quantities that appear in the RHS of equation (4.16).

Lemma 4.15. $L \geq \frac{n}{2} + \sqrt{\frac{n\ln n}{6}}$.

Proof. By applying Corollary 4.4 to the random variable $B \sim \mathcal{B}(n, 1/2)$ with $\delta = \sqrt{\frac{\ln n}{6n}}$ (observe that $\delta \leq 1/10$ since n is large), we have $\Pr\left[B \geq \frac{n}{2} + \sqrt{\frac{n\ln n}{6}}\right] \geq \frac{1}{n\sqrt{2}}$ for the random variable $B \sim \mathcal{B}(n, 1/2)$. The lemma follows by the definition of L. $\square$

Lemma 4.16. $U \leq \frac{n}{2} + \sqrt{n\ln n}$.

Proof. A simple application of the Chernoff bound (inequality (4.1) from Lemma 4.2) to the binomial random variable $B \sim \mathcal{B}(n, 1/2)$ yields $\Pr\left[B \geq \frac{n}{2} + \sqrt{n \ln n}\right] \leq \frac{1}{n^2} \leq \frac{1}{3e^2\sqrt{6}}$. The lemma then follows by the definition of U. $\square$

Lemma 4.17. *For the random variable $B \sim \mathcal{B}(n, 1/2)$, it holds that*

$$\Pr[B = x] \geq \frac{2L - n}{n} \cdot \Pr[B \geq x],$$

for every integer $x \geq L$.

Proof. Consider integers x, y with $L \leq x \leq y$. By the definition of the binomial distribution $\mathcal{B}(n, 1/2)$, we have

$$\frac{\Pr[B = y]}{\Pr[B = x]} = \frac{\binom{n}{y}}{\binom{n}{x}} = \frac{x!(n-x)!}{y!(n-y)!} \leq \left(\frac{n-x}{x}\right)^{y-x} \leq \left(\frac{n-L}{L}\right)^{y-x}.$$

Hence,

$$\begin{aligned}\Pr[B \geq x] &= \sum_{y=x}^{n} \Pr[B = y] \leq \Pr[B = x] \cdot \sum_{y=x}^{n} \left(\frac{n-L}{L}\right)^{y-x} \\ &\leq \frac{L}{2L - n} \cdot \Pr[B = x] \leq \frac{n}{2L - n} \cdot \Pr[B = x],\end{aligned}$$

and the lemma follows by rearranging. $\square$

The proof of Lemma 4.18 below follows a similar roadmap with the proof of Lemma 4.12 in Section 4.3.2 but is considerably simpler. In the proof, we will use the following claim, which will also be useful later. The proof follows easily by the definition of the binomial distribution.

Claim 4.17.1. *For the random variable $B \sim \mathcal{B}(n, 1/2)$ and integers x and y with $x \leq y$, it holds that*

$$\Pr[B = x] \leq \left(\frac{y}{n-y}\right)^{y-x} \cdot \Pr[B = y].$$

Proof. By the definition of the binomial distribution $\mathcal{B}(n, 1/2)$, we have

$$\frac{\Pr[B = x]}{\Pr[B = y]} = \frac{\binom{n}{x}}{\binom{n}{y}} = \frac{y!(n-x)!}{x!(n-y)!} \leq \left(\frac{y}{n-y}\right)^{y-x}. \qquad \square$$

Lemma 4.18. *For the random variable $B \sim \mathcal{B}(n, 1/2)$, it holds that*

$$\Pr[B = U] \leq \frac{8}{3e\sqrt{6}} \frac{\sqrt{\ln n}}{n^{3/2}}.$$

Proof. By Claim 4.17.1, we have

$$\frac{\Pr[B = U]}{\Pr[B = y]} \leq \left(\frac{y}{n-y}\right)^{y-U} \leq \exp\left(\frac{(y-U)(2y-n)}{n-y}\right) \tag{4.17}$$

for every integer $y > U$. The second inequality follows since $e^z \geq 1 + z$ for $z \in \mathbb{R}$. By selecting y such that

$$\frac{(y-U)(2y-n)}{n-y} \leq 1, \tag{4.18}$$

we get

$$\frac{\Pr[B = U]}{\Pr[B = y]} \leq e. \tag{4.19}$$

Solving inequality (4.18), we get that the range of values for y so that (4.19) is true satisfies

$$U < y \leq \frac{2U + n - 1 + \sqrt{(2U-n)^2 + 6n - 4U + 1}}{4}.$$

Hence, the number of integer values for y so that $y > U$ and (4.18) is satisfied is at least

$$\begin{aligned} &\frac{2U + n - 1 + \sqrt{(2U-n)^2 + 6n - 4U + 1}}{4} - U - 1 \\ &= \frac{\sqrt{(2U-n)^2 + 6n - 4U + 1} - 2U + n - 5}{4}. \end{aligned} \tag{4.20}$$

Now observe that the quantity at the RHS of (4.20) is non-increasing with respect to U

since its derivative

$$\frac{2U-n-1}{2\sqrt{(2U-n)^2+6n-4U+1}}-\frac{1}{2}$$

is non-positive. So, we can bound the RHS of (4.20) from below using the upper bound on U from Lemma 4.16. We get

$$\begin{aligned}
&\frac{\sqrt{(2U-n)^2+6n-4U+1}-2U+n-5}{4}\\
&=\frac{16n-24U-25}{4\left(\sqrt{(2U-n)^2+6n-4U+1}+2U-n+5\right)}\\
&\geq\frac{4n-24\sqrt{n\ln n}-25}{4\left(\sqrt{4n\ln n+4n-4\sqrt{n\ln n}+1}+2\sqrt{n\ln n}-n+5\right)}\\
&\geq\frac{1}{8}\sqrt{\frac{n}{\ln n}}.
\end{aligned}$$

In the last inequality, we have used $24\sqrt{n\ln n}+25\leq n$ to lower-bound the numerator by $3n$ and $5\leq\sqrt{n\ln n}$ to upper-bound the parenthesis in the denominator by $6\sqrt{n\ln n}$.

Now, let $r=\left\lceil\frac{1}{8}\sqrt{\frac{n}{\ln n}}\right\rceil$. Multiplying inequality (4.19) by $1/r$ and summing these inequalities for $y=U+1,...,U+r$, we have

$$\Pr[B=U]\leq\frac{e}{r}\sum_{y=U+1}^{U+r}\Pr[B=y]\leq\frac{e}{r}\cdot\Pr[B>U]\leq\frac{8}{3e\sqrt{6}}\frac{\sqrt{\ln n}}{n^{3/2}},$$

as desired. The last inequality follows by the definition of U. □

Lemma 4.19. $U-L\geq\frac{1}{6}\sqrt{\frac{n}{\ln n}}$.

Proof. Let $B \sim \mathcal{B}(n, 1/2)$. Using the definition of U and L, we have

$$\begin{aligned}
\frac{2}{3n} \leq \frac{1}{n\sqrt{2}} - \frac{1}{3e^2 n\sqrt{6}} &\leq \Pr[L \leq B \leq U] \\
&\leq \sum_{x=L}^{U} \left(\frac{U}{n-U}\right)^{U-x} \cdot \Pr[B = U] \\
&\leq \left(\frac{U}{n-U}\right)^{U-L+1} \cdot \frac{n-U}{2U-n} \cdot \frac{8}{3e\sqrt{6}} \frac{\sqrt{\ln n}}{n^{3/2}} \\
&\leq \exp\left(\frac{2U-n}{n-U}(U-L+1)\right) \cdot \frac{n-L}{2L-n} \cdot \frac{8}{3e\sqrt{6}} \frac{\sqrt{\ln n}}{n^{3/2}} \\
&\leq \exp\left(\frac{2U-n}{n-U}(U-L+1)\right) \cdot \frac{2}{3en}. \qquad (4.21)
\end{aligned}$$

The first inequality is obvious, while the second one uses the definition of U and L (recall that $\Pr[B \geq L] \geq \frac{1}{n\sqrt{2}}$ and $\Pr[B > U] \leq \frac{1}{4e^2 n\sqrt{3}}$). The third inequality follows by Claim 4.17.1. The fourth inequality follows by Lemma 4.18, the fifth one follows since $U \geq L$ and by the definition of U, and the sixth one follows by Lemma 4.15 and the fact $L \geq n/2$.

By Lemma 4.16 and due to the high value of n, $2n - 3U \geq n/3$. Hence, inequality (4.21) implies that

$$U - L \geq \frac{2n-3U}{2U-n} \geq \frac{1}{6}\sqrt{\frac{n}{\ln n}},$$

as desired. ☐

Lemma 4.20. *Let $B \sim \mathcal{B}(n, 1/2)$ and $B' \sim \mathcal{B}(n-1, 1/2)$. For every integer $x \in [L+1, U]$, $\Pr[B' = x] \geq \frac{2}{3}\Pr[B = x]$.*

Proof. By the definition of the binomial distribution, Lemma 4.16, and the facts that $x \leq U$ and that n is large (the last two imply that $x \leq 2n/3$), we have

$$\Pr[B' = x] = \binom{n-1}{x} 2^{-n+1} = 2\frac{n-x}{n}\binom{n}{x} 2^{-n} \geq \frac{2}{3}\Pr[B = x].$$ ☐

We are now ready to bound $\mathbb{E}[(L - n/2)\,\mathbb{1}\{D\}]$ from below. Using equation (4.16), and the lemmas above, we have

$$\begin{aligned}
\mathbb{E}(L - n/2)\,\mathbb{1}\{D\} &\geq \frac{1}{2}\left(L - \frac{n}{2}\right) \cdot \sum_{d=L+1}^{U} \binom{n}{2} \Pr[B' = d]^2 \Pr[B \leq d-1]^{n-2} \\
&\geq \frac{2}{9}\left(L - \frac{n}{2}\right) \cdot \sum_{d=L+1}^{U} \binom{n}{2} \Pr[B = d]^2 \Pr[B \leq d-1]^{n-2} \\
&\geq \frac{8}{9n^2}\left(L - \frac{n}{2}\right)^3 \sum_{d=L+1}^{U} \binom{n}{2} \Pr[B \geq d]^2 \Pr[B \leq d-1]^{n-2} \\
&\geq \frac{8}{9n^2}\left(L - \frac{n}{2}\right)^3 (U - L)\binom{n}{2}\left(\frac{1}{3e^2 n\sqrt{6}}\right)^2 \left(1 - \frac{1}{n\sqrt{2}}\right)^{n-2} \\
&\geq \frac{1}{6561\, e^5 \sqrt{6}} \cdot \ln n.
\end{aligned}$$

The second inequality follows by Lemma 4.20. The third inequality follows by Lemma 4.17. The fourth inequality follows by the definition of L and U. Finally, the fifth inequality follows by Lemma 4.15, Lemma 4.19, and the fact $\left(1 - \frac{1}{n\sqrt{2}}\right)^{n-2} \geq 1/e$. Theorem 4.14 follows. □

4.5 Multiplicative Approximation and Voter Correlation

In the following, we briefly justify two main decisions that we have taken. First, we show that knowing the prior cannot help us improve the approximation ratio of 2 that is best possible for worst-case inputs. This explains why we have completely ignored the study of multiplicative approximations when prior information is available.

We extend the approximation ratio ρ of a mechanism f against a prior $\mathbf{P}$ as follows:

$$\rho = \frac{\mathbb{E}_{\mathbf{x}\sim\mathbf{P}}[\Delta(\mathbf{x})]}{\mathbb{E}_{\mathbf{x}\sim\mathbf{P}}[d_{f(\mathbf{x})}(\mathbf{x})]}$$

Theorem 4.21. *For every $\epsilon > 0$, no impartial selection mechanism has an approximation ratio better than $2 - \epsilon$ against all uniform priors.*

Proof. Consider uniform instances with two nodes u and v of popularity p. Clearly, $\mathbb{E}[\Delta(\mathbf{x})] = 1 - (1-p)^2 = 2p - p^2$. We show that for every impartial mechanism f, it holds $\mathbb{E}[d_{f(\mathbf{x})}(\mathbf{x})] \leq p$. The theorem then follows by taking p to be sufficiently small.

Indeed, consider the profile consisting of the two directed edges between u and v and let q_u and q_v be the probabilities that the winner is node u and node v, respectively. Impartiality means that node u is the winner with probability q_u at the profile consisting only of the directed edge from v to u and node v is the winner at the profile consisting only of the directed edge from u to v with probability q_v. Overall,

$$\mathbb{E}[d_{f(\mathbf{x})}(\mathbf{x})] = (q_u + q_v) \cdot p^2 + q_u \cdot p \cdot (1-p) + q_v \cdot (1-p) \cdot p \leq (q_u + q_v) \cdot p \leq p.$$

Notice that our argument includes randomized mechanisms that may return no winner with positive probability at some profiles. □

Second, we show that our assumption about voter independence is crucial since, otherwise, even our most appealing AVD mechanism has linear additive approximation.

Example 8. Consider the following instance with $8k+2$ nodes partitioned into sets of nodes A and B of $4k$ nodes each and two additional nodes a and b. Node a is approved by no node with probability $1/2$ and all the $4k$ nodes of set A with probability $1/2$ (i.e., there is correlation between the votes in A). Similarly, and independently from the approvals to node a, node b is approved by no node with probability $1/2$ and by all nodes of set B with probability $1/2$. Notice that there is always a tie and hence AVD always selects the default node, which cannot have expected in-degree higher than $2k$. The expected highest in-degree is $4k$ with probability $3/4$ and 0 with probability $1/4$, i.e., an expected highest in-degree of $3k$. Hence, the additive approximation is k, i.e., linear in the number of nodes.

Chapter 5

Truthful Budget Aggregation with Proportionality Guarantees

This chapter is based on joint work with George Christodoulou and Ioannis Caragiannis. The problem is introduced in Section 1.3. Part of this chapter was published in [28].

5.1 Our Results

In this chapter we focus on the problem of truthful budget aggregation, where voters are tasked to split an exogenously given amount of money among various projects. As an illustrative example, consider a city council inquiring the residents on how to divide the upcoming year's budget on education, among a list of publicly funded schools. Each citizen proposes an allocation for the budget and the city council uses a suitable aggregation mechanism to allocate the budget among the schools. This is an example of a participatory budgeting problem, an exciting and relatively new framework engaging citizens in public decision-making, which we have introduced in Section 1.3.

A natural way to aggregate the proposals in the above example is to compute the arithmetic mean for each project and assign exactly that portion of the budget to each school. This method (or variations[1] of it) is used in practice in economics and sports. An example is the score computation for competing athletes by a board of subjective judges.

[1] A usual variation is the *trimmed mean* mechanism, where some of the extreme bids are discarded. This is done to discourage a single voter to heavily influence a particular alternative.

In competitive diving, for example, seven judges [2] propose scores for each dive. Each athlete's score for a dive is computed as the arithmetic mean from the judges' proposals after discarding the top and the bottom two proposed scores [52]. In another example, two well-known financial benchmarks for interest rates, the London Interbank Offered Rate (Libor) and the Euro Interbank Offered Rate (Euribor), are calculated using the arithmetic mean of rates proposed by various banks, excluding a percentage of extreme values. See the discussion in [88, 69] for more on these examples and [86, 87] for some other applications. Assigning the budget proportionally comes with some perks: It can be easily described, it's calculated efficiently, and it scales naturally to any number of projects.

As we have already discussed in the introductory section, allocating the budget proportionally is susceptible to manipulation. Indeed, consider a simple example with two projects and one hundred voters. Fifty voters propose a $50\% - 50\%$ allocation, while the other fifty voters propose a $100\% - 0\%$ allocation. Hence, the proportional allocation is $75\% - 25\%$. Assume now, that one voter changes her $50\% - 50\%$ proposal to $0\% - 100\%$. This turns the aggregated division to $74.5\% - 25.5\%$, a division which is closer to the $50\% - 50\%$ proposal that she prefers. Hence, she may have an incentive to misreport her most preferred allocation to obtain a better outcome, according to her preference.

Truthful mechanisms, i.e. mechanisms nullifying the incentives for strategic manipulation have already been proposed in the literature, for voters with ℓ_1 preferences. We have already present the most important of these mechanisms in Section 1.3. Recall that, under ℓ_1 preferences, a voter has an ideal division in mind and suffers a disutility equal to the ℓ_1 distance from her ideal division.

While truthful mechanisms protect a participatory budgeting process from strategic behavior, they also may lead to results that cannot be considered fair by any reasonable means. Recall the following example from Section 1.3: $2k+1$ voters provide budget proposals over two projects. The first k voters are strong supporters of the first project, and they all propose the whole budget to be spent on that project, while the remaining $k+1$ voters are equally strong supporters of the second project and behave similarly. Then, any utilitarian social welfare maximizing mechanism[3] should assign the whole budget to the second project. To cope with this issue, Freeman et al. [44] proposed the property of *proportionality*. A mechanism is *proportional* if, in any input consisting only of *single-minded* voters (voters

[2] this is the number of judges used for the individual events in the Olympics.

[3] we have seen such mechanisms, which are also truthful, in the introduction, e.g., the Fractional Knapsack Voting and the Utilitarian Moving Phantom mechanism.

who fully assign the budget to a single project), each project receives the proportion of the voters supporting that project.

The INDEPENDENT MARKETS, proposed by Freeman et al. [44] is a proportional mechanism mechanism which belongs to the broader class of truthful mechanisms, the *moving phantom mechanisms*. A moving phantom mechanism for n voters and m projects, returns for each project the median between the voters' proposals for that project and $n+1$ carefully selected *phantom values*. The selection of the phantom values is crucial: it ensures both the strategy-proofness of the mechanism, as well as its ability to return a feasible aggregated division, i.e. that the portions sum up to 1. For example, the INDEPENDENT MARKETS mechanism, places the $n + 1$ phantom values uniformly in the interval $[0, x]$, for some $x \in [0, 1]$, that guarantees feasibility.

While proportionality is a natural fairness property, it is defined only under a limited scope: A proportional mechanism guarantees to provide the proportional division *only when all voters are single-minded*, and provides no guarantee for all other inputs. In this work, we move one step further and we address the question: "How far from the proportional division can the outcome of a truthful mechanism be?"

Building on the work of Freeman et al. [44], we propose a more robust measure, and we extend the notion of proportionality as follows: Given any input of budget proposals, we define the *proportional division* as the coordinate-wise mean of the proposals and then we measure the ℓ_1 distance between the outcome of any mechanism and the proportional division. We call this metric the *ℓ_1-loss*. We say that a mechanism is α-approximate if the maximum ℓ_1-loss, over all preference profiles, is upper bounded by α. Hence, an approximation guarantee of 0 implies that the mechanism always returns the proportional division, while an approximation guarantee of 2 provides no information.

5.1.1 Contribution

We expand the notion of proportionality due to [44] by proposing a quantitative worst-case measure that compares the outcome of a mechanism with the proportional division. We evaluate this measure on truthful mechanisms, focusing on the critical class of moving phantom mechanisms [44]. Our main objective is to design truthful mechanisms with a small α-approximation. We are able to provide optimal mechanisms for the case of two and three projects.

For the case of two projects, we show that the *Uniform Phantom* mechanism from [44] is

1/2-approximate. Then, for the case of three projects, we first examine the Independent Markets mechanism, and we show that this mechanism cannot be better than 0.6862-approximate. We then propose a new, proportional moving phantom mechanism which we call the Piecewise Uniform mechanism, which is $(2/3+\epsilon)$-approximate, for some $\epsilon \leq 10^{-5}$. The analysis of this mechanism is substantially more involved than the case of two projects, and en route to proving the approximation guarantee, we characterize the instances bearing the maximum ℓ_1-loss for any moving phantom mechanism.

We complement our results by showing matching impossibility results: First, we show that no α-approximate truthful mechanism exists for $\alpha < 1/2$, implying that the Uniform Phantom mechanism is the best possible among all truthful mechanisms when the budget is to be divided among two projects. Furthermore, we show that no α-approximate moving phantom mechanism exists for any $\alpha < 1 - 1/m$. This implies that our result for three projects is effectively the best possible within the family of moving phantom mechanisms.

5.1.2 Roadmap

The rest of this chapter is structured as follows. Section 5.2 introduces formally the models and the definitions we are using in this chapter. Section 5.3 is devoted on our positive results. More precisely, Section 5.3.1 is devoted to the analysis of the Uniform Phantom mechanism, an optimal mechanism for two projects, and Section 5.3.2 is devoted to the analysis of the Piecewise Uniform mechanism, an effectively optimal mechanism for three projects. Section 5.4 is devoted to impossibility results. Section 5.4.1 presents a lower bound of 1/2 for any truthful mechanism, Section 5.4.2 is devoted to the $1 - 1/m$ lower bound for any phantom mechanism and Section 5.4.3 is devoted to the 0.6862 lower bound for the Independent Markets mechanism. Section 5.4.4 presents lower bounds for arbitrarily many projects. Finally, in Section 5.5 we present in detail how the Piecewise Uniform belongs in the class of moving phantom mechanisms.

5.2 Preliminaries

Let $N = [n]$ be a set of voters and $M = [m]$ be a set of projects, for $n \in \mathbb{N}_{\geq 2}$ and $m \in \mathbb{N}_{\geq 2}$. Let $\mathcal{D}(m) = \{\mathbf{x} \in [0,1]^m : \sum_{j \in [m]} x_j = 1\}$. This set is also known as the *standard simplex* [20]. We call a *division* among m projects any tuple $\mathbf{x} \in \mathcal{D}(m)$. Let $d(\mathbf{x}, \mathbf{y}) = \sum_{j \in [m]} |x_j - y_j|$ denote the ℓ_1 distance between the divisions $\mathbf{x}$ and $\mathbf{y}$. Voters have structured preferences over budget divisions. Each voter $i \in [n]$ has a most preferred division, her *peak*, $\mathbf{v}_i^*$, and for each division $\mathbf{x}$, she suffers a disutility equal to $d(\mathbf{v}_i^*, \mathbf{x})$, i.e. the ℓ_1 distance between her peak $\mathbf{v}^*$ and $\mathbf{x}$. A related setting to ℓ_1 preferences is proposed by Nehring and Puppe [78]. In this setting voters have the utility function $u_{\mathbf{x}}(\mathbf{q}) = \sum_{j \in [m]} \min(q_j, x_j)$. The tuple $\mathbf{x}$ provides *funding caps* for each project. Hence, each voter has linear preferences over the allocated values for each project, given that this amount is lower than the cap x_j. The special case where each voter i has a funding cap equal to the peak v^*, is equivalent to ℓ_1 preferences.

Each voter $i \in [n]$ reports a division $\mathbf{v}_i$. These divisions form a *preference profile* $\mathbf{V} = (\mathbf{v}_i)_{i \in [n]}$. We use the notation $\mathbf{V}_{-i}$ to denote the reports of all voters par $i \in [n]$. A *budget aggregation mechanism* $f : (\mathcal{D}(m))^n \to \mathcal{D}(m)$ uses the proposed divisions to decide an aggregate division $f(\mathbf{V})$. A mechanism f is *continuous* when the function $f : (\mathcal{D}(m))^n \to \mathcal{D}(m)$ is continuous. A mechanism is *anonymous* if the output is independent of any voters' permutation and *neutral* if any permutation of the projects (in the voters' proposals) permutes the outcome accordingly.

In this thesis, we focus on *truthful* mechanisms, i.e. mechanisms where no voter can alter the aggregated division to her favor, by misreporting her preference.

Definition 9. (Truthful Budget Aggregation Mechanism) A budget aggregation mechanism f is truthful if, for all preference profiles $\mathbf{V}$, voters i, and divisions $\mathbf{v}_i^*$ and $\mathbf{v}_i$, $d(f(\mathbf{V}_{-i}, \mathbf{v}_i)) \geq d(f(\mathbf{V}_{-i}, \mathbf{v}_i^*))$.

A large part of this work is concerned with the class of *Moving Phantom Mechanisms*, proposed in [44].

Definition 10 (Moving Phantom Mechanism)**.** Let $\mathcal{Y} = \{y_k : k \in \{0..n\}\}$ be a family of functions such that, for every $k \in \{0..n\}$, $y_k : [0,1] \to [0,1]$ is a continuous, weakly increasing function with $y_k(0) = 0$ and $y_k(1) = 1$. In addition, $y_0(t) \leq y_1(t) \leq ... \leq y_n(t)$ for every $t \in [0,1]$. The set $\mathcal{Y}$ is called a *phantom system*. For any valid phantom system, a *moving phantom mechanism* $f^{\mathcal{Y}}$, is defined as follows: For any profile $\mathbf{V}$ and any project $j \in [m]$,

$$f_j^{\mathcal{Y}}(\mathbf{V}) = \text{med}\left(\mathbf{V}_{i \in [n], j}, (y_k(t^*))_{k \in \{0..n\}}\right) \tag{5.1}$$

for some

$$t^* \in \left\{ t : \sum_{j\in[m]} \operatorname{med}\left(\mathbf{V}_{i\in[n],j}, (y_k(t))_{k\in\{0..n\}}\right) = 1 \right\}.$$

The function $y_k(t)$ is used to derive the value of the k-th phantom. For each $t \in [0,1]$, the phantom system returns the tuple $(y_0(t), ..., y_n(t))$ with $n+1$ phantom values. We denote with t^*, any value in $[0,1]$ where the tuple of phantom values is sufficient for the sum of the coordinate-wise medians to be equal to 1. Freeman et al. in [44] show that one such t^* always exists and, in case of multiple candidate values, the specific choice of t^* does not affect the outcome. When the phantom system is clear from the context, we use f instead of $f^{\mathcal{Y}}$.

The following theorem from [44] is crucial for this work.

Theorem 5.1. *[44] Every moving phantom mechanism is truthful and always returns feasible divisions.*

Each median in Definition 10 can be computed using a sorted array with $2n+1$ slots, numbered from 1 to $2n+1$. The median value is located in slot $n+1$. Throughout this work we refer to the slots 1 to n as the *lower* slots, and $n+2$ to $2n+1$ as the *upper* slots.

In Algorithm 12 we present a generic procedure for the computation of moving phantom mechanisms. Each moving phantom mechanism is characterized by its phantom system. Once a valid phantom system is defined, the procedure becomes an algorithm for computing the precise moving phantom mechanism. Freeman et al. [44] show that when the phantom functions are *piecewise-linear*, the normalization parameter t^* is rational and can be represented with polynomially many bits, hence the binary search procedure admits polynomial computational time in the size of the input. All the mechanisms we are considering in this thesis use piecewise-linear phantom functions.

In Chapter 1 we have introduced two moving phantom mechanisms. The first one is the Independent Markets mechanism which uses the phantom system

$$\mathcal{Y}^{\mathbf{IM}} = \{y_k(t) = \min(k \cdot t, 1) : k \in \{0..n\}\}. \tag{5.2}$$

The second is the Utilitarian Moving Phantom mechanism, which uses the phantom system $\mathcal{Y}^*$, which is described by the following phantom functions:

$$y_k(t) = \begin{cases} 0 & 0 \le t \le \frac{k}{n+1}, \\ t(n+1) - k & \frac{k}{n+1} < t < \frac{k+1}{n+1}, \\ 1 & \frac{k+1}{n+1} \le t \le 1, \end{cases} \tag{5.3}$$

for all $k \in [n]$. Both these mechanisms are due to [44]. Two examples of these phantom systems are presented in Figure 1.13.

For a given preference profile $\mathbf{V}$, let

$$\bar{\mathbf{V}} = \left(\frac{1}{n} \sum_{i \in [n]} v_{i,j} \right)_{j \in [m]}$$

be the *proportional division* of $\mathbf{V}$. A *single-minded* voter is a voter $i \in [n]$ such that $v_{i,j} = 1$ for some project $j \in [m]$. A budget aggregation mechanism is called *proportional*, if for any preference profile $\mathbf{V}$ consisted solely of single-minded voters, it holds $f(\mathbf{V}) = \bar{\mathbf{V}}$.

For a given budget aggregation mechanism f and a preference profile $\mathbf{V}$, we define the ℓ_1-loss as the ℓ_1 distance between the outcome $f(\mathbf{V})$ and the proportional division $\bar{\mathbf{V}}$, i.e.

$$\ell(\mathbf{V}) = d(f(\mathbf{V}), \bar{\mathbf{V}}) = \sum_{j \in [m]} \left| f_j(\mathbf{V}) - \bar{\mathbf{V}}_j \right|. \tag{5.4}$$

We say that a budget aggregation mechanism is α-approximate when the ℓ_1-loss for any preference profile is no larger than α. We note that no mechanism can be more than 2-approximate, as the ℓ_1 distance between any two arbitrary divisions is at most 2.

5.3 Upper Bounds

In this section we present mechanisms with small approximation guarantees for $m = 2$ and $m = 3$. As we will see later on, the upper bound for the case of 2 projects is optimal among any truthful mechanism, while the upper bound for 3 projects is practically optimal among all moving phantom mechanisms.

Algorithm 12: Generic Moving Phantom

Input: collection of preference profiles $\mathbf{V} \in (\mathcal{D}(m))^n$
Output: division $\mathbf{f} = (f_1, ..., f_m)$
Parameters: number of voters n, number of projects m, $n+1$ functions $(y_0(t), ..., y_n(t))$, error tolerance $\epsilon \geq 0$

▷ `Binary search over` $t \in [0, 1]$
$t^* \leftarrow 1$;
$\delta \leftarrow 1$;
$k \leftarrow 1$;
while $|\delta| > \epsilon$ **do**
  for $j \leftarrow 1$ **to** m **do**
    $f_j = \text{med}\left((\mathbf{V}_{i,j})_{i\in[n]}, (p_k(t^*))_{k\in\{0,\ldots,n\}}\right)$;
  $s \leftarrow 1 - \sum_{j=1}^{m} f_j$;
  if $\delta > t^* + \epsilon$ **then**
    $t^* \leftarrow t^* - 2^{-k}$
  else if $\delta < t^* - \epsilon$ **then**
    $t^* \leftarrow t^* + 2^{-k}$
  $k \leftarrow k + 1$;
return $\mathbf{f}$;

5.3.1 Two Projects

For the case of two projects, we focus on the Uniform Phantom mechanism, as defined in [44], for which we show a 1/2-approximation. Freeman et al. [44] have already shown that this mechanism is the unique truthful and proportional mechanism for $m = 2$. Here, we show that this mechanism yields the best possible ℓ_1-loss for $m = 2$.

The Uniform Phantom mechanism returns the median between $n+1$ phantom values, uniformly distributed over the $[0, 1]$ line, and the voters' reports for each project, i.e.,

$$f_j = \text{med}\left(\mathbf{V}_{i\in[n],j}, (k/n)_{k\in\{0..n\}}\right),$$

for $j \in \{1, 2\}$.

Later, in Theorem 5.9, we show that 1/2 is the best approximation we can achieve by any truthful mechanism.

Theorem 5.2. *For $m = 2$, the* Uniform Phantom *mechanism is* 1/2*-approximate.*

Proof. Let f be the Uniform Phantom mechanism, and let $\mathbf{V}$ be a preference profile. Let

$f(\mathbf{V}) = (x, 1-x)$ and $\bar{\mathbf{V}} = (\bar{v}, 1-\bar{v})$. The loss of the mechanism for $\mathbf{V}$ is

$$\ell(\mathbf{V}) = 2\left|x - \bar{v}\right|. \tag{5.5}$$

Let $k \in \{0..n\}$ be the minimum phantom index such that $x \leq \frac{k}{n}$. This implies that the phantoms with indices $k, ..., n$ are located in the slots $n+1$ to $2n+1$. These phantoms are exactly $n+1-k$ i.e. exactly k voters' reports are located in the same area. Since all values in these slots are at least equal to the median we get that

$$\begin{aligned} \frac{k}{n} \cdot x \leq \bar{v} \leq \frac{n-k}{n} \cdot x + \frac{k-1}{n} + \frac{1}{n} \cdot \mathbb{1}\left\{x = k/n\right\} \\ + \frac{x}{n} \cdot \mathbb{1}\left\{x < k/n\right\} \end{aligned} \tag{5.6}$$

The first inequality holds, since exactly k voters' reports have value at least equal to the median x. For the second inequality, we note that exactly $n-k$ voters' reports have value at most x, while at least $k-1$ voters' reports can have value at most 1. If the median is equal to k/n, we can safely assume that this is a phantom value, and there should be exactly k values upper bounded by 1. Otherwise, if the median is strictly smaller than k/n, then x should be a voter's report and exactly $k-1$ voters' reports are located in the upper slots.

By removing x from both inequalities in 5.6 we get:

$$\begin{aligned} \frac{k}{n} \cdot x - x \leq \bar{v} - x \leq \frac{k-1}{n} + \frac{1}{n} \cdot \mathbb{1}\left\{x = k/n\right\} \\ + \frac{x}{n} \cdot \mathbb{1}\left\{x < k/n\right\} - \frac{k}{n} \cdot x. \end{aligned} \tag{5.7}$$

When the median is a phantom value, i.e. $x = \frac{k}{n}$, inequalities 5.7 imply that

$$\left|\bar{v} - x\right| \leq \max\left\{x\left(1 - \frac{k}{n}\right), \frac{k}{n}\left(1 - x\right)\right\} = \frac{k}{n}\left(1 - \frac{k}{n}\right),$$

which is maximized to for $k = n/2$ to a value no greater than $1/4$. When x is a voter's report, i.e. $\frac{k-1}{n} < x < \frac{k}{n}$, inequalities 5.7 imply that

$$
\begin{aligned}
|\bar{v} - x| &\leq \max\left\{x\left(1-\frac{k}{n}\right), \frac{k-1}{n}(1-x)\right\} \\
&< \max\left\{\frac{k}{n}\left(1-\frac{k}{n}\right), \frac{k-1}{n}\left(1-\frac{k-1}{x}\right)\right\}. \quad (5.8)
\end{aligned}
$$

Both quantities in the maximum operator are upper bounded by 1/4. The theorem follows. □

5.3.2 Three Projects

We now turn our attention to cases where the budget should be split over three projects, and we provide a $(2/3+\epsilon)$-approximate truthful mechanism for some $\epsilon \leq 10^{-5}$. This mechanism belongs to the family of moving phantom mechanisms, and it is also proportional. In the following, we describe the mechanism, and then prove the approximation guarantee. Later, in Theorem 5.10, we show that 2/3 is the best possible guarantee among the class of moving phantom mechanisms.

The Piecewise Uniform mechanism

The Piecewise Uniform mechanism uses the phantom system $\mathcal{Y}^{\text{PU}} = \{y_k(t) : k \in \{0..n\}\}$, for which

$$
y_k(t) = \begin{cases} 0 & \frac{k}{n} < \frac{1}{2} \\ \frac{4tk}{n} - 2\,t & \frac{k}{n} \geq \frac{1}{2} \end{cases} \quad (5.9)
$$

for $t < 1/2$, while

$$
y_k(t) = \begin{cases} \frac{k(2\,t-1)}{n} & \frac{k}{n} < \frac{1}{2} \\ \frac{k(3-2\,t)}{n} - 2 + 2\,t & \frac{k}{n} \geq \frac{1}{2} \end{cases} \quad (5.10)
$$

for $t \geq 1/2$. This mechanism belongs to the family of moving phantom mechanisms[4]:

[4]Note that this mechanism does not entirely fit the Definition 10 since $y_k(t) < 1$, for all $k \in \{0..n-1\}$. This can fixed easily however, with an alternative definition, where all phantom functions are shifted slightly

each $y_k(t)$ is a continuous, weakly increasing function, and $y_k(t) \geq y_{k-1}(\mathrm{t})$ for $k \in \{0..n-1\}$ and any $t \in [0,1]$. We call a phantom with index $k < n/2$ a *black* phantom, and a phantom with index $k \geq n/2$, a *red* phantom.

(a) (b)

Figure 5.1: Examples of the PIECEWISE UNIFORM mechanism, with 5 voters. The dashed lines correspond to the phantom values, the small rectangles correspond to the medians, and the thick lines correspond to the voters' reports. In the first example $\mathbf{v}_1 = \mathbf{v}_2 = (3/8, 3/8, 1/4)$, $\mathbf{v}_3 = (1/8, 1/2, 3/8)$, $\mathbf{v}_4 = (7/16, 9/16, 0)$ and $\mathbf{v}_5 = (5/8, 1/16, 5/16)$ and $t^* = 3/8$, while the final outcome is $(3/8, 3/8, 1/4)$. In the second example, $\mathbf{v}_1 = (1, 0, 0)$, $\mathbf{v}_2 = (1/2, 1/2, 0)$, $\mathbf{v}_3 = (0, 2/3, 1/3)$, $\mathbf{v}_4 = (1/3, 5/9, 1/9)$, $\mathbf{v}_5 = (3/8, 3/8, 1/4)$ and $t^* = 49/64$.

This mechanism can been seen as a combination of two different mechanisms: For $t < 1/2$, the mechanism uses $n/2$ phantom values equal to 0, and the rest are uniformly located in $[0, y_n(t)]$. For $t \geq 1/2$, the mechanism assigns half of the phantoms uniformly in $[0, y_{\lfloor n/2 \rfloor}(t)]$, while the rest are uniformly distributed in $[y_{\lceil n/2 \rceil}, 1]$. See the examples of Figure 5.1, for an illustration.

We emphasize here that the PIECEWISE UNIFORM mechanism admits polynomial time-computation using a binary search algorithm, since $\mathcal{Y}^{\mathrm{PU}}$ is a *piecewise linear phantom system* (see Theorem 4.7 from [44]; see also Algorithm 12 for possible implementation).

We continue by showing that this mechanism is proportional. Note that this is not necessarily needs to hold, to show the desired approximation guarantee, but it is a nice extra feature of our mechanism.

Theorem 5.3. *The* PIECEWICE UNIFORM *mechanism is proportional.*

Proof. Consider any preference profile which consists exclusively of single-minded voters. Note that by using $t = 1$, the phantom with index k has the value k/n, for any $k \in \{0..n\}$. Let that $a_j \in \{0..n\}$ be the number of 1-valued proposals on project j. Consequently,

to the left and a third set of linear functions are added, such that and $y_k(1) = 1$ for all $k \in \{0..n\}$. See Section 5.5 for a detailed explanation.

$n - a_j$ is the number of 0-valued proposals. Then the median in each project is exactly the phantom value a_j/n, i.e. the proportional allocation. □

The Upper Bound

Overview The analysis for the upper bound is substantially more involved than the analysis for the case of two projects. Here we present an outline of our technique.

We first provide a characterization of the worst-case preference profiles (i.e. profiles that may yield the maximum loss) in Theorem 5.5. This characterization states that essentially all worst-case preference profiles belong to a specific family, which we call *three-type* profiles (see Definition 11). The family of three-type profiles depends crucially on the moving phantom mechanism used. Given a moving phantom mechanism, Lemma 5.6 characterizes further the family of three-type for that mechanism.

We combine Theorem 5.5 and Lemma 5.6 to build a Non-Linear Program (NLP; see Figure 5.5) which explores the space created by the worst-case instances. Finally we present the optimal solution of the NLP in Theorem 5.8.

Characterization of Worst-Case Instances We concentrate on a family of preference profiles which are maximal (with respect to the loss) in a local sense: A preference profile $\mathbf{V}$ is *locally maximal* if, for all voters $i \in [n]$, it holds that $\ell(\mathbf{V}) \geq \ell(\mathbf{V}_{-i}, \mathbf{v}'_i)$ for any division $\mathbf{v}'_i$. In other words, in such profiles, any single change in the voting divisions cannot increase the ℓ_1-loss. Inevitably, any profile which may yield the maximum loss, belongs to this family, and we can focus our analysis on such profiles. Our characterization shows that the class of locally maximal preference profiles and the class of three-type profiles are equivalent with respect to ℓ_1-loss, for any phantom mechanism.

Definition 11 (three-type profiles)**.** For any phantom mechanism f, a preference profile $\mathbf{V}$ is called a *three-type* profile if every voter $i \in [n]$ belongs to one of the following classes:

1. *fully-satisfied* voters, where voter i proposes a division equal to the outcome of the mechanism, i.e. $f(\mathbf{V}) = \mathbf{v}_i$,

2. *double-minded* voters, where voter i agrees with the outcome in one project, i.e. $v_{i,j} = f_j(\mathbf{V})$ for some $j \in [3]$, while $v_{i,j'} = 1 - f_j(\mathbf{V})$ for some different project j', and

3. *single-minded* voters, where $v_{i,j} = 1$ for some project $j \in [3]$.

To build intuition, we provide the following example:

Example 9 (three-type profile)**.** Consider a phantom mechanism f, and the preference profile $\mathbf{V}$ with 5 voters: $\mathbf{v}_1 = (1, 0, 0)$ and $\mathbf{v}_2 = (0, 0, 1)$, which are single-minded voters, $\mathbf{v}_3 = (1/2, 1/2, 0)$, $\mathbf{v}_4 = (0, 1/4, 3/4)$ and $\mathbf{v}_5 = (1/2, 1/4, 1/4)$. Then, if $f(\mathbf{V}) = \mathbf{v}_5$, the preference profile $\mathbf{V}$ is a three-type profile for mechanism f. Voter 5 is a fully-satisfied voter, while voters 3 and 4 are double-minded voters.

In Theorem 5.5 that follows, we show that for any locally maximal preference profile $\mathbf{V}$, there exists a three-type profile $\hat{\mathbf{V}}$ (not necessarily different than $\mathbf{V}$) for which $\ell(\hat{\mathbf{V}}) \geq \ell(\mathbf{V})$. Therefore, we can search for the maximum ℓ_1-loss by focusing only on the profiles described in Definition 11. The following lemma is an important stepping stone for the proof of Theorem 5.5.

Lemma 5.4. *Let f be a phantom mechanism for $m = 3$, $\mathbf{V}$ a preference profile and $i \in [n]$, a voter which is neither single-minded, double-minded nor fully-satisfied. Let $\mathbf{v}_i$ be voter's i proposal. Then there exists a division $\mathbf{v}'_i$ such that $\ell(\mathbf{V}_{-i}, \mathbf{v}'_i) \geq \ell(\mathbf{V})$. Furthermore, when $\ell(\mathbf{V}_{-i}, \mathbf{v}'_i) = \ell(\mathbf{V})$ the division $\mathbf{v}'_i$ is double-minded, single-minded or fully-satisfied, and $f(\mathbf{V}) = f(\mathbf{V}_{-i}, \mathbf{v}'_i)$.*

Proof. Let $\bar{v}_j = \sum_{i=1}^{n} v_{i,j}$ for $j \in [3]$. For convenience we also write $\mathbf{f} = (f_1, f_2, f_3)$ instead of $f(\mathbf{V})$. We will prove the lemma by constructing the division $\mathbf{v}'_i$. For that, we alter the proposals only on two projects, and we keep the proposal for the third project invariant, in such a way that $\mathbf{v}'_i$ is a valid division. Our first attempt is to strictly increase the loss. When we fail to do that, we create $\mathbf{v}'_i$ in such a way that the loss is not decreasing.

The following claim, allows us to focus our analysis only on two cases.

Claim 5.4.1. *Let that $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be valid divisions over m projects, and let that $\mathbf{x} \neq \mathbf{y}$. Then there exists a pair of projects such that either i) $x_j \leq \min\{y_j, z_j\}$ and $x_{j'} \geq \max\{y_{j'}, z_{j'}\}$ or ii) $y_j \leq x_j \leq z_j$ and $z_{j'} \leq x_{j'} \leq y_{j'}$.*

Proof. We firstly notice that there exists two projects, say 1 and 2, such that $x_1 < y_1$ and $y_2 < x_2$, since $\mathbf{x} \neq \mathbf{y}$. There exist four possible relations between x_1, x_2 and z_1, z_2: If $x_1 \leq z_1$ and $z_2 \leq x_2$ then condition (i) is satisfied. Similarly, if $z_1 \leq x_1$ and $x_2 \leq z_2$, condition (ii) is satisfied. If $x_1 \leq z_1$ and $x_2 \leq z_2$, we notice that there exists a different project, say project 3, such that $z_3 \leq x_3$; otherwise $\sum_{j \in [m]} z_j > 1$. We have now two cases, according to the relation between y_3 and x_3: when $y_3 \leq x_3$, condition (i) holds between projects 1 and 3;

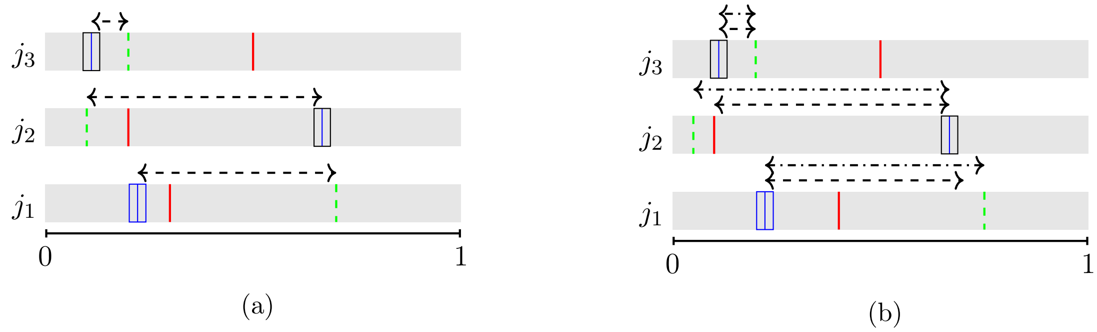


Figure 5.2: An example where the loss can be increased by a single change in voter's i proposed division. Note that voter i is neither fully-satisfied, single-minded nor double-minded. The voter's proposals are depicted with solid vertical lines, the mean with dashed vertical lines and the outcome of the mechanism is depicted with vertical solid lines inside a rectangle. By moving the proposals of voter i away from the median, the loss strictly increases.

when $x_3 \leq y_3$, condition (ii) holds between projects 2 and 3. Similar arguments holds for the case $z_1 \leq x_1$ and $z_2 \leq x_2$. $\square$

By Claim 5.4.1, we can assume without loss of generality that either i) $f_1 \leq \min\{v_{i,1}, \bar{v}_1\}$ and $f_2 \geq \max\{v_{i,2}, \bar{v}_2\}$ or ii) $v_{i,1} \leq f_1 \leq \bar{v}_1$ and $v_{i,2} \leq f_2 \leq \bar{v}_2$.

Case (i): If $v_{i,2} > 0$, we can increase the loss as follows: we move $v_{i,1}$ to $v'_{i,1} = v_{i,1} + \epsilon$ and $v_{i,2}$ to $v'_{i,2} = v_{i,2} - \epsilon$, for some $0 < \epsilon \leq \min\{v_{i,2}, 1 - v_{i,1}\}$. This increases $\bar{v}_1$ to $\bar{v}'_1 = \bar{v}_j + \epsilon/n$ and decreases $\bar{v}_2$ to $\bar{v}'_2 = \bar{v}_2 - \epsilon/n$. Note that these moves do not affect the outcome of f, since no voters' reports move from the lower to the upper slots or vice versa. Also, $\mathbf{v}'_{i,3} = \mathbf{v}_{i,3}$, hence $\bar{v}'_3 = \bar{v}_3$. Thus $\ell(\mathbf{V}_{-i}, \mathbf{v}'_i) > \ell(\mathbf{V})$. See also Figure 5.2.

If $v_{i,2} = 0$, i can be transformed to a single or double-minded voter, without decreasing the loss. Note that $v_{i,1} = 1 - v_{i,3}$. When $v_{i,3} \leq f_3$, $v_{i,3}$ moves to $v'_{i,3} = 0$ and $v_{i,1}$ moves to $v'_{i,1} = 1$ to create a single-minded division. When $v_{i,3} > f_3$, we can move $v_{i,3}$ to $v'_{i,3} = f_3$ and $v_{i,1}$ to $v'_{i,1} = 1 - f_3$ (note that $1 - v_{i,3} < 1 - f_3$) to create a double-minded division $\mathbf{v}'_i$. In any case, $v'_{i,1} = v_{i,1} + \epsilon$ and $v'_{i,3} = v_{i,3} - \epsilon$, where $\epsilon = v_{i,3} \cdot \mathbb{1}\{v_{i,3} \leq f_3\} + (v_{i,3} - f_3) \cdot \mathbb{1}\{v_{i,3} > f_3\}$. Also, $f(\mathbf{V}_{-i,}, \mathbf{v}'_i) = f(\mathbf{V})$ and $\bar{v}'_2 = \bar{v}_2$. Thus,

$$\begin{aligned}\ell(\mathbf{V}_{-i,}\mathbf{v}'_i) &= \bar{v}_1 + \epsilon - f_1 + f_2 - \bar{v}_2 + |\bar{v}_3 - \epsilon - f_3| \\ &\geq \ell(\mathbf{V}).\end{aligned}$$

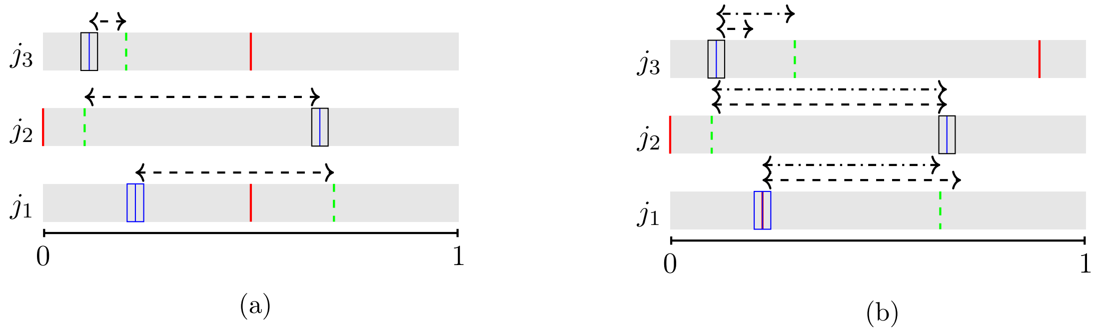


Figure 5.3: An example where the loss cannot be increased by a single change in voter's i proposed division, without changing the outcome of the mechanism. Note that voter i is neither fully-satisfied, single-minded nor double-minded. The voters proposals are depicted with solid vertical lines, the mean with dashed vertical lines and the outcome of the mechanism is depicted with vertical solid lines inside a rectangle. We tranform the voter to a double minded voter and the loss is preserved.

The inequality holds due to $|x| \geq x$ for $x \in \mathbb{R}$. See also Figure 5.3.

Case (ii): Recall that $\bar{v}_1 \leq f_1 \leq v_{i,1}$ and and $v_{i,2} \leq f_2 \leq \bar{v}_2$. If there exists some $0 < \epsilon \leq \min\{v_{i,1} - f_1, f_2 - v_{i,2}\}$ we can strictly increase the loss by $2\epsilon/n$ by moving $v_{i,1}$ to $v'_{i,j} - \epsilon$ and $v'_{i,2} = v_{i,2} + \epsilon$. When no such ϵ exists, either $f_1 = v_{i,1}$ or $f_2 = v_{i,2}$ (note that this cannot happen for both projects; this would lead to $\mathbf{v}_i = \mathbf{f}$). Assume that $v_{i,1} = f_1$. We firstly note that for project 3, $v_{i,3} \geq f_3$; otherwise, $\sum_{j\in[3]} v_{i,j} < 1$. We also note that $v_{i,2} < f_2$ and $v_{i,3} > f_3$, otherwise $\mathbf{v}_i = \mathbf{f}$. We will transform $\mathbf{v}_i$ to a fully-satisfied voter, proposing $\mathbf{v}'_i$, without decreasing the loss. This is done by increasing $v_{i,2}$ to $v'_{i,2} = f_2$ and decreasing $v'_{i,3}$ to $v'_{i,3} = f_3$. Hence $\bar{v}'_2 = \bar{v}_2 + \frac{f_2 - v_{i,2}}{n}$ and $\bar{v}'_3 = \bar{v}_3 - \frac{v_{i,3} - f_3}{n} = \bar{v}_3 - \frac{f_3 - v_{i,2}}{n}$ (recall that $v_{i,1} = f1$). Note that $f(\mathbf{V}_{-i,}, \mathbf{v}'_i) = f(\mathbf{V})$ and $\bar{v}'_1 = \bar{v}_1$. Let $\epsilon = \bar{v}_2 + \frac{f_2 - v_{i,2}}{n}$ and

$$\begin{aligned}\ell(\mathbf{V}_{-i,}\mathbf{v}'_i) &= \bar{v}_1 - f_1 + f_2 + \epsilon - \bar{v}_2 + |\bar{v}_3 - \epsilon - f_3| \\ &\geq \ell(\mathbf{V}).\end{aligned}$$

A symmetric argument works when $v_{i,2} = f_2$. □

Theorem 5.5. *Let f be a phantom mechanism for $m = 3$ and let $\mathbf{V}$ be a locally maximal preference profile, i.e. $\ell(\mathbf{V}) \geq \ell(\mathbf{V}_{-i}, \mathbf{v}'_i)$, for any $i \in [n]$ and any division $\mathbf{v}'_i$. Then, there exists a three-type profile $\hat{\mathbf{V}}$ such that $\ell(\hat{\mathbf{V}}) \geq \ell(\mathbf{V})$.*

Proof. Let S denote the set of single-minded, double-minded or fully-satisfied voters (for mechanism f and for profile $\mathbf{V}$) and let $\bar{S} = [n] \setminus S$.

If $\bar{S} = \emptyset$, $\mathbf{V}$ is a three-type profile. Hence $\hat{\mathbf{V}} = \mathbf{V}$ and the theorem holds trivially. Otherwise, let $i \in \bar{S}$. By Lemma 5.4, we know that we can transform $\mathbf{v}_i$ to $\mathbf{v}'_i$ such that either (a) $\ell(\mathbf{V}_{-i}, \mathbf{v}'_i) > \ell(\mathbf{V})$ or (b) i becomes a double-minded, single-minded or fully-satisfied voter, $f(\mathbf{V}) = f(\mathbf{V}_{-i}, \mathbf{v}'_i)$ and $\ell(\mathbf{V}) = \ell(\mathbf{V}_{-i}, \mathbf{v}'_i)$. When (a) holds, clearly profile $\mathbf{V}$ is not locally maximal. Hence, we can assume that (b) holds for all voters in $\bar{S}$ and we can create $\hat{\mathbf{V}}$ by transforming all voters in $\bar{S}$ to single-minded, double-minded or fully satisfied, one-by-one. By Lemma 5.4, both the outcome and the loss stay invariant in each transformation. Hence, $f(\hat{\mathbf{V}}) = f(\mathbf{V})$ and $\ell(\hat{\mathbf{V}}) = \ell(\mathbf{V})$. The theorem follows. □

From now on, we focus on three-type profiles, and in the following we define variables to describe them. A three-type profile $\mathbf{V}$ can be presented using 13 independent variables:

- $\mathbf{x} = (x_1, x_2, x_3)$, the division of the fully satisfied voters,
- a_1, a_2, a_3, three integer variables counting the single-minded voter towards each project,
- $b_{1,2}, b_{1,3}, b_{2,1}, b_{2,3}, b_{3,1}, b_{3,2}$, six integer variables counting the double-minded voters (e.g. $b_{2,1}$ counts the voters proposing $(1 - x_2, x_2, 0)$)
- and the total number of voters n.

We also use $A = \sum_{j \in [3]} a_j$ and $B = \sum_{j,k \in [3], k \neq j} b_{k,j}$ to count the single-minded and the double-minded voters, respectively. Consequently, the number of fully satisfied voters is $C = n - A - B$. These profiles can have at most 8 distinct voters' reports: values x_1, x_2 and x_3, from fully-satisfied and double-minded voters, values $1 - x_1$, $1 - x_2$ and $1 - x_3$ which we call *complementary values* from the double-minded voters and, reports with values equal to 1 and 0. Note that, apart from values 0 and 1, in project 1 we can find values x_1, $1 - x_2$ and $1 - x_3$, in project 2 values x_2, $1 - x_1$ and $1 - x_3$ and finally in project 3 the values x_3, $1 - x_1$ and $1 - x_2$.

At this point, we break our analysis into two branches. First, we assume that $x_j > 0$, for all $j \in [3]$ from now on. We will examine the cases were $x_j = 0$ for some $j \in [3]$ independently.

Case 1: No zero values exists in the outcome Recall that Definition 11 demands that $f(\mathbf{V}) = \mathbf{x}$. To ensure this, we prove the following lemma.

Lemma 5.6. *Let that $x_j > 0$, for all $j \in [3]$. Let $z_j = a_j + \sum_{k\in[3]\setminus\{j\}} b_{k,j}$ and $q_j = \sum_{k\in[3]\setminus\{j\}} b_{j,k}$. For any moving phantom mechanism f, defined by the phantom system $\mathcal{Y} = \{y_k(t) : k \in \{0..n\}\}$, and any three-type profile $\mathbf{V}$, then $f(\mathbf{V}) = \mathbf{x}$ if and only if*

$$y_{z_j}(t^*) \le x_j \le y_{q_j+z_j+C}(t^*) \tag{5.11}$$

for any

$$t^* \in \left\{ t : \sum_{j\in[m]} \operatorname{med}\left(\mathbf{V}_{i\in[n],j}, (y_k(t))_{k\in\{0..n\}}\right) = 1 \right\}.$$

Proof. First note that for $x_j > 0$ for all $j \in [3]$, all complementary values $1 - x_1$, $1 - x_2$ and $1 - x_3$ are located in the upper slots. Assume otherwise, that there exists some complementary value, say $1 - x_2$ such that $1 - x_2 \le x_1$. Then $1 \le x_1 + x_2$, which is not possible when $x_3 > 0$. In addition, all 1-valued voters' reports should be located in the upper slots, while all 0-valued voter reports should be located in the lower slots. Note also that $z_j = a_j + \sum_{k\in[3]\setminus\{j\}} b_{k,j}$ counts exactly the voters' reports that must be located in the upper slots.

(if direction) Let $\mathbf{V}$ be a three-type profile and let $f(\mathbf{V}) = \mathbf{x}$ for some $t^* \in [0, 1]$. Assume, for the sake of contradiction that $y_{z_j}(t^*) > x_j$. This implies that the $n + 1 - z_j$ phantoms with indices $z_j, ..., n$ are located in the upper slots (i.e. the n higher slots). Since at least z_j voters' reports are also be located in the upper slots there exists at least $n + 1$ values for n slots. A contradiction. Suppose now that that $y_{z_j+q_j+C}(t^*) < f_j(\mathbf{V})$. This implies that $z_j + C + 1$ phantom values (those with indices $0, ..., z_j + C)$) are located in the lower slots (i.e. the n lower slots). The voters' reports with value 0 must be also be located in the lower slots, since $0 < x_j$ for any $\in [3]$. There are exactly $A + B - q_j - z_j = n - C - q_j - z_j$ such values. Hence at least $n + 1$ values should be located in the lower slots. A contradiction.

(only if direction) Let that inequalities 5.11 hold and let $\mathbf{V}$ be a three-type profile. Assume for the sake of contradiction that there exists a project $j \in [3]$ such that $f_j(\mathbf{V}) < x_j$. Hence, the C values of the possibly fully satisfied voters, plus the q_j values equal to x_j should be located in the upper slots. The 1-valued voters' reports, which count to a_1 and the complementary values, which count to $\sum_{k\in[3]\setminus\{j\}} b_{k,j}$ are also located in the upper slots. These count to $z_j + q_j + C$ values. Furthermore, since $x_j \le y_{z_j+q_j+C}(t^*)$, another $n - C - z_j - q_j + 1$ phantom values should be located in the upper slots. A contradiction. Similarly, assume for the sake of contradiction that there exists a profile $j \in [3]$ such that

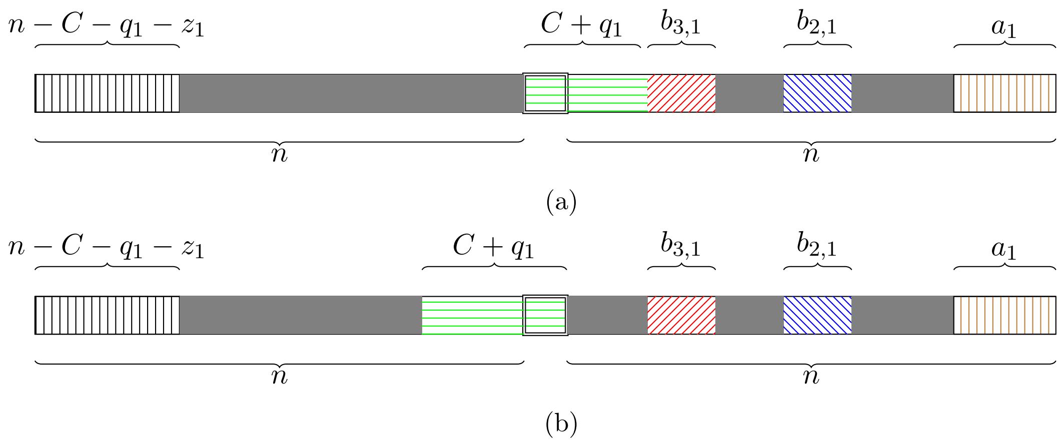


Figure 5.4: The positioning of phantom and voters reports in a single project. The 5 patterned intervals represent the voters' reports. The solid, dark intervals represent the phantom values. The double-lined rectangle in the middle represents the median. Shape 5.4a represents the case where the $C + b_{1,2} + b_{1,3} = C + q_1$ voters' reports with values x_1 are located in the top $n + 1$ slots. The $z_1 = a_1 + b_{2,1} + b_{3,1}$ voters' reports with values 1 and $1 - x_2$ and $1 - x_3$ must also be located in the top $n + 1$ slots.

$f_j(\mathbf{V}) > x_j$. Then the $C + q_j$ values should be located in the lower slots, along with the $n - C - q_j - z_j$ voters' reports equal to 0. Since $y_{z_j}(t^*) \leq x_j$, $z_j + 1$ phantom values are also located in the lower slots. In total $n + 1$ values, a contradiction. □

We note that the only-if direction is not required for the proof of Theorem 5.8, but it is a nice feature that we include for the sake of completeness.

A Non-Linear Program We show that the Piecewise Uniform mechanism is $(2/3+\epsilon)$-approximate, for some $\epsilon \leq 10^{-5}$, using a Non-Linear Program. The feasible region of this program is defined by the class of three-type profiles, and we search for the maximum ℓ_1-loss among them. For simplicity, we firstly normalize all of our variables with n. We introduce new variables $\hat{a}_j = a_j/n$ for $j \in [3]$ and $\hat{b}_{j,j'} = b_{j,j'}/n$ for $j, j' \in [3]$, $j \neq j$, and $\hat{C} = C/n$. We also use a relaxed version of the Piecewise Uniform mechanism: For every $x \in [0, 1]$:

$$\hat{y}(x,t) = \begin{cases} 0 & 0 \le t < \frac{1}{2} \textbf{ and } x \le \frac{1}{2} \\ 4tx - 2t & 0 < t < \frac{1}{2} \textbf{ and } x > \frac{1}{2} \\ x(3-2t) - 2 + 2t & \frac{1}{2} \le t \le 1 \textbf{ and } x > \frac{1}{2} \\ x(2t-1) & \frac{1}{2} \le t \le 1 \textbf{ and } x \le \frac{1}{2}. \end{cases}$$

We introduce also variables for the mean for each project $j \in [3]$:

$$\bar{v}_j = \hat{a}_j + \sum_{k \in [3] \setminus \{j\}} (1 - x_k)\hat{b}_{k,j} + x_j \left(\hat{C} + \sum_{k \in [3] \setminus \{j\}} \hat{b}_{j,k} \right)$$

The Non-Linear Program is presented in Figure 5.5. Inequalities 5.13 and 5.14 ensure that we are searching over all three-type profiles for the Piecewise Uniform mechanism. Crucially, any profile which does not meet these two conditions cannot have $\mathbf{x}$ as the outcome (see Lemma 5.6). Finally, we let the program optimize over any $t^* \in [0, 1]$. Lemma 5.6 ensures that any value t^* that satisfies inequalities 5.13 and 5.14 will return a valid outcome.

Maximum Loss Computation To compute the maximum value of the NLP in Figure 5.5, we break this program in simpler programs, based on 3 conditions; first, depending on whether $t^* \le 1/2$ or not, second, according to the signs of the $\bar{v}_j - x_j$ terms on the objective function (in order to remove the absolute values), and finally, according to the types of the phantoms enclosing the medians.

To deal with the signs of the $\bar{v}_j - x_j$ terms, we define *sign patterns*, as tuples in $\{+, -\}^3$. For example the sign pattern $(+, +, -)$ implies that $\bar{v}_1 \ge x_1$ and $\bar{v}_2 \ge x_2$, while $x_3 \ge \bar{v}_3$. We note that we cannot have the same sign is all projects, unless the loss is equal to 0. To see this, assume otherwise that there exists a sign pattern with the same sign in all projects, say $(+, +, +)$, while the loss is strictly positive. Then, $\sum_{j \in [3]} \bar{v}_j - x_j = \sum_{j \in [3]} \bar{v}_j - \sum_{j \in [3]} -x_j = 0$, a contradiction. Hence, we only need to check the patterns $(+, -, -)$ and $(+, +, -)$.

We need also to address the discontinuities in function $\hat{y}$, with respect to the first argument. For this, we use the tuple (b, r) to distinguish weather the median lies between two red, two black, or between a red an a black phantom. By noting that $\hat{z}_j > 1/2$ implies that $\hat{C} + \hat{q}_j + \hat{z}_j > 1/2$, we can safely assume that no median is upper bounded by a black phantom and lower bounded by a red phantom, and we define *phantom patterns*, as tuples in $\{(b, b), (b, r), (r, r)\}^3$. We build a quadratic program for each phantom pattern.

$$\text{maximize} \sum_{j=1}^{3} |\bar{v}_j - x_j| \tag{5.12}$$

subject to

$$\begin{aligned}
&\sum_{j=1}^{3} x_j = 1 \\
&\hat{A} = \sum_{j=1}^{3} \hat{a}_j \\
&\hat{B} = \sum_{j,k \in [3], j \neq k} \hat{b}_{k,j} \\
&\hat{z}_j = \hat{a}_j + \sum_{k \in [3] \setminus \{j\}} \hat{b}_{k,j} && \forall j \in [3] \\
&\hat{q}_j = \sum_{k \in [3] \setminus \{j\}} \hat{b}_{j,k} && \forall j \in [3] \\
&x_j \geq \hat{y}\left(\hat{z}_j, t^*\right), && \forall j \in [3] \quad (5.13) \\
&x_j \leq \hat{y}\left(\hat{C} + \hat{q}_j + \hat{z}_j, t^*\right), && \forall j \in [3] \quad (5.14) \\
&\hat{A} + \hat{B} \leq 1 \\
&x_j \geq 0, a_j \geq 0, && \forall j \in [3] \\
&b_{k,j} \geq 0, && \forall j, k \in [3], j \neq k \\
&1 \geq t^* \geq 0.
\end{aligned}$$

Figure 5.5: The Non-Linear Program used to upper bound the maximum ℓ_1-loss for the PIECEWISE UNIFORM mechanism.

In total, we end up with $2\times2\times27 = 108$ Quadratic Programs with Quadratic Constraints (QPQC). In Figure 5.11 we present in detail on of these programs. Specifically, we present the case where $t > 1/2$ with sign pattern $(+,-,-)$ and phantom pattern $((r,r),(b,b),(b,b))$. This quadratic program corresponds to a case yielding s large ℓ_1-loss.

In order to prove Theorem 5.8 which follows, we first compute the maxima for 27 programs corresponding to the case $t > 1/2$ and the sign pattern $(+,-,-)$. For the remaining QPQCs, i.e. for $t \geq 1/2$ and the sign pattern $(+,+,-)$ and for $t < 1/2$ for both sign patterns, we check whether any of these cases can yield loss greater than 2/3. We present one example in Figure 5.12. All these programs turn out infeasible, i.e. non of these cases yield ℓ_1 loss greater than 2/3 plus a computational error term which is upper bounded by 10^{-5}.

Case 2: Zero values exists in the outcome To complete our analysis for the upper bound, we need to consider also the cases where the outcome of the mechanism in a three-type profile $\mathbf{x}$ includes values equal to 0. First, note that when the outcome includes two 0 values, i.e. $\mathbf{x} = (1,0,0)$, any three-type $\mathbf{V}$ contains only single-minded voters. Since the mechanism is proportional, the loss is 0, and the Piecewise Uniform mechanism is optimal for such inputs.

In the case where there exists only one 0 value, say $x_3 = 0$, we need to check two possibilities. If $t > 1/2$, there can be only one phantom value equal to 0, and for $x_3 = 0$, at least n voter's reports should be equal to 0. Thus all voters propose 0 for project 3. This can be reduced to the case of 2 projects. The Piecewise Uniform mechanism can ensure a feasible solution by using the phantoms $(k/n)_{k\in\{0..n\}}$, using $t^* = 1$. Theorem 5.2 shows that the ℓ_1-loss in this case cannot be higher than 1/2.

Finally we tackle the remaining case where $x_3 = 0$, $x_1 > 0$, $x_2 > 0$ and $t \leq 1/2$. We will build a different Non-Linear Program to show that this case cannot yield a loss higher than $1/2 + \epsilon$ for some $\epsilon \leq 10^{-5}$. Recall first that $b_{1,2}$ and $b_{2,1}$ counts divisions $(x_1, x_2, 0)$, $b_{1,3}$ counts divisions $(1,0,0)$, $b_{2,3}$ counts divisions $(0, x_2, x_1)$, $b_{3,1}$ counts divisions $(x_1, 0, x_2)$ and $b_{3,2}$ counts divisions $(0,1,0)$. Notice that for each project $j \in \{1,2\}$ the only possible voters' reports are $0, x_j$ and 1. Hence, for project 1 there exists $a_1 + b_{1,3}$ voters' reports with value 1, $b_{2,1} + b_{1,2} + b_{3,1}$ voters' reports with values x_1 and $a_2 + a_3 + b_{2,3} + b_{3,2}$ voters' reports with value 0. Similarly, there exists $a_2 + b_{3,2}$ voters' reports with value 1, $b_{2,1} + b_{1,2} + b_{2,3}$ voters' reports with value x_2 and $a_1 + a_3 + b_{3,1} + b_{1,3}$ voters' reports with value 0. For project 3, there exists a_3 voters' reports with value 1, $b_{2,3}$ voters' reports with value x_1 and $b_{3,1}$ voters reports with value x_2. For project 3, all complementary values are positive, i.e. they are located in the upper slots. This is not always the case for projects 1 and 2, where we can only guarantee that a_1 and a_2 voters' reports (i.e. the 1 valued reports) are located in the upper slots, respectively. On the other side, there exists $C + b_{2,3} + b_{3,2} + a_2 + a_3$ zero values in project 1 and $C + b_{1,3} + b_{3,1} + a_1 + a_3$ from the double-minded, single-minded and the zeros of the fully-satisfied voters. In the following lemma, we use this information to impose sufficient and necessary conditions (similar to Lemma 5.6) for the mechanism to return $\mathbf{x}$ as an outcome.

The following lemma imposes necessary conditions (similar to Lemma 5.6) for the mechanism to return $\mathbf{x}$ as an outcome, for the projects 1 and 2.

Lemma 5.7. *Let that $x_1 > 0$, $x_2 > 0$, $x_3 = 0$. For any moving phantom mechanism f, defined by the phantom system $\mathcal{Y} = \{y_k(t) : k \in \{0..n\}\}$, and a three-type profile $\mathbf{V}$ then $f(\mathbf{V}) = \mathbf{x}$ if and only if:*

$$y_{a_1+b_{1,3}}(t^*) \leq x_1 \leq y_{n-a_2-a_3-b_{2,3}-b_{3,2}}(t^*) \tag{5.15}$$

$$y_{a_2+b_{3,2}}(t^*) \leq x_2 \leq y_{n-a_1-a_3-b_{1,3}-b_{3,1}}(t^*) \tag{5.16}$$

for any

$$t^* \in \left\{ t : \sum_{j\in[m]} \text{med}\left(\mathbf{V}_{i\in[n],j}, (y_k(t))_{k\in\{0..n\}}\right) = 1 \right\}.$$

Proof. We focus on inequality 5.15. Similar arguments can be used to show inequality 5.16.

(if direction) Let $\mathbf{V}$ be a three-type profile and $f(\mathbf{V}) = \mathbf{x}$ for some $t^* \in [0, 1]$. Assume for the sake of contradiction that $y_{a_1+b_{1,3}}(t^*) > x_1$. This implies that the $n - a_1 - b_{1,3} + 1$ phantoms with indices $a_1 + b_{1,3}, ..., n$ are located in the upper slots. Since at least $a_1 + b_{1,3}$ voters' reports should be located in the upper slots (since they have the value 1), at least $n + 1$ values should be located in the lower slots. A contradiction. Suppose now, that $y_{n-a_2-a_3+b_{2,3}+b_{3,2}}(t^*) < f_1(\mathbf{V})$. Then, $n - a_2 - a_3 + b_{2,3} + b_{3,2} + 1$ phantoms are located in the lower slots. Also, at least $a_3 + a_2 + b_{2,3} + b_{3,2}$ zero values should be located in the lower slots. These are $n + 1$ values, while the lower slots are n. A contradiction.

(only if direction) Let that inequalities 5.15 hold, and consider a three-type profile $\mathbf{V}$, where $f_1(\mathbf{V}) < x_1$. Hence the C voters reports from the fully-satisfied voters plus the $b_{1,2} + b_{2,1} + b_{3,1}$ reports equal to x_1 from the double-minded voters and the $a_1 + b_{1,3}$ 1-valued reports, should be located in the upper slots. From inequality 5.15, there exists

$a_2 + a_3 + b_{2,3} + b_{3,2} + 1$ phantoms in the upper slots. These are $n+1$ values, which cannot fit in the n upper slots.

Suppose now, that $x_1 < f_1(\mathbf{V})$. Then the C voters reports from the fully-satisfied voters plus the $b_{1,2} + b_{2,1} + b_{3,1}$ reports equal to x_1 from the double-minded voters and the $a_2 + a_3 + b_{2,3} + b_{3,2}$ voters' reports equal to 0 should be located in the lower slots. Hence, $C + b_{1,2} + b_{1,3} + b_{2,1} + a_1 + b_{3,1} + a_2 + b_{2,3} + b_{3,2}$ voters' reports are located in the lower slots. From inequality 5.15, at least $a_1 + b_{1,3} + 1$ voter reports are located in the lower slots, hence at least $n+1$ values are located in the lower slots. A contradiction. □

Using now Lemma 5.7, we can safely assume that $n - a_2 - a_3 - b_{2,3} - b_{3,2} > 1/2$ and $n - a_1 - a_3 - b_{1,3} - b_{3,1} > 1/2$; otherwise at least one of the x_1 or x_2 is 0, and we have already seen that the mechanism would be have 0 loss in this case. Hence, both projects are upper bounded by a red phantom. We don't have any guarantee for the lower phantoms, hence we use a phantom pattern[5] for each project $j \in \{1, 2\}$: $(r, r)_j$, means that x_j is located between two red phantoms and $(r, b)_j$ means that x_j is upper bounded by a red phantom and lower bounded by a black phantom. We check all 4 possible sign patterns, for the projects 1 and 2. We don't need to examine project 3, since $x_3 \leq y_k(t)$ for any $k \in \{0..n\}$ and any $t \in [0, 1/2]$.

Since $x_3 = 0$, the only sign values we need to tackle are $(-, -, +)$ and $(-, +, +)$. Note that $\bar{v}_3 - x_3 \geq 0$. Hence we need to check 2×4 QPQCs. In Figure 5.10 we present the Non-Linear Program for these special cases.

Quadratic Programs' with Quadratic Constraints Solutions We solve all QPQCs using the Gurobi optimization software [54]. The solver models our programs as Mixed Integer Quadratic Programs and uses the *spatial Branch and Bound Method* (see [66]) with various heuristics, to return a global maximum, when the program is feasible. The solver computes arithmetic solutions with 10^{-5} error tolerance.

Theorem 5.8. *The* PIECEWISE UNIFORM *mechanism is* $(2/3 + \epsilon)$*-approximate, for some* $\epsilon \leq 10^{-5}$.

Proof. Theorem 5.5 states that the maximum loss for any moving phantom mechanism happens in a three-type profile. The Non-Linear Program in Figure 5.5 searches for the profile with maximum loss, over all three-type profiles. The latter is guaranteed by Lemma

[5] Note that these are of a different style from the case of no zeros in the outcome.

Phantoms	**Status**	**Lower bound**	**Upper bound**	**Gap**
$(b,b),(b,b),(b,b)$	OPTIMAL	0.333332	0.333333	1.83e − 6
$(b,b),(b,b),(b,r)$	OPTIMAL	0.357003	0.357007	3.69e − 6
$(b,b),(b,b),(r,r)$	OPTIMAL	0.357003	0.357010	7.38e − 6
$(b,b),(b,r),(b,r)$	OPTIMAL	0.499998	0.500014	1.62e − 5
$(b,b),(b,r),(r,r)$	OPTIMAL	0.499999	0.500010	1.11e − 5
$(b,b),(r,r),(b,b)$	OPTIMAL	0.357004	0.357008	4.27e − 6
$(b,b),(r,r),(r,r)$	OPTIMAL	0.000000	0.000000	0.00
$(b,r),(b,b),(b,b)$	OPTIMAL	0.666667	0.666667	−2.22e − 16
$(b,r),(b,b),(b,r)$	OPTIMAL	0.666666	0.666668	1.92e − 6
$(b,r),(b,b),(r,r)$	OPTIMAL	0.529134	0.529141	7.08e − 6
$(b,r),(b,r),(b,r)$	OPTIMAL	0.666667	0.666672	5.37e − 6
$(b,r),(b,r),(r,r)$	OPTIMAL	0.500000	0.500006	5.67e − 6
$(b,r),(r,r),(b,b)$	OPTIMAL	0.529134	0.529139	5.40e − 6
$(b,r),(r,r),(r,r)$	OPTIMAL	0.000000	0.000000	0.00
$(r,r),(b,b),(b,b)$	OPTIMAL	0.666667	0.666669	1.86e − 6
$(r,r),(b,b),(b,r)$	OPTIMAL	0.666666	0.666667	5.70e − 7
$(r,r),(b,b),(r,r)$	OPTIMAL	0.527863	0.527866	3.25e − 6
$(r,r),(b,r),(b,r)$	OPTIMAL	0.666666	0.666667	1.12e − 6
$(r,r),(b,r),(r,r)$	OPTIMAL	0.500000	0.500004	3.87e − 6
$(r,r),(r,r),(b,b)$	OPTIMAL	0.527864	0.527867	2.55e − 6

Table 5.1: The bounds computed by the QPQCs, for $t > 1/2$ and the sign pattern $(+,-,-)$. The programs without feasible solutions are not presented, as well as symmetric cases. The lower bound corresponds to the largest loss for a feasible solution computed by the solver. The upper bound corresponds to the smaller non-feasible lower bound computed be the solver. The last column shows the gap between them. Gaps smaller than 10^{-5} are insignificant due to the tolerance of the solver.

5.6. We first solve 27 QPQCs, corresponding to the sign pattern $(+,-,-)$ and $t^* > 1/2$. The maximum value is no higher than $2/3+\epsilon$, where ϵ is due to the error tolerance of the solver. Table 5.1 shows analytically the upper bounds computed for each one of the 27 QPQCs (excluding some symmetric cases).

For the other 81 QPQCs we check whether any of them yields loss at least 2/3. For that, we add the constraint $\sum_{j=1}^{3} s(j)(\bar{v}_j - x_j) \geq 2/3$, where $s(j)$ denotes the sign for project j according to the sign pattern and we search for any feasible solution. Eventually, no feasible solution exists, i.e. there exists no three-type without zeros in the outcome with loss at least $2/3+\epsilon$.

Signs	**Phantoms**	**Status**	**Lower bound**	**Upper bound**	**Gap**
$(-,+,+)$	$(r,r),(r,r)$	INFEASIBLE	–	–	–
$(-,+,+)$	$(b,r),(r,r)$	OPTIMAL	0.000000	0.000000	0.00
$(-,+,+)$	$(r,r),(b,r)$	OPTIMAL	0.000000	0.000000	0.00
$(-,+,+)$	$(b,r),(b,r)$	OPTIMAL	0.250002	0.250009	7.49e − 6
$(-,-,+)$	$(r,r),(r,r)$	INFEASIBLE	–	–	–
$(-,-,+)$	$(b,r),(r,r)$	OPTIMAL	0.000000	0.000000	0.00
$(-,-,+)$	$(r,r),(b,r)$	OPTIMAL	0.000000	0.000000	0.00
$(-,-,+)$	$(b,r),(b,r)$	OPTIMAL	0.499999	0.500003	3.43e − 6

Table 5.2: The upper bounds computed by the QPQCs, for $t < 1/2$ and $x_3 = 0$. The lower bound corresponds to the largest loss for a feasible solution computed by the solver. The upper bound corresponds to the smaller non-feasible lower bound computed be the solver. The last column shows the gap between them. Gaps smaller than 10^{-5} are insignificant due to the tolerance of the solver.

Finally, we address the case where the outcome includes at least one 0 value. For that we solve 8 QPQCs, obtained from the NLP in Figure 5.10. No program yields a loss higher than $1/2 + \epsilon$. Table 5.2 shows analytically the bounds obtained.

□

5.4 Lower Bounds

In this section, we provide impossibility results for our proposed measure. Theorem 5.9 shows that no truthful mechanism can be less than 1/2-approximate. Theorem 5.10 focuses on the class of moving phantom mechanisms and shows that no such mechanism can be less than $(1 - 1/m)$-approximate. Theorem 5.11 shows that the Independent Markets mechanism is 0.6862-approximate. Finally, we present lower bounds for large m: a combined lower bound of $2 - \frac{8}{m^{1/3}}$ for both the Independent Markets mechanism and the Piecewise Uniform mechanism, and a lower bound of $2 - \frac{4}{m+1}$ for any mechanism which maximizes the social welfare.

5.4.1 A Lower Bound for any Truthful Mechanism

In the following, we show that truthfulness inevitably admits ℓ_1-loss at least 1/2 in the worst case. We recall that the Uniform Phantom mechanism achieves this bound for $m = 2$.

Theorem 5.9. *No truthful mechanism can achieve ℓ_1-loss less than* $1/2$.

Proof. Let f be a truthful mechanism over m projects. Consider a profile with 2 voters $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2)$, such that $\mathbf{v}_1 = (1, 0, ..., 0)$ and $\mathbf{v}_2 = (0, 1, 0, ..., 0)$ and let that $f(\mathbf{V}) = (x_1, ..., x_m) = \mathbf{x}$ for some $x \in \mathcal{D}(m)$. Consider also the profile $\mathbf{V}' = (\mathbf{x}, \mathbf{v}_2)$. Due to truthfulness, then $f(\mathbf{V}') = \mathbf{x}$. Assume otherwise, that $f(\mathbf{V}) = \mathbf{x}' \neq \mathbf{x}$; when voter's 1 peak is at $\mathbf{x}$, i.e. $\mathbf{v}_1^* = \mathbf{x}$, then the disutility for voter 1 when proposing $\mathbf{v}_1$ is:

$$d(f(\mathbf{V}), \mathbf{v}_1^*) = d(\mathbf{x}, \mathbf{x}) = 0, \tag{5.17}$$

while the disutility for voter 1 when proposing $\mathbf{v}_1^*$ is

$$d(f(\mathbf{V}'), \mathbf{v}_1^*) = d(\mathbf{x}', \mathbf{x}) > 0, \tag{5.18}$$

a contradiction. With a similar argument we can show that for $\mathbf{V}'' = (\mathbf{v}_1, \mathbf{x})$, $f(\mathbf{V}'') = \mathbf{x}$. Hence, the ℓ_1-loss for these two preference profiles is:

$$\ell(\mathbf{V}') = \left|x_1 - \frac{x_1}{2}\right| + \left|x_2 - \frac{1 + x_2}{2}\right| + \sum_{j=3}^{m} \left|x_j - \frac{x_j}{2}\right| = 1 - x_2, \text{ and}$$
$$\ell(\mathbf{V}'') = \left|x_1 - \frac{1 + x_1}{2}\right| + \left|x_2 - \frac{x_2}{2}\right| + \sum_{j=3}^{m} \left|x_j - \frac{x_j}{2}\right| = 1 - x_1.$$

The optimal mechanism should minimize the quantity $\max\{1 - x_1, 1 - x_2\}$, given that $x_1 + x_2 \leq 1$ and $x_1 \geq 0$, $x_2 \geq 0$. Note that $x_1 \leq 1 - x_2$, hence $\max\{1 - x_1, 1 - x_2\} \geq \max\{1 - x_1, x_1\}$ which is minimized for $x_1 = 1/2$ to a value at least $1/2$. □

5.4.2 A Lower Bound for any Moving Phantom Mechanism

In this subsection we present a preference profile where any phantom mechanism yields loss equal to $1 - 1/m$. We recall that the Piecewise Uniform mechanism nearly achieves this bound for $m = 3$.

Theorem 5.10. *No moving phantom mechanism can achieve ℓ_1 loss less than $1 - 1/m$, for any $m \geq 2$.*

Proof. Let $n \geq 2$ and even, and let $S = \{1, ..., n/2\}$, $Q = \{n/2+1, ..., n\}$ be two sets of voters. Let f be a moving phantom mechanism defined over m projects and consider the preference profile $\mathbf{V} = (\mathbf{v}_1, ..., \mathbf{v}_{n/2}, \mathbf{v}_{n/2+1}, ..., \mathbf{v}_n)$. Each voter $i \in S$ proposes the divisions $\mathbf{v}_i = (1, 0, ..., 0)$ each all voter $k \in Q$ proposes the division $\mathbf{v}_k = (1/m, ..., 1/m)$. Let y_i, for $\in \{0, .., n\}$ denote the i-th phantom value and let that $(y_i)_{i \in \{0,...,n\}}$ induce a valid outcome for the moving phantom mechanism. Hence, the outcome of the mechanism is equal to:

$$f_1(\mathbf{V}) = \text{med}\left(\overbrace{\tfrac{1}{m}, ..., \tfrac{1}{m}}^{n/2}, y_0, ..., y_n, \overbrace{1, ..., 1}^{n/2} \right) = x$$

while, for $j \in \{2, ..., m\}$

$$f_j(\mathbf{V}) = \text{med}\left(\overbrace{0, ..., 0}^{n/2}, y_0, ..., y_n, \overbrace{\tfrac{1}{m}, ..., \tfrac{1}{m}}^{n/2} \right) = z.$$

We will show that both $x \leq 1/m$ and $z \leq 1/m$, which implies that $x = z = 1/m$ for the outcome to sum up to 1, and thus being a valid outcome to moving phantom mechanism. Assume, for the sake of contradiction, that either $x > 1/m$ or $z > 1/m$. Starting from the case $z > 1/m$, note that $z \leq x$. For the outcome to sum up to 1, i.e. $x + (m-1)z = 1$, it must be $x < 1/m$, a contradiction.

We continue with the case $x > 1/m$. Then, $n/2$ voters' reports with value equal to $1/m$ should be located in the n lower slots, in the computation of the median for the first project. Hence, at most $n/2$ phantoms can be located in the n lower slots, and the phantom with index $n/2$ should be located in one of the $n+1$ slots higher than the median, i.e. $y_{n/2} \geq x > 1/m$. Observe also that $z < 1/m$, otherwise $x + (1-m)z > 1$. This imply that the $n/2$ voters' reports with value equal to $1/m$ should be located in the n upper slots, in the computation of the medians for the projects 2 to m. Hence, at most $n/2$ phantom values should be located in the n lower slots, and the phantom with index $n/2$ should be located in one of the $n+1$ slots lower than the median, i.e. $y_{n/2} \leq z < 1/m$. A contradiction.

Eventually, for a valid outcome of the mechanism it must be that $x = z = 1/m$ and the ℓ_1-loss becomes:

$$\ell(\mathbf{V}) = \left| \frac{1}{2} - \frac{1}{m} \right| + (m-1) \left| \frac{1}{2 \cdot m} \right| = 1 - \frac{1}{m}. \tag{5.19}$$

□

5.4.3 A Lower Bound for the Independent Markets Mechanism

In this subsection, we present a class of instances where the Independent Markets mechanism from [44] yields loss at least 0.6862, for large enough n. Recall that the Independent Markets mechanism utilizes the phantoms $(\min\{k \cdot t, 1\})_{k\in\{0..n\}}$.

Theorem 5.11. *The* Independent Markets *mechanism is at least* 0.6862*-approximate for three projects.*

Proof. Let f be the Independent Markets mechanism and let $\rho = 2 - \sqrt{2}$. Consider a preference profile $\mathbf{V}$ with n voters, where $\lfloor n\rho \rfloor$ voters propose the division $(1, 0, 0)$ while $\lceil n(1-\rho) \rceil$ voters propose the division $\mathbf{x} = (\sqrt{2}-1, 1-\sqrt{2}/2, 1-\sqrt{2}/2)$. See also Figure 5.6. Let that $t = \frac{\sqrt{2}}{2n}$. Then, $x_1 = n\rho t \geq \lfloor n\rho \rfloor\, t$, i.e $\lfloor n\rho \rfloor + 1$ phantom values with indexes 0 to $\lfloor n\rho \rfloor$ are at most equal to x_1. Hence, there exists $n+1$ values (phantoms and voters' reports) at most equal to x_1, thus $f_1(\mathbf{V}) = x_1$. Similarly, $x_j = n(1-\rho)\, t \leq \lceil n(1-\rho) \rceil\, t$ for $j \in \{1, 2\}$, i.e. the $n + 1 - \lfloor n\rho \rfloor$ phantom values with indices $\lfloor n\rho \rfloor$ to n are at least equal to x_j. Hence there exists $n+1$ values at least equal to x_j, thus $f_j(\mathbf{V}) = x_j$ for $j \in \{2, 3\}$.

The ℓ_1-loss for the preference profile $\mathbf{V}$ is

$$\ell(\mathbf{V}) = \left(3 - 2\sqrt{2}\right)\left(1 - \frac{\lceil n(1-\rho) \rceil}{n}\right) + \frac{\lfloor n\rho \rfloor}{n} \geq 0.6862,$$

for $n \geq 2 \cdot 10^4$. □

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0.4142 & 0.2929 & 0.2929 \\ \vdots & \vdots & \vdots \\ 0.4142 & 0.2929 & 0.2929 \end{array}\right) \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \lfloor n\rho \rfloor \text{ voters} \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} \lceil n(1-\rho) \rceil \text{ voters} \end{array}$$

Figure 5.6: The preference profile which yields a loss at least 0.6862 for the Independent Markets mechanisms.

5.4.4 Lower Bounds for Many Projects

In this section we provide two impossibility results for large number of projects. These results show that some known mechanisms may yield a loss that goes to 2, as m grows large. For the two proportional mechanisms, Independent Markets and Piecewise Uniform mechanisms, we use the same construction to show that the ℓ_1-loss can be as large as $2 - \frac{8}{m^{1/3}}$, for large m. Then, we focus on the mechanisms maximizing the social welfare and we show an even higher lower bound, at $2 - \frac{4}{m}$ for every $m \geq 3$.

Proportional Mechanisms

Theorem 5.12. *Both the* Piecewice Uniform *mechanism and the* Independent Markets *mechanisms yields loss at least* $2 - \frac{8}{m^{1/3}}$ *for* $m > 8$.

Proof. We will construct a preference profile with m projects and m voters. In this profile a supermajority of the voters (denoted by the integer variable z) are single-minded, towards a unique project each. The rest $m - z$ are fully-satisfied.

Let $z = \lfloor m - m^{2/3} \rfloor$ and $a = (m-z)^2$ for some $m \geq 8$. Let $\mathbf{s}^j$ be the division such that $\mathbf{s}^j_j = 1$ for $j \in [m]$ and $\mathbf{x}$ be the division such that $x_j = \frac{1}{a+z}$ for $j \in [z]$ and $x_j = \frac{m-z}{a+z}$ for $j \in \{z+1, ..., m\}$. Consider an instance $\mathbf{V}$ such that $\mathbf{v}_i = \mathbf{s}^i$ for each $i \in [z]$ while for each $i \in \{z+1, ..., m\}$, $\mathbf{v}_i = \mathbf{x}$.

We will first show that the Piecewise Uniform mechanism returns the division $\mathbf{x}$ for any $m \geq 8$. Let that $t^* = \frac{1}{2} + \frac{1}{2(a+z)}$. Let that $m - z \leq m/2$. This is true for every $m > 8$ and implies all phantom values with indices at most $m - z$ are black phantoms, i.e. $y(1, t^*) = \frac{1}{a+z}$. Hence, $f_j(\mathbf{V}) = \frac{1}{a+z}$ for projects $j \in [z]$. To see this, notice in these projects there exists only one voter's report with value strictly higher than $\frac{1}{a+z}$ (the report of the single-minded voter), an well as $m-1$ phantom values (those with indices 2 to m). Since $\frac{1}{a+z}$ is the largest of the other $m+1$ values. Similarly, $y(m-z, t^*) = \frac{m-z}{a+z}$ and $f_j(\mathbf{V}) = \frac{m-z}{a+z}$ for projects $j \in \{z+1, ..., m\}$; there exists z voters' reports with value 0 and $m - z + 1$ phantom values smaller than $y(m-k, t^*)$, while the smaller value of the rest is equal to $\frac{m-z}{a+z}$. In a similar manner, by using $t^* = \frac{1}{a+z}$, the Independent Markets mechanism returns the same outcome.

To compute the loss for $\mathbf{V}$ for the outcome $\mathbf{x}$, first notice that $|\bar{v}_j - f_j| = \frac{1}{m} + \frac{m-z}{m}\frac{1}{a+z} - \frac{1}{a+z} = \frac{a}{m(a+z)}$ for all $j \in [z]$ and $|\bar{v}_j - f_j| = \frac{m-z}{a+z} - \frac{m-z}{m} \cdot \frac{m-z}{a+z} = \frac{(m-z)z}{m(a+z)}$. (Note that $\frac{1}{m} + \frac{m-z}{m}\frac{1}{a+z} \geq \frac{1}{a+z}$ and $\frac{m-z}{a+z} \leq \frac{m-z}{m} \cdot \frac{m-z}{a+z}$). Hence the loss is equal to

$$
\begin{aligned}
\ell(\mathbf{V}) &= z \cdot \frac{a}{m(a+z)} + (m-z) \cdot \frac{(m-z)z}{m(a+z)} \\
&= 2 \cdot \frac{a}{a+z} \cdot \frac{z}{m} \\
&> 2 \left(\frac{m^{4/3}}{m^{4/3}+2m} \right) \left(\frac{m - m^{2/3} - 1}{m} \right) \\
&\geq 2 \left(1 - \frac{2}{m^{1/3}} \right)^2 \\
&\geq 2 - \frac{8}{m^{1/3}}
\end{aligned} \tag{5.20}
$$

For the first inequality we have used the following facts: $z > m - m^{2/3} - 1$, $a \geq m^{4/3}$ and $a + z \leq m^{4/3} + 2m$ for $m \geq 5$. The first one is due to $x \geq \lfloor x \rfloor > x - 1$ for any $x \in \mathbb{R}$. This implies alos that $a = (m - \lfloor m - m^{2/3} \rfloor)^2 \geq m^{4/3}$. Also, notice that $a + z = (m - \lfloor m - m^{2/3} \rfloor)^2 + \lfloor m - m^{2/3} \rfloor \leq (m^{2/3}+1)^2 + m - m^{2/3} + 1 \leq m^{4/3} + 2m$. Second inequality is due to $\left(\frac{m^{4/3}}{m^{4/3}+2m} \right) \geq 1 - \frac{2}{m^{1/3}}$ for $m \geq 0$ and $\left(\frac{m - m^{2/3} - 1}{m} \right) \geq 1 - \frac{2}{m^{1/3}}$ for $m \geq 1$. The last inequality is due to $\left(1 - \frac{c}{x}\right)^2 \geq \left(1 - \frac{2c}{x}\right)$ for $c \geq 0$ and $x > 0$. $\square$

$$
\begin{array}{c}
\overbrace{\hphantom{1 \quad 0 \quad 0 \quad \cdots \quad 0}}^{z} \quad \overbrace{\hphantom{0 \quad \cdots \quad 0}}^{m-z} \\
\left(
\begin{array}{ccccc|ccc}
1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\
\hline
\frac{1}{a+z} & \frac{1}{a+z} & \frac{1}{a+z} & \cdots & \frac{1}{a+z} & \frac{m-z}{a+z} & \cdots & \frac{m-z}{a+z} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
\frac{1}{a+z} & \frac{1}{a+z} & \frac{1}{a+z} & \cdots & \frac{1}{a+z} & \frac{m-z}{a+z} & \cdots & \frac{m-z}{a+z}
\end{array}
\right)
\begin{array}{l}
\left.\vphantom{\begin{array}{c}1\\0\\0\\\vdots\\0\end{array}}\right\} z \\
\left.\vphantom{\begin{array}{c}1\\\vdots\\1\end{array}}\right\} m-z
\end{array}
\end{array}
$$

Figure 5.7: The preference profile which yields a loss at least $2 - \frac{4}{m^{1/3}}$ for tthe Piecewise Uniform and the Independent Markets mechanisms.

Utilitarian mechanisms

In this part, we focus on a family of budget aggregation mechanisms, which we call *utilitarian* mechanisms. A utilitarian mechanism returns an aggregated division which maximizes the social welfare. This is done by minimizes the ℓ_1 distance between the outcome and each voter's proposal, i.e. for a utilitarian mechanism f and a preference profile $\mathbf{V} = (\mathbf{v}_1, ..., \mathbf{v}_n)$, then

$$f(\mathbf{V}) \in \underset{\mathbf{x}\in\mathcal{D}(m)}{\operatorname{argmin}} \sum_{i\in[n]} d(\mathbf{v}_i, \mathbf{x}). \tag{5.21}$$

Mechanisms that maximize the social welfare have been examined in the literature (see Section 1). One of them is a moving phantom mechanism and is proven to be the unique Pareto Optimal mechanism in this family, providing a dichotomy between Proportional the Pareto Optimal moving phantom mechanisms. Here we show that any utilitarian mechanism, inevitably yields very high loss. In contrast to the two proportional mechanisms, these mechanisms yields quite large loss even for small m, e.g. for three project, the ℓ_1-loss is at least 1.

Theorem 5.13. *For any utilitarian mechanism, there exists an instance with ℓ_1-loss equal to $2 - \frac{4}{m+1}$.*

Proof. Let $\mathbf{s}^j$ be the division such that $s^j_j = 1$ for $j \in [m]$. Consider a preference profile with m projects and $m+1$ voters. Each voter $i \in [m+1]$ is single-minded, and $\mathbf{v}_i = \mathbf{s}^i$ when $i \mod m = 0$. Hence, project 1 its fully supported by two voters, while the other projects are supported by one voter each.

Let f be a utilitarian mechanism and let $f(\mathbf{V}) = \mathbf{x}$. Let $x_1 = 1 - \epsilon$ and $\sum_{j=2}^{m} x_j = \epsilon$ for some $\epsilon \in [0, 1]$. We call the quantity $\sum_{i\in[n]} d(\mathbf{v}_i, \mathbf{x})$ the *social cost* (SC). Recall that a utilitarian mechanism minimizes the social cost.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Figure 5.8: An example of the construction used in Theorem 5.13, with 5 voters and 4 projects.

$$\begin{aligned} \text{SC} &= \sum_{i=1}^{m+1} d(\mathbf{v}_i, \mathbf{x}) \\ &= \sum_{i=1}^{m+1} \sum_{j=1}^{m} |v_{i,j} - x_j| \\ &= 2(1 - x_1) + 2\sum_{j=2}^{m} x_j + \sum_{i=2}^{m} (2 - 2x_i) \\ &= 2\epsilon + 2(m-1) \end{aligned} \tag{5.22}$$

The social cost is minimized for $\epsilon = 0$, i.e. the only possible outcome of a utilitarian mechanism for this preference profile is $x_1 = 1$ and x_j for $j \in \{2, ..., m\}$ for $j \in [m]$.

On the other hand, the proportional division assigns $\bar{\mathbf{V}}_1 = \frac{2}{m+1}$ and $\bar{\mathbf{V}}_j = \frac{1}{n+1}$ for any $j \in \{2, ..., m\}$. Eventually, the ℓ_1 loss is:

$$\begin{aligned} \ell(\mathbf{V}) &= \left(1 - \frac{2}{m+1}\right) + \sum_{j=2}^{m} \frac{1}{m+1} \\ &= 2 - \frac{4}{m+1}. \end{aligned} \tag{5.23}$$

□

5.5 Auxiliary Result

Here we present in detail how the Piecewise Uniform mechanism is a moving phantom mechanism. For that, we will present an alternative phantom system $\mathcal{Y}^{\mathrm{PU}'}$ which satisfies Definitions 2. Then, we show that the Piecewise Uniform mechanism simulates this new alternative definition.

The alternative phantom system is $\mathcal{Y}^{\mathrm{PU}'} = \{y'_k(t) : k \in \{0..n\}\}$, for which

$$y'_k(t) = \begin{cases} 0 & \frac{k}{n} \leq \frac{1}{2} \\ \frac{2tk}{n(\frac{1}{2}-\epsilon)} - \frac{t}{\frac{1}{2}-\epsilon} & \frac{k}{n} > \frac{1}{2} \end{cases} \tag{5.24}$$

for $t < 1/2 - \epsilon$,

$$y'_k(t) = \begin{cases} \frac{k(2\,t-1)}{n} + \frac{2k\epsilon}{n} & \frac{k}{n} \leq \frac{1}{2} \\ \frac{k(3-2\,t)}{n} - 2 + 2\,t + 2\epsilon - \frac{2k\epsilon}{n} & \frac{k}{n} > \frac{1}{2}, \end{cases} \tag{5.25}$$

for $1/2 - \epsilon \leq t < 1 - \epsilon$, and

$$y'_k(t) = \frac{k}{n}\left(\frac{1-t}{\epsilon}\right) + \frac{t-1}{\epsilon} + 1, \tag{5.26}$$

for $t > 1 - \epsilon$, for some $0 < \epsilon < 1/2$.

This phantom system satisfies Definition 2: All phantom functions are continuous, $y_k(0) = 0$ and $y_k(1) = 1$ for all $k \in \{0..n\}$. Also, $y_{k+1}(t) \geq y_k(t)$ for all $k \in \{0..n-1\}$.

We notice also that there exists no feasible solution for $t > 1 - \epsilon$, for this phantom system. For $t = 1 - \epsilon$ the phantoms of the mechanism described by the phantom system $\mathcal{Y}^{\mathrm{PU}'}$ are exactly the phantoms used by the Uniform Phantom mechanism. In the following lemma we show that the sum of the medians of the Uniform Phantom mechanism is at least 1. Hence, any phantom returned by the phantom system $\mathcal{Y}^{\mathrm{PU}'}$ with $t > 1 - \epsilon$, would return an outcome that sums to a value strictly larger than 1.

Lemma 5.14. *Let* $\mathbf{x}$ *be the outcome of the Uniform Phantom mechanism on an arbitrary preference profile for some* $m \geq 3$. *Then* $\sum_{j\in[m]} x_j \geq 1$.

Proof. Let k_j be the largest index such that $\frac{k_j}{n} \leq x_j$. Assume for the sake of contradiction that $\sum_{j\in[m]} x_j < 1$, i.e. $\sum_{j\in[m]} \frac{k_j}{n} < 1$. In the slots 1 to $n+1$, there exists exactly k_j phantom values, for each project $j \in [m]$. Hence, there exists exactly $n - k_j$ voters' reports in the same slots. In total, there exists $mn - \sum_{j\in[m]} k_j > n(m-1)$ voter's reports in the slots 1 to $n+1$. Similarly, there exists exactly k_j voters' reports in each project j in the upper slots (slot $n+2$ to slot $2n+1$). Hence in total there exists exactly $\sum_{j\in[m]} k_j < n$ upper slots filled by phantom values out of mn. Since there are mn slots in the upper phantoms, there should be at least $n(m-1)+1$ voter's reports, but we already know that at least $n(m-1)$ out of nm voters reports are located either in the lower slots or in the slots of the median. A contradiction. □

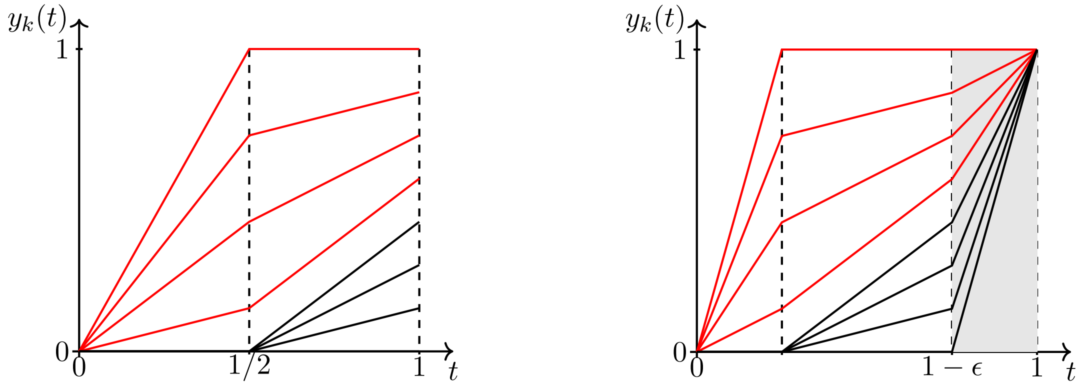
Finally, we will show that the two phantom systems $\mathcal{Y}^{\text{PU}}$ and $\mathcal{Y}^{\text{PU}'}$ describe the same moving phantom mechanism.

Lemma 5.15. *The phantom systems $\mathcal{Y}^{PU}$ and $\mathcal{Y}^{PU'}$ implement the same moving phantom mechanism.*

Proof. We use $y_k(t)$ to denote the functions from the phantom system $\mathcal{Y}^{\text{PU}}$ and $y'_k(t)$ to denote the functions for the phantom system $\mathcal{Y}^{\text{PU}'}$. Consider any preference profile $\mathbf{V}$. Let $f_j(\mathbf{V}) = \text{med}\left(\mathbf{V}_{i\in[n]}, (y_k(t))_{k\in\{0..n\}}\right)$ and $f'_j(\mathbf{V}) = \text{med}\left(\mathbf{V}_{i\in[n]}, (y'_k(t'))_{k\in\{0..n\}}\right)$, for suitable $t \in [0,1]$ and $t' \in [0,1]$ such that $\sum_{j\in m} f_j(\mathbf{V}) = 1$ and $\sum_{j\in[m]} f'_j(\mathbf{V}) = 1$. We will show that the Piecewise Uniform implements the mechanism described by $\mathcal{Y}^{\text{PU}'}$, i.e. for each $t' \in [0,1]$ which yields a feasible solution, there exists some t which yields the same solution using the phantom system $\mathcal{Y}^{\text{PU}}$. Since the phantom system $\mathcal{Y}^{\text{PU}'}$ describes a moving phantom mechanism, the $\mathcal{Y}^{\text{PU}}$ describes also a phantom mechanism.

We consider the phantom system $\mathcal{Y}^{\text{PU}'}$. Assume that $t' > 1-\epsilon$. By Lemma 5.14, $\sum_{j\in[m]} f'_j(\mathbf{V}) > 1$, i.e. no feasible outcome is possible with $t' > 1-\epsilon$. Assume now that $1/2 - \epsilon < t' \leq 1-\epsilon$. Then, by using $t = t' + \epsilon$, the tuples $(y'_k(t'))_{k\in\{0..n\}}$ and $(y_k(t))_{k\in\{0..n\}}$ are equivalent, hence $f'(\mathbf{V}) = f(\mathbf{V})$. Finally, assume that $f(\mathbf{V})$ uses some $t \leq 1/2 - \epsilon$ and returns a feasible solution. Then, by using $t' = \frac{t}{1-2\epsilon}$, the tuples $(y'_k(t'))_{k\in\{0..n\}}$ and $(y_k(t))_{k\in\{0..n\}}$ are equivalent, hence $f'(\mathbf{V}) = f(\mathbf{V})$.

□



(a) $\mathcal{Y}^{\mathrm{PU}'}$ (b) $\mathcal{Y}^{\mathrm{PU}'}$

Figure 5.9: The phantom systems $\mathcal{Y}^{\mathrm{PU}}$ and $\mathcal{Y}^{\mathrm{PU}'}$ for $n = 7$. No feasible solution is possible for any t in the shaded area in Figure 5.9b.

$$\text{maximize} \quad x_1 - \bar{v}_1 + \sum_{j=2}^{3} \bar{v}_j - x_j \tag{5.27}$$

subject to

$$\begin{aligned}
\sum_{j=1}^{3} x_j &= 1 \\
\hat{A} &= \sum_{j=1}^{3} \hat{a}_j \\
\hat{B} &= \sum_{j,k \in [3], j \neq k} \hat{b}_{k,j} \\
\hat{y}_{a_1+b_{1,3}}(t^*) &\leq x_1 \leq \hat{y}_{1-\hat{a}_2-\hat{a}_3-\hat{b}_{2,3}-\hat{b}_{3,2}}(t^*) \\
\hat{y}_{a_2+b_{3,2}}(t^*) &\leq x_2 \leq \hat{y}_{1-\hat{a}_1-\hat{a}_3-\hat{b}_{1,3}-\hat{b}_{3,1}}(t^*) \\
\hat{A} + \hat{B} &\leq 1 \\
x_j &\geq 0, a_j \geq 0, && \forall j \in [3] \\
b_{k,j} &\geq 0, && \forall j, k \in [3] \\
0 &\leq t^* \leq 1/2.
\end{aligned}$$

Figure 5.10: The Non-Linear Program computing the maximum loss for the special case $t < 1/2$, $(-, +, +)$, $x_3 = 0$.

$$
\begin{aligned}
\text{maximize} \quad & \bar{v}_1 - x_1 + x_2 - \bar{v}_2 + x_3 - \bar{v}_3 \\
\text{subject to} \quad & \\
& \sum_{j=1}^{3} x_j = 1 \\
& \hat{A} = \sum_{j=1}^{3} \hat{a}_j \\
& \hat{B} = \sum_{j,k \in [3], j \neq k} \hat{b}_{k,j} \\
& \hat{z}_j = \hat{a}_j + \sum_{k \in [3] \setminus \{j\}} \hat{b}_{k,j} && \forall j \in [3] \\
& \hat{q}_j = \sum_{k \in [3] \setminus \{j\}} \hat{b}_{j,k} && \forall j \in [3] \\
& x_1 \geq \hat{z}_1 (3 - 2t^*) - 2 + 2t^*, \\
& \hat{z}_1 \geq 1/2 \\
& \hat{C} + \hat{z}_1 + \hat{q}_1 \geq 1/2 \\
& x_1 \leq \left(\hat{C} + \hat{z}_1 + \hat{q}_1 \right) (3 - 2t^*) - 2 + 2t^* \\
& x_j \geq \hat{z}_j (2t^* - 1), && \forall j \in \{2, 3\} \\
& x_j \leq \left(\hat{C} + \hat{z}_j + \hat{q}_j \right) (2t^* - 1), && \forall j \in \{2, 3\} \\
& \hat{z}_j \leq 1/2 && \forall j \in \{2, 3\} \\
& \hat{C} + \hat{z}_j + \hat{q}_j \leq 1/2 && \forall j \in \{2, 3\} \\
& \hat{A} + \hat{B} \leq 1 \\
& x_j \geq 0, a_j \geq 0, && \forall j \in [3] \\
& b_{k,j} \geq 0, && \forall j, k \in [3] \\
& 1 \geq t^* \geq 1/2.
\end{aligned}
\tag{5.28}
$$

Figure 5.11: The Quadratic Program with Quadratic Constraints for the maximum loss computation: case $t > 1/2$, $(+, -, -)$, $((r, r), (b, b), (b, b))$. The inequalities referring to function $\hat{y}$ are now replaced by new 4 new inequalities, using the help of the phantom pattern. This quadratic program yields a close-to-maximum upper bound.

$$
\begin{aligned}
&\text{maximize} \quad 1 \\
&\text{subject to} \\
&2/3 \leq \bar{v}_1 - x_1 + \bar{v}_2 - x_2 + x_3 - \bar{v}_3 && && (5.29) \\
&\sum_{j=1}^{3} x_j = 1 \\
&\hat{A} = \sum_{j=1}^{3} \hat{a}_j \\
&\hat{B} = \sum_{j,k \in [3], j \neq k} \hat{b}_{k,j} \\
&\hat{z}_j = \hat{a}_j + \sum_{k \in [3] \setminus \{j\}} \hat{b}_{k,j} && \forall j \in [3] \\
&\hat{q}_j = \sum_{k \in [3] \setminus \{j\}} \hat{b}_{j,k} && \forall j \in [3] \\
&x_j \geq \hat{z}_j (2t^* - 1), && \forall j \in [3] \\
&x_j \leq \left(\hat{C} + \hat{z}_j + \hat{q}_j \right) (2t^* - 1), && \forall j \in [3] \\
&\hat{z}_j \leq 1/2, && \forall j \in [3] && (5.30) \\
&\hat{C} + \hat{z}_j + \hat{q}_j \leq 1/2, && \forall j \in [3] && (5.31) \\
&\hat{A} + \hat{B} \leq 1 \\
&x_j \geq 0, a_j \geq 0, && \forall j \in [3] \\
&b_{k,j} \geq 0, && \forall j, k \in [3] \\
&0 \leq t^* \leq 1/2.
\end{aligned}
$$

Figure 5.12: The Quadratic Program with Quadratic Constraints for the maximum loss computation: case $t \leq 1/2$, $(+,+,-)$, $((b,b),(b,b),(b,b))$. This program has no feasible solutions, i.e. for this specific case the loss cannot exceed 2/3, plus the computational error from the solver.

Chapter 6

Conclusion and Future directions

In this chapter, we wrap-up our main results, discuss various open problems, and explore some directions for future research projects.

6.1 Impartial Selection with Additive Approximation Guarantees

In Chapter 3 we propose the design of impartial selection mechanisms that minimize the worst-case additive approximation, i.e. the maximum, over all possible nomination graphs, difference between the maximum in-degree and in-degree of the selected node. We have provided upper and lower bounds for two models of nomination graphs: The multiple-nomination model where the nodes can nominate multiple other nodes and even abstain, and the more strict, single-nomination model where each node must nominate exactly one other node as the winner. Some of our bounds are not tight, and further investigation is needed.

The most important open problem is to close the gap for the additive approximation bounds on the multiple-nomination model. While we do not know any mechanism with additive approximation less than $n-1$ (See Section 3.5), our lower bound is just 3 (See Section 3.4.2). Increasing this lower bound seems to be a challenging problem.

A second important question is whether we can have a characterization for the family of *Sample* mechanisms, which we define in Section 3.4.1. This family of mechanisms arises naturally and many of the mechanisms presented here are included in this family. In this thesis we were able to characterize the much-restricted family of *Strong Sample* mechanism for the single-nomination model as a family of mechanisms where the sample set, a set of nodes that define the outcome, must be defined independently of the graph, otherwise, impartiality fails. A similar result for the much wider family of Sample mechanisms would be a significant step for the understanding of impartial mechanisms.

Finally, we note that our characterization of strong sample mechanisms covers only the single-nomination model. A further step is to expand this characterization (if this is possible) to the multiple-nomination model.

6.2 Impartial Selection with Prior Information

In Chapter 4 we propose impartial mechanisms which use prior information to yield better additive approximation bounds. Our polylogarithmic upper bound in Section 4.3.2 shows that prior information can yield dramatic improvements on the performance of simple impartial selection mechanisms. It also gives hope that AVD could be similarly efficient for the more general opinion poll instances. Unfortunately, this is not true as the following counter-example indicates.

Indeed, starting from a uniform instance with $n+1$ nodes of popularity $p = 1/2$, we add a new copy j' for each node j. Also, for every edge (i, j) realized in the original instance, we add the edge (i, j'). In this way, we construct opinion poll instances with $2(n+1)$ nodes, where no node can ever beat all the other nodes. Hence, in such instances, AVD will behave as the Fixed Winner mechanism in the original instance and will always return the default node as winner. By applying Theorem 4.13 we obtain the following negative result for AVD.

Theorem 6.1. *When applied on opinion poll instances, the AVD mechanism has expected additive approximation* $\Omega(\sqrt{n \ln n})$.

We should note that the above construction is fragile, in the sense that it exploits a very specific aspect of the mechanism. So, still, the quest of designing deterministic mechanisms that achieve polylogarithmic additive approximation in the opinion poll model is very important and challenging. A starting step could be to restrict our attention to the instances considered in [57], in which every voter approves exactly one other candidate.

Finally, throughout Chapter 4, we have assumed that the prior information is reliable. This should not be expected to be always the case in practice. We expect that our results on the constant mechanism still hold if we have a rough estimate of the highest in-degree. Highest accuracy seems to be necessary to recover our polylogarithmic upper bound though.

This issue is also related to the strengths of prior-independent mechanisms (e.g., see Section 4.3 of [55]) and needs to be investigated further.

6.3 Truthful Budget Aggregation with Proportionality Guarantees

In Chapter 5 we focus on a problem where a set of strategic individuals need to collectively divide a budget over a set of projects. Each individual comes with a proposal on how the budget should be spent, and an aggregation mechanism needs to decide how eventually the budget should be spent. We consider a family of provably truthful mechanisms, and we are interested in quantifying the loss due to truthfulness, compared to the average proposal. We were able to identify effectively optimal truthful mechanisms for the case of 2 and 3 projects.

We have proposed the PIECEWISE UNIFORM mechanism and we have shown that this mechanism is practically optimal with respect to the worst-case ℓ_1-loss, achieving an upper bound not higher than $2/3+10^{-5}$. Since we have shown that no moving phantom mechanism can achieve an ℓ_1-loss smaller than 2/3, a natural question is whether a more precise analysis can scrap this 10^{-5} gap. The PIECEWISE UNIFORM mechanism is also proven to be proportional. A weakness of this mechanism is that it does not satisfy k-proportionality, as proposed in [44]. For example, when the input consist of only 4 voters, two of which propose the division $(0.5, 0.5, 0)$, one proposes the division $(0.5, 0, 0.5)$ and the fourth voter proposes the division $(0, 0.5, 0.5)$, the PIECEWISE UNIFORM mechanism returns the division $(0.5, 0, 0.5)$[1], while the proportional division is $(0.375, 0.25, 0.375)$ and the ℓ_1-loss is equal to 1/2. An open question here is whether there exists any proportional, 2-proportional and optimal with respect to ℓ_1-loss, moving phantom mechanism.

A natural and more interesting open question, is what happens for a large number of projects. In Section 5.4.4 we have shown that all the mechanisms we have considered fail terribly when the number of projects is large, i.e. their ℓ_1-loss is asymptotically close to 2, the maximum possible value for this measure. This leads us to ask the following: Is there any moving phantom mechanism that can restrict the ℓ_1-loss to some constant strictly smaller than 2 for an arbitrary number on projects? Or there exist some general way that our

[1] To see why, notice that the PIECEWISE UNIFORM uses the phantom values 0,0,0,0.5, and 1, obtained with $t = 1/2$.

measure reaches its worst-case scenario of 2 as the number of projects increases, regardless, for any truthful mechanism. In the latter scenario, a good place to start is our lower bound construction for Theorem 5.12.

We note here, that in practice we cannot expect the number of projects to be very large. Hence, even if eventually we cannot find truthful mechanisms which keeps the ℓ_1-loss lower than 2 for any number of projects, its worthy to try and find optimal mechanisms for a small or moderate number of projects.

Going beyond the moving phantom mechanism, we can still ask the questions: Is there any mechanism with worst-case ℓ_1-loss less than 2/3 for the case of three projects? Our lower bound for any truthful mechanism is just 1/2. While this lower bound is very simple, and probably a more sophisticated construction may answer the question negatively, we should note that until now the existence of one such mechanism is still possible.

Finally, with have shown that our ℓ_1-loss measure is incompatible with the utilitarian welfare, in the sense that social welfare maximization leads to very large ℓ_1-loss. There are two other popular measures in the literature (see e.g. [5]), the *egalitarian* welfare, which minimizes the worst-case disutility for any single voter and the *Nash* welfare, which is seen as a compromise between the other two measures. An interesting question is to find connections between the ℓ_1-loss and the other two measures. An even more compelling question, is to find optimal moving phantom mechanisms for these two popular measures.

Bibliography

[1] N. Alon, A. Shapira, and B. Sudakov. Additive approximation for edge-deletion problems. *Annals of Mathematics*, 170(1):371–411, 2009.

[2] N. Alon, F. Fischer, A. Procaccia, and M. Tennenholtz. Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 101–110, 2011.

[3] R. B. Ash. *Information Theory*. Courier Corporation, 1990.

[4] H. Attiya and J. Welch. *Distributed Computing: Fundamental, Simulations, and Advanced Topics*. John Wiley and Sons, Inc., 2nd edition, 2004.

[5] H. Aziz and N. Shah. Participatory budgeting: Models and approaches. In *Pathways Between Social Science and Computational Social Science*, pages 215–236. Springer, 2021.

[6] H. Aziz, B. E. Lee, and N. Talmon. Proportionally representative participatory budgeting: Axioms and algorithms. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 23–31, 2018.

[7] H. Aziz, A. Bogomolnaia, and H. Moulin. Fair mixing: the case of dichotomous preferences. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, pages 753–781, 2019.

[8] H. Aziz, O. Lev, N. Mattei, J. S. Rosenschein, and T. Walsh. Strategyproof peer selection using randomization, partitioning, and apportionment. *Artificial Intelligence*, 275:295–309, 2019.

[9] H. Aziz, A. Lam, B. E. Lee, and T. Walsh. Strategyproof and proportionally fair facility location. *arXiv preprint arXiv:2111.01566*, 2021.

[10] Y. Babichenko, O. Dean, and M. Tennenholtz. Incentive-compatible diffusion. In *Proceedings of the 27th International Conference on World Wide Web (WWW)*, pages 1379–1388, 2018.

[11] Y. Babichenko, O. Dean, and M. Tennenholtz. Incentive-compatible selection mechanisms for forests. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*, pages 111–131, 2020.

[12] S. Barberà. Strategyproof social choice. *Handbook of Social Choice and Welfare*, 2: 731–831, 2011.

[13] G. Benade, S. Nath, A. D. Procaccia, and N. Shah. Preference elicitation for participatory budgeting. *Management Science*, 2020.

[14] J. Beus. Why (almost) everything you knew about Google CTR is no longer valid. Retrieved from: `https://www.sistrix.com/blog/why-almost-everything-you-knew-about-google-ctr-is-no-longer-valid`, 07 2020. Accessed: 2021-10-10.

[15] A. Bjelde, F. Fischer, and M. Klimm. Impartial selection and the power of up to two choices. *ACM Transactions on Economics and Computation*, 5(4):21, 2017.

[16] D. Black. On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–34, 1948.

[17] A. Bogomolnaia, H. Moulin, and R. Stong. Collective choice under dichotomous preferences. *Journal of Economic Theory*, 122(2):165–184, 2005.

[18] B. Bollobás. *Random Graphs.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2nd edition, 2001.

[19] N. Bousquet, S. Norin, and A. Vetta. A near-optimal mechanism for impartial selection. In *Proceedings of the 10th International Conference on Web and Internet Economics (WINE)*, pages 133–146, 2014.

[20] S. Boyd, S. P. Boyd, and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[21] J. Brands, F. Steutel, and R. Wilms. On the number of maxima in a discrete sample. *Statistics and Probability Letters*, 20(3):209–217, 1994.

[22] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[23] Y. Cabannes. Participatory budgeting: a significant contribution to participatory democracy. *Environment and Urbanization*, 16(1):27–46, 2004.

[24] Q. Cai, A. Filos-Ratsikas, and P. Tang. Facility location with minimax envy. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 137–143, 2016.

[25] I. Caragiannis. Recent advances in large-scale peer grading. In U. Endriss, editor, *Trends in Computational Social Choice*, chapter 17, pages 327–344. AI Access, 2017.

[26] I. Caragiannis, A. Procaccia, and N. Shah. Truthful univariate estimators. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning (ICML)*, volume 48, pages 127–135, 2016.

[27] I. Caragiannis, G. Christodoulou, and N. Protopapas. Impartial selection with additive approximation guarantees. In *Algorithmic Game Theory: 12th International Symposium (SAGT)*, pages 269–283, 2019.

[28] I. Caragiannis, G. Christodoulou, and N. Protopapas. Truthful aggregation of budget proposals with proportionality guarantees. To Appear: In Proceedings of the 36th AAAI Conference on Artificial Intelligence, 2022.

[29] H. Chan, A. Filos-Ratsikas, B. Li, M. Li, and C. Wang. Mechanism design for facility location problems: A survey. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence, (IJCAI)*, pages 4356–4365, 8 2021. Survey Track.

[30] G. Christodoulou, E. Koutsoupias, and A. Kovács. On the Nisan-Ronen conjecture. In *Proceedings of the 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 839–850, 2022.

[31] E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 11:17–33, 1971.

[32] Y. J. Cohler, J. K. Lai, D. C. Parkes, and A. D. Procaccia. Optimal envy-free cake cutting. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI)*, 2011.

[33] G. de Clippel, H. Moulin, and N. Tideman. Impartial division of a dollar. *Journal of Economic Theory*, 139(1):176–191, 2008.

[34] B. Dean. Here's what we learned about organic click through rate. Retrieved from: `https://backlinko.com/google-ctr-stats`, 2019. Accessed: 2021-10-10.

[35] O. Dekel, F. Fischer, and A. D. Procaccia. Incentive compatible regression learning. *Journal of Computer and System Sciences*, 76(8):759–777, 2010.

[36] S. Dobzinski, N. Nisan, and M. Schapira. Truthful randomized mechanisms for combinatorial auctions. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 644–652, 2006.

[37] C. Duddy. Fair sharing under dichotomous preferences. *Mathematical Social Sciences*, 73:1–5, 2015.

[38] B. Eisenberg and G. Stengle. Minimizing the probability of a tie for first place. *Journal of Mathematical Analysis and Applications*, 198(2):458–472, 1996.

[39] B. Eisenberg, G. Stengle, and G. Strang. The asymptotic probability of a tie for first place. *The Annals of Applied Probability*, 3(3):731–745, 1993.

[40] E. Elkind, M. Lackner, and D. Peters. Structured preferences. In U. Endriss, editor, *Trends in Computational Social Choice*, chapter 10, pages 187–207. AI Access, 2017.

[41] P. Erdős and R. J. Wilson. On the chromatic index of almost all graphs. *Journal of Combinatorial Theory, Series B*, 23(2-3):255–257, 1977.

[42] B. Fain, A. Goel, and K. Munagala. The core of the participatory budgeting problem. In *Proceedings of the 12th International Conference on Web and Internet Economics (WINE)*, pages 384–399, 2016.

[43] F. Fischer and M. Klimm. Optimal impartial selection. *SIAM Journal on Computing*, 44(5):1263–1285, 2015.

[44] R. Freeman, D. M. Pennock, D. Peters, and J. W. Vaughan. Truthful aggregation of budget proposals. *Journal of Economic Theory*, 193:105234, 2021.

[45] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

[46] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[47] A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41: 587–601, 1973.

[48] A. Goel, A. K. Krishnaswamy, S. Sakshuwong, and T. Aitamurto. Knapsack voting for participatory budgeting. *ACM Transactions on Economics and Computation*, 7 (2):8:1–8:27, 2019.

[49] M. X. Goemans. Minimum bounded degree spanning trees. In *Proceedings of the 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 273–282, 2006.

[50] J. Goldman and A. D. Procaccia. Spliddit: Unleashing fair division algorithms. *ACM SIGecom Exchanges*, 13(2):41–46, 2015.

[51] S. Gonçalves. The effects of participatory budgeting on municipal expenditures and infant mortality in Brazil. *World Development*, 53:94–110, 2014. Decentralization and Governance.

[52] C. Grannan. How is diving scored? Retrieved from: `https://www.britannica.com/story/how-is-diving-scored`. Accessed: 2021-10-15.

[53] T. Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973.

[54] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2021. URL `https://www.gurobi.com`.

[55] J. D. Hartline. *Bayesian Mechanism Design*. Foundations and trends in theoretical computer science. Now Publishers, 2013.

[56] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[57] R. Holzman and H. Moulin. Impartial nominations for a prize. *Econometrica*, 81(1): 173–196, 2013.

[58] A. Kahng, Y. Kotturi, C. Kulkarni, D. Kurokawa, and A. D. Procaccia. Ranking wily people who rank each other. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1087–1094, 2018.

[59] G. Kendall and L. J. Lenten. When sports rules go awry. *European Journal of Operational Research*, 257(2):377–394, 2017.

[60] D. Kurokawa, O. Lev, J. Morgenstern, and A. D. Procaccia. Impartial peer review. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 582–588, 2015.

[61] R. Laraki and E. Varloot. Level-strategyproof belief aggregation mechanisms. *arXiv preprint arXiv:2108.04705*, 2021.

[62] J.-F. Laslier and M. R. Sanver, editors. *Handbook on Approval Voting*. Springer, 2010.

[63] P. Legros and E. Cantillon. The Nobel prize: What is mechanism design and why does it matter for policy-making? *VoxEU. org. Vox*, 18, 2007.

[64] B. Lehmann, D. Lehmann, and N. Nisan. Combinatorial auctions with decreasing marginal utilities. *Games and Economic Behavior*, 55(2):270–296, 2006.

[65] D. Lehmann, L. I. O'Callaghan, and Y. Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM*, 49(5):577–602, 2002.

[66] L. Liberti. Introduction to global optimization. *Ecole Polytechnique*, 2008.

[67] T. Lindner. *Zur Manipulierbarkeit der Allokation öffentlicher Güter: Theoretische Analyse und Simulationsergebnisse*. KIT Scientific Publishing, 2011.

[68] T. Lindner, K. Nehring, and C. Puppe. Allocating public goods via the midpoint rule. In *Proceedings of the 9th International Meeting of the Society for Social Choice and Welfare*, 2008.

[69] P. Louis, M. Nuñez, and D. Xefteris. Trimming extreme opinions in preference aggregation. University of Cyprus Working Papers in Economics, University of Cyprus, Department of Economics, 2019.

[70] T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2011.

[71] A. Mackenzie. Symmetry and impartial lotteries. *Games and Economic Behavior*, 94: 15–28, 2015.

[72] A. Mackenzie. An axiomatic analysis of the papal conclave. *Economic Theory*, 69: 713–743, 2020.

[73] N. Mattei, P. Turrini, and S. Zhydkov. PeerNomination: Relaxing exactness for increased accuracy in peer selection. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 393–399, 2020.

[74] S. McNulty. An unlikely success: Peru's top-down participatory budgeting experience. *Journal of Public Deliberation*, 8(2), 2012.

[75] M. Michorzewski, D. Peters, and P. Skowron. Price of fairness in budget division and probabilistic social choice. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, pages 2184–2191, 2020.

[76] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[77] H. Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.

[78] K. Nehring and C. Puppe. Resource allocation by frugal majority rule. Technical report, KIT Working Paper Series in Economics, 2019.

[79] N. Nisan. Introduction to mechanism design (for computer scientists). In N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors, *Algorithmic Game Theory*, chapter 9, pages 209–242. Cambridge University Press, 2007.

[80] N. Nisan and A. Ronen. Algorithmic mechanism design. *Games and Economic behavior*, 35(1-2):166–196, 2001.

[81] K. Nissim, R. Smorodinsky, and M. Tennenholtz. Approximately optimal mechanism design via differential privacy. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS)*, page 203–213, 2012.

[82] M. Okamoto. Some inequalities relating to the partial sum of binomial probabilities. *Annals of the Institute of Statistical Mathematics*, 10(1):29–35, 1958.

[83] PFA Player of the Year: Kevin de Bruyne and Beth England named 2020 winners. British Broadcasting Corporation (BBC), Retrieved from: `https://www.bbc.com/sport/football/54044861`, 09 2020. Accessed: 2022-04-15.

[84] A. D. Procaccia and M. Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation*, 1(4):1–26, 2013.

[85] C. Puppe and J. Rollmann. Mean versus median voting in multi-dimensional budget allocation problems. A laboratory experiment. *Games and Economic Behavior*, 130: 309–330, 2021.

[86] R. Renault and A. Trannoy. Protecting minorities through the average voting rule. *Journal of Public Economic Theory*, 7(2):169–199, 2005.

[87] R. Renault and A. Trannoy. Assessing the extent of strategic manipulation: the average vote example. *SERIEs*, 2(4):497–513, 2011.

[88] F. Rosar. Continuous decisions by a committee: median versus average mechanisms. *Journal of Economic Theory*, 159:15–65, 2015.

[89] A. E. Roth and M. Sotomayor. Two-sided matching. *Handbook of Game Theory with Economic Applications*, 1:485–541, 1992.

[90] A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488, 2004.

[91] T. Roughgarden. *Twenty lectures on Algorithmic Game Theory*. Cambridge University Press, 2016.

[92] T. Roughgarden and M. Sundararajan. Quantifying inefficiency in cost-sharing mechanisms. *Journal of the ACM*, 56(4):1–33, 2009.

[93] M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

[94] J. Schummer and R. V. Vohra. Mechanism design without money. In N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors, *Algorithmic Game Theory*, chapter 10, pages 243–299. Cambridge University Press, 2007.

[95] P. Serafino, C. Ventre, and A. Vidali. Truthfulness on a budget: trading money for approximation through monitoring. *Autonomous Agents and Multi-Agent Systems*, 34 (1):1–24, 2020.

[96] C. Su. From Porto Alegre to New York City: Participatory budgeting and democracy. *New Political Science*, 39(1):67–75, 2017.

[97] S. Tamura. Characterizing minimal impartial rules for awarding prizes. *Games and Economic Behavior*, 95:41–46, 2016.

[98] S. Tamura and S. Ohseto. Impartial nomination correspondences. *Social Choice and Welfare*, 43(1):47–54, 2014.

[99] Z. Tang, C. Wang, and M. Zhang. Price of fairness in budget division for egalitarian social welfare. In *Proceedings of the 14th International Conference on Combinatorial Optimization and Applications (COCOA)*, pages 594–607, 2020.

[100] Terry claims player of year award. British Broadcasting Corporation (BBC), Retrieved from: `http://news.bbc.co.uk/sport1/hi/football/eng_prem/4478329`, 04 2005. Accessed: 2022-04-15.

[101] T. N. Tideman and F. Plassmann. Paying the partners. *Public Choice*, 136(1):19–37, 2008.

[102] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37, 1961.

[103] B. Wampler. Participatory budgeting: Core principles and key impacts. *Journal of Public Deliberation*, 2012.

[104] B. Wampler, S. McNulty, and M. Touchton. Participatory budgeting: Spreading across the globe. *Transparency & Accountability Initiative*, 2018.

[105] E. Williams, E. St Denny, and D. Bristow. Participatory budgeting: An evidence review, 2019.

[106] D. P. Williamson and D. B. Shmoys. *The design of approximation algorithms*. Cambridge university press, 2011.

[107] Y. Xu, H. Zhao, X. Shi, and N. B. Shah. On strategyproof conference peer review. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence, (IJCAI)*, pages 616–622, 7 2019.

[108] X. Zhang, Y. Zhang, and D. Zhao. Incentive compatible mechanism for influential agent selection. In *Algorithmic Game Theory: 14th International Symposium (SAGT)*, pages 79–93, 2021.