DECIDING WHO, WHAT, WHY, AND HOW: AGGREGATING PREFERENCES OVER AGENTS, ALTERNATIVES, AXIOMS, AND RULES

Ben Abramowitz

Submitted in Partial Fullfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Approved by:
Elliot Anshelevich, Chair
Malik Magdon-Ismail
Nicholas Mattei
Lirong Xia

Department of Computer Science
Rensselaer Polytechnic Institute
Troy, New York

[August 2021]
Submitted June 2021

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

First, I am incredibly grateful to my advisor Elliot Anshelevich for his guidance, patience, and support. When I first visited RPI, Elliot's former student Onkar told me how lucky I was to have Elliot as an advisor. Over the past few years I have discovered how right Onkar was. In addition to my advisor, I have been fortunate to collaborate with a number of other brilliant and talented coauthors including Wennan Zhu, Nicholas Mattei, Udi Shapiro, Nimrod Talmon, Davide Grossi, and Edith Elkind. I have also had the joy of working alongside many others at IBM Research, The Weizmann Institute of Science, and RPI. I would especially like to thank Wennan for making working in the office so much more enjoyable, Nick for taking me under his wing at IBM, Udi for inviting me to work with his group as a visiting student, and K. Brent Venable for first introducing me to research. I couldn't be more excited to return to Tulane as a postdoctoral student to work with Nick this fall. I also thank my committee members Lirong Xia and Malik Magdon-Ismail, as well as Tracy Hoffman, Chris Coonrad, Shannon Carrothers, and the late Terry Hayden for making my experience at RPI, and my graduation, possible.

I don't know where I would be without my loved ones. I am endlessly grateful to my parents for all that they've given me and all that they've sacrificed. I am lucky to have grown up with my grandparents as a second set of parents and an older sister I look up to. And I am luckier still to have such a wonderful extended family and to have Gianna in my life. To all of my friends and family across the U.S. and around the world, thank you for all of your love and support. I hope to see you soon.

In short, thanks to everyone above for putting up with me and all of my nonsense.

ABSTRACT

Collective decision making requires dealing with competing preferences. Preference aggregation problems are so ubiquitous that the reader probably needs little convincing that such problems can be complex and difficult to solve. There are many types of preferences, many different ways of expressing them, many different spaces of possible outcomes, and many different objectives to consider. While there is no one-size-fits-all method to aggregate preferences for all collective decision problems, we can design preference aggregation mechanisms using reasonable models of agent preferences and behavior and sensible measures of performance for specific classes of problems.

This dissertation is organized into three parts. Part I uses economic utility to describe agents' (numeric) preferences and to define objectives. The challenge here is that agents are only able to present limited information about their preferences. Part I focuses on implementable algorithms that take in agent preferences and produce an outcome that is provably close to maximizing utility when there is insufficient information to compute the optimal outcome. Chapter 4 looks at a class of problems where agents have preferences over one another, while Chapter 5 looks at agent preferences over inert alternatives.

Part II moves away from utilitarian reasoning and considers the interplay between two types of preference objects – agents and non-agent alternatives. Chapter 7 has agents who must each be assigned, or assign themselves, to a single alternative. The agents have preferences based upon both the alternative to which they are assigned and how many other agents are assigned to it. The central consideration of Chapter 7 is what outcomes are stable in the sense that agents cannot deviate profitably to yield an outcome they prefer. Chapter 8 has agents using elections and a method of proxy voting to empower other agents as representatives to make decisions on their behalf. Following Chapter 5, agents' preferences over potential representatives are based on the "distance" between them according to a certain metric. While in Chapter 5 an arbitrary metric is considered, Chapter 8 uses an explicit metric, and while the goal in Chapter 5 is to minimize the sum of agent costs, the goal in Chapter 8 is to select using indirect methods the same outcome that would be selected if all the agents voted directly without representatives.

Whereas the agents throughout Parts I-II only have preferences over non-agent alternatives, agents, or a combination of the two, the agents in Part III have a new type of

preference object. In Part III agents share the concerns that any mechanism designer might. The agents now care how collective decisions are made, not only about the outcome. Chapter 10 demonstrates that when agents agree on how decisions should be made, they can construct a set of self-regulating rules, or a constitution capable of amending itself. Lastly, Chapter 11 extends the traditional models of Social Choice to reveal that even when agents disagree on how decisions should be made rational collective choices are possible.

CHAPTER 1
INTRODUCTION

Life would be much simpler if everyone always agreed. There would be no arguments or politics. In reality, collective decisions give rise to conflicts. People, or *agents* in the broader sense, can hold divergent views about what is or about what ought to be. The following chapters are concerned with the latter; designing and analyzing procedures for making collective decisions when agents have competing *preferences* for how the world should be.

Preference aggregation procedures (functions) convert preferences (input) into outcomes (output). For any preference aggregation problem, an obvious starting point is to ask, "what constitutes a good outcome or makes one outcome better than another?" There are many approaches to answering this question. A standard approach in economic theory is to assume that all agents will experience a personal *utility* - a generic measure of goodness - determined by the outcome of the decision. Modeling agent preferences as numeric *cardinal utility functions* creates a way to express concepts like "good", "better", and "best" in quantitative terms. Concrete objectives can be specified in units of utility. The measure used in Part I to characterize the quality of outcomes is the sum of all agent utilities, also called the social welfare.

If all agents can reveal their cardinal utility functions, maximizing social welfare becomes an optimization problem. The trouble with the utility maximization approach just described is that these numeric values are an unnatural way for a person to express their preferences in many settings. In practice, people express their preferences in various ways, but usually not as a real-valued utility function. For example, it is more reasonable to assume that given a set of alternatives, agents can report orderings, or rankings, over the alternatives from most favorite to least favorite. The catch is that these *ordinal preferences* may not contain sufficient information to find a utility maximizing outcome. If we do not have enough information to determine which outcome is best, then we must resort to finding an outcome that is guaranteed to be at least good, or close to the best.

Mechanisms that guarantee performance close to the best that can be done are referred to as approximation algorithms. Approximation algorithms are valuable not because they give guarantees in absolute terms but because the outcomes they select are assured to be good relative to the best we could possibly do. Hidden in this statement is a subtle ambiguity.

What is the best we could do? Given a fixed set of available alternatives, one can compare the aggregate utility of any chosen outcome to the aggregate utility of the best alternative in that set. The results in Part I show that we can find outcomes provably close to optimal social welfare using minimal preference information. One of the major lessons of Part I is that the more information we have about agents' preferences, the better we can do in terms of social welfare. However, a deficit of preference information is not the only obstacle to preference aggregation and social welfare is not the only way to evaluate a decision outcome.

Suppose the students in a class are to be put into pairs to work on an assignment. If each student derives a utility based on who their partner is, we can try to find a matching that maximizes social welfare or approximates maximum social welfare. However, if students are able to change who their partners are the nature of the problem changes. If there are two students who are not paired with each other who each prefer the other to their current partner, they might leave their current partners and form a new pair. If a subset of agents can coordinate to alter the status quo in a way that they all prefer, we say that the status quo is *unstable*. When there is no central authority to impose an outcome, attention turns to finding and characterizing *stable* outcomes when they exist. In extreme cases of instability when agents associate freely a subset of disgruntled agents may break off, or fork, from a community to form a new community of their own. The problem of forking is the subject of Chapter 7 in Part II.

While social welfare and stability can characterize an outcome, whether a decision is good or bad can depend on more than just the outcome. One might care about *how* a decision is made. If a natural first question is about what constitutes a good outcome, the next question is logically, "what makes for a good decision other than a good outcome?" Overall, the properties of the decision making process may matter in ways not reflected solely in the outcome. For example, one may want a decision making rule to be non-dictatorial, meaning that there is no single agent such that the outcome is determined entirely by their individual preference, regardless of the preferences of the other agents. A desirable property, or *axiom*, in egalitarian settings is that all agents be treated equally, meaning that if the agents are relabeled but the collection of preferences remains otherwise the same the outcome should not change (e.g. one person, one vote). In Chapter 8 of Part II, this form of equality is the ideal to be emulated, rather than some intrinsic feature of the outcome.

The reason preference aggregation is hard is that we can't have everything we want.

There are objectives that cannot be pursued simultaneously and trade-offs that must be made. Non-utilitarian work in Social Choice Theory centers heavily on *impossibility theorems* that force us to choose between desirable properties when they cannot all be guaranteed. The most famous of these is Arrow's General Impossibility Theorem.

In Part I and Part II, we as mechanism designers have implicitly decided what properties matter (e.g., welfare maximization, stability, equality). In a sense this is inescapable. However, just as an agent who single-handedly determines the outcome of a collective decision is a dictator, an agent who single-handedly determines how preferences are to be aggregated becomes dictator of norms. In light of impossibility results, it must be admitted that reasonable people can disagree on what makes a good outcome or a good decision. We must acknowledge that the agents being modeled are not necessarily so different from us, and might care about any of the same aspects of decisions that we care about as mechanism designers. This point brings us to the culminating question of Part III: "How can agents with competing preferences over alternatives, agents, axioms, and/or rules come together to make decisions that everyone involved thinks are good or acceptable?"

If everyone agrees on what axioms or normative principles must be upheld, then there can be a "right" way to aggregate agent preferences to make collective decisions. Chapter 10 of Part III assumes some agreement among agents on what type of voting rule they should use. The agents must select a voting rule by which to make decisions, including the decision to amend the rule. In essence, the agents are agreeing upon axioms that specify a rudimentary *constitution* capable of aggregating their adverse preferences over rules.

Unfortunately, even moderate agreement about how decisions should be made cannot always be assumed. Finally, Chapter 11 of Part III introduces a framework for preference aggregation when agents do not agree on how decisions should be made or what makes for a good outcome and there may be no central authority to enforce outcomes. The key idea here is that agents with the same information can differ in both their preferences about what ought to be and their *counterfactual reasoning* about what could have been. The axioms of Arrow's General Impossibility Theorem are considered as a first application of this framework, followed by a new analysis of the constitution from Chapter 10.

In the end, whether in small groups or large societies, preference aggregation problems can be complex and difficult to solve. There is not always a single "best" outcome, and even when there is we may not have enough information to find it. According to a single set of

criteria there may be a "right" way to make a decision, but reasonable people can disagree on what criteria are important. Even if we can compute the "best" outcome the "right" way according to some chosen standard, it may not be enforceable and there's no guarantee the outcome will be stable. People might try to change the rules, or leave the community entirely. The chapters ahead do not offer a grand theory for preference aggregation to resolve these issues, but rather a set of approaches that one can take to designing and analyzing the means by which certain collective decisions can be made.

CHAPTER 2
BACKGROUND AND PRELIMINARIES

The problems and methods used in the following chapters span an array of fields within the disciplines of economics and computation, including game theory, social choice, algorithm design, mechanism design, computational complexity, voting theory, and democracy. The results in these chapters come from published and unpublished work [1]–[5]. The models, assumptions, and objectives therein are influenced by works in sociology, biology, philosophy, computation, and other fields [6]–[18]. To understand the chapters ahead and how they fit into the existing literature it helps to first have a light introduction to some of the concepts and terminology involved.

Firstly, throughout the following chapters there will be references to preference aggregation functions, rules, mechanisms, algorithms, procedures, and processes. Some preference aggregation functions that are specific to particular classes of problems have special names such as voting rules, approximation algorithms, and social welfare functions. These are all ways of referring to a process by which agent preferences are taken as input and an outcome is output. Their use depends on the type of problem at hand, although any distinctions are not of much concern here.

2.1 Problem Types

There are many different types of preference aggregation problems. It is convenient to characterize these problems first by what types of outcomes are sought and what options are available. The way in which we should expect agents to express their preferences, and the mechanism we choose to aggregate them, will depend on what type of outcome is desirable.

Subsets and Social Orders In a simple single-winner voting scenario, the agents are tasked with selecting a single winning element from a set. For example, they might be electing someone to hold public office from a field of candidates or alternatives. The space of possible outcomes is exactly the set of available candidates. A multi-winner voting problem is similar, except that the space of possible outcomes is now a collection of subsets of candidates in the power set. Multi-winner voting is appropriate for electing committees or parliaments. Rather than selecting a single winner or subset of winners, one might want to produce a relation

over a set of alternatives. There are many possible relations one might want to output, but perhaps the most commonly studied are orderings. One can take a set of alternatives and produce a social order, or ranking over the alternatives, from "best" to "worst" according to some criterion. One common application that requires producing orderings from agent preferences is in athletic competitions that require scoring from judges such as gymnastics, figure skating, and boxing.

Graph Formation Suppose a class of students are allowed to work with partners on a project. One could allow the students to choose their own partners, or they could express their preferences to the instructor to assign them to pairs. This is an example of a matching problem where the goal is to match students with a single partner. A convenient way to model this problem is as a complete graph where each student is a node and each edge is a possible partnership. The space of outcomes to the matching problem is all of the subgraphs in which no two edges are adjacent, so every agent has only one partner.

We can generalize from matching problems to consider the class of all problems that can be modeled as selecting a subgraph from a set of allowable subgraphs. These types of problems are referred to as graph formation problems or network formation problems. In the matching problem and similar graph formation problems the nodes of our graph represent the agents. We can generalize further to consider a broader range of problems when nodes can represent non-agent alternatives as well.

Suppose our class of students are to be divided up to perform various tasks in groups. Students have preferences for both the task they are assigned to perform and who it is that they will be working with. To model this graphically, create a set of nodes for the students and a set of nodes for the alternatives. An assignment corresponds to a collection of disjoint cliques (complete subgraphs containing all edges between all pairs of nodes inside the clique) in which each clique contains exactly one alternative.

If we go back to the problem of selecting a subset of alternatives we can model this graphically as well. Create a set of nodes for the agents and a set of nodes for the alternatives, and an edge between every agent-alternative pair. This forms a bipartite graph, consisting of two sets of nodes in which no node has an edge to any other node in its same set. If our goal is to select a single winner, then the space of outcomes is all subgraphs that consist of all the edges incident to a single alternative node. Similarly, if our goal is to select a subset

of winners, each possible outcome can be represented by a subgraph consisting of exactly all of the edges between the alternatives in that subset and the agents.

Alternatives vs. Outcomes In the simplest case, agents are faced with a set (or space) of possible outcomes and report their preferences directly over the possible outcomes. However, it is often unreasonable to expect agents to express their preferences over all possible outcomes. Sometimes the set of possible outcomes may be too large or complex for each agent to express their preferences over all possible outcomes (e.g. the set of all possible group task assignments). Agents might also care only about parts or aspects of the outcome relevant to them (e.g. with whom they get partnered). It is therefore important to differentiate between the *alternatives* over which agents express their preferences and the *outcomes* that are produced because these sets may not be the same, or even contain the same type of elements. The set of alternatives will usually be denoted by X and the set of outcomes by Y.

2.2 Preference Information

There are three basic categories into which agent preferences fall. The first category is cardinal. Cardinal preferences are numeric. In particular, agents may experience a numeric utility – a generic real-valued measure of goodness or happiness – for each possible outcome. Assuming that agents all have utilities in the same units makes their preferences directly comparable in quantitative terms. For each possible outcome y, each agent has a utility $u_i(y) \in \mathbb{R}$. In Chapter 4 the utility of all agents is non-negative. By contrast, in Chapter 5 each agent incurs a non-positive utility, referred to as a cost.

One form of cardinal preference information that does not reveal utilities directly is that of preference strengths. For example, an agent might report the ratio of their utilities between two possible outcomes $\frac{u_i(x)}{u_i(y)}$ without reporting their exact utility for either one.

The second category is ordinal preferences. Ordinal preferences are orders, or rankings, over some set of elements. Orders can be strict (a $\succ$ b) or weak (a $\succeq$ b). When ordinal information is available but not all pairs of elements can be compared the result is called a partial order. Notice that preference orders can always be deduced from cardinal utilities but not the other way around: $u_i(x) \geq u_i(y) \Rightarrow x \succeq_i y$.

The third and final category of preference information is categorical. Agents might

place a set of alternatives into different categories. For example, when there is only a single alternative presented as a proposal, an agent might approve of it or not. With multiple alternatives, an agent can partition a set of alternatives into disjoint sets based on whether they approve or do not approve of the alternative.

Incomplete Information We cannot always expect agents to report their full preferences. A running theme throughout the ensuing chapters will be that agents report some information about their preferences without being able to communicate their preferences in full. For example, an agent might possess cardinal utilities but only report their ordinal preferences induced by these utilities. There are a number of reasons why this would be the case. A common reason is that cardinal utility functions are too cumbersome or unnatural to express. For instance, it is obviously more natural for a person to say something like ,"I prefer A over B" than to say, "I have a utility of 8 for A and 5 for B." Even when cardinal utilities can be assumed to exist, this does not entitle us to assume that we have direct access to this information.

Another reason we may not have complete information about agent preferences is that the agents use some hierarchical decision structure, or system of representation, to make decisions. For example, shareholder proxy voting and representative democracies both involve some agents making decisions on behalf of others, so only the reported preferences of the representatives or proxies are known.

2.3 Evaluating and Comparing Outcomes

To design and compare preference aggregation mechanisms for a particular problem we need to have some criteria by which to evaluate and compare them.

Social Welfare and Social Cost When agents have cardinal utility functions, our primary concern is the maximization of *social welfare* or the sum of agent utilities. If each agent i has a utility of $u_i(y) \in \mathbb{R}$ for the outcome y, then the social welfare of y is $SW(y) = \sum_{\text{agents } i} u_i(y)$. We would like to find an outcome y that maximizes the social welfare. When agents experience costs instead of utilities we want to minimize cost rather than maximize it. In the realm of costs, our goal becomes to minimize the sum of agent costs or *social cost*: $SC(y) = \sum_{\text{agents } i} u_i(y)$.

Approximation Algorithms Sometimes it is impossible for us to find an optimal (e.g. welfare maximizing) outcome. In particular, we must deal with cases where we do not have sufficient information to determine what outcome is optimal. Even if we have the power to brute force search over all possible outcomes, we may not know enough about agents' cardinal preferences to determine the exact social welfare yielded by any outcome. For example, if agents report preference orderings over outcomes induced by their underlying cardinal utilities, we cannot determine from this alone what outcome maximizes utility. Suppose all agents but one prefer A to B, and the last agent prefers B to A, and we do not know their cardinal utilities. Does A or B yield a higher social welfare? It is tempting to say A, particularly when the number of agents is large, but this relies on assumptions about how large agent utilities are relative to one another.

When we do not have enough information to compute the optimal (e.g. welfare maximizing) outcome, our focus shifts instead to getting close to or *approximating* optimal. We can choose objectives other than maximizing social welfare and minimizing social cost. Whichever way we define OPT, as long as we can define some notion of similarity or distance between potential outcomes we can quantify how closely any outcome approximates optimal. In upcoming chapters we will adopt techniques from the study of *approximation algorithms* and improve upon them. Conversely, some of our results will be quite general and applicable to problems beyond preference aggregation.

Worst-Case Approximations We can characterize the quality of outcomes based on how close they are to optimal in terms of social welfare in relative terms. Namely, if OPT is the welfare maximizing outcome for some instance, we want to be able to compute an outcome y for that instance such that $\frac{SW(OPT)}{SW(y)}$ is bounded. The lower the bound the better because it means we are closer to optimizing social welfare in the worst case. Keep in mind that we want to design algorithms that are able to achieve these low approximation factors on all instances when we do not have enough information to determine what outcome is optimal.

Spatial Preferences One common way to model agent preferences is as points in a policy space. Every agent and alternative is represented by a point in space, and the distance between points $d(x, y) \geq 0$ is inverse to their similarity. Each agent's ideal alternative would be at the same point in space that they are, and the closer an alternative is to

them the more they like it. Agents will prefer alternatives that are closer to them in space $d(i,x) < d(i,y) \Rightarrow x \succ_i y$. In other words, the distance between an agent and an alternative represents the agent's cost if that alternative is selected as the outcome.

Metric Distortion When agents have spatial preferences, the distances are not arbitrarily. Typically they are assumed to obey the triangle inequality for any three points: $d(x,y) + d(y,z) \geq d(x,z)$. When the preference space obeys this inequality, it is called a metric space. For a given mechanism, the worst-case approximation factor it achieves on any instance (where the true distances are unknown) is called the *distortion* of that mechanism. The metric distortion of a mechanism is an upper bound on the ratio $\frac{SC(y)}{SC(OPT)}$ over all possible instances, where y is the selected outcome and OPT is the cost-minimizing outcome.

If the winning outcome for an instance I is y^*, then the approximation factor for this instance is

$$\delta_I = \frac{SC(x^*)}{SC(y^*)} \tag{2.1}$$

If for all instances I, a mechanism F always chooses an outcome such that δ_I is bounded above by some constant δ, we say that the mechanism F yields a δ-approximation to maximum Social Welfare, or that F has a distortion of δ.

$$\delta = \max_I \delta_I \tag{2.2}$$

Stability One important non-utilitarian characteristic of outcomes is stability. Recall the problem of matching. Suppose we have four students u, v, x and y and place them in two pairs (u,v) and (x,y). However, u and x would prefer to be paired with each other than with their current partner. If they have the freedom to change partners, then the original matching is *unstable* because u and x will leave their partners to pair with each other. For some problems a significant challenge is to design mechanisms that always return stable outcomes when they exist. Note that the social welfare optimizing solutions may not be stable for all problem instances.

2.4 Properties of Mechanisms

When designing mechanisms for preference aggregation problems, we may care about more than qualities of the outcome. When we care about *how* a decision is made, we must

pay attention to properties of the mechanism itself. Several such properties, or axioms, will be introduced in the chapters ahead. Two basic properties that are helpful to know up front are computational efficiency and strategyproofness.

Computational Efficiency Computational efficiency refers to how much computation needs to be done, or how many steps need to be taken, to solve a problem. In particular, the standard is to look at the worst-case asymptotic runtime as a function of the size of the input. Algorithms that run in time polynomial in the size of the input are said to be efficient, while algorithms whose runtime is exponential in the size of the input are inefficient.

Strategyproofness When agents are truthful, their reported preferences are induced by, and consistent with, their true preferences. When agents are strategic, this is no longer the case. In some situations, an agent might benefit by misreporting their preferences. For example, if voters can only vote for one of three candidates and a voter's favorite candidate has no chance of winning, they might vote for their second favorite candidate. This agent is said to be strategic, because they are not reporting their actual favorite. One desirable trait in mechanisms is that they are strategyproof, so no single agent can get an outcome they prefer by misreporting their preferences. If no subset of agents can collaboratively change their reported preferences to lead a mechanism to produce an outcome they all prefer, then the mechanism is said to be group strategyproof. Note that strategyproofness and stability can be parameterized by the number of agents who would need to cooperate to change the outcome in a way that they all prefer.

PART I

IMPLICIT UTILITARIAN MODELS OF PREFERENCE AGGREGATION

CHAPTER 3
CONTRIBUTIONS, BACKGROUND, AND RELATED WORK

The line of thought in Social Choice Theory following Arrow has historically focused on choosing which axiomatic properties to satisfy rather than quantitative measures of the quality of the solution. The motivation for this approach was historically rooted in part in the rejection of interpersonal comparisons of utility [19]. However, under the influence of Harsanyi [20] and Sen [6], [21], among others, utilitarian models with cardinal information have become increasingly popular. The measure of distortion was introduced by [22] in the context of spatial voting, (see also [23]–[26] for discussion). Since then, the distortion, or approximation factor, of various ordinal utility maximization mechanisms has been studied, particularly for matchings [27]–[35] and social choice [23], [24], [33], [36]–[54]. As a rule of thumb, the more information one has about agents' latent cardinal preferences, the better approximation to maximum utility one can obtain [25], [32]–[34], [44], [49], [55].

In many social choice problems, agents are assumed to have utility functions. An agent's utility function specifies their utility or cost for each possible outcome. Agents prefer outcomes for which they have a higher utility or lower cost. If agents were able to report their exact utility functions, then for any fixed objective one could determine the optimal

Portions of this chapter previously appeared as: B. Abramowitz and E. Anshelevich, "Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences," in *Proc. 32nd AAAI Conf. Artif. Intell.*, S. A. McIlraith and K. Q. Weinberger, Eds., Feb. 2018, pp. 894–901.

Portions of this chapter previously appeared as: B. Abramowitz and E. Anshelevich, "Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences - Full Version," Nov. 2017, arXiv:abs/1711.10588. (Retrieved on 7/18/2021).

Portions of this chapter previously appeared as: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," in *Web and Internet Econ. - 15th Int. Conf. Proc.*, vol. 11920, I. Caragiannis, V. S. Mirrokni, and E. Nikolova, Eds., Dec. 2019, pp. 3–16, doi: 10.1007/978-3-030-35389-6_1.

Portions of this chapter previously appeared as: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," Jun. 2019, arXiv:1906.10562. (Retrieved on 7/18/2021).

solution to any instance of a social choice problem by aggregating the utilities of all agents for each possible outcome, to compute the value of each outcome. For example, if the objective is to maximize the sum of agent utilities, or social welfare, one could sum up the utilities of all agents for each outcome and determine which outcome yields the highest total utility. Unfortunately, this brute force approach is rarely viable.

There are at least two reasons why the approach just described is unrealistic. The first reason, as mentioned above, is that there are far too many potential outcomes for an agent to specify a preference over all of them directly, so instead they express compact preferences over some set of alternatives that can be used to select an outcome. The second reason is that people typically do not report their preferences as numerical utility functions. It is unnatural, even bizarre, for a person to say something like, "I have a utility of 6 for option A and a utility of 4 for option B." It is much more common for a person to say, "I prefer A to B", "I prefer A much more than B", or even "I like A twice as much as B." Even when utilities can be assumed to exist, it is more likely that agents can report preference rankings over alternatives, a most preferred alternative, or even some information about preference strengths [56]–[60], but not their true utility function.

Chapter 4 and Chapter 5 use what are called implicit utilitarian models. Agents are assumed to have cardinal utility functions, or cost functions, and report preference information that is derived from and consistent with their true utilities. This derivative information is then used to compute an outcome that is guaranteed to be close to optimal.

3.1 Graph Formation Using Only Ordinal Information

In Chapter 4, agents express purely ordinal information. Each agent is represented as a node, and an edge between two nodes represents a connection between them. Every agent receives some non-negative utility for connecting to each other agent. Rather than reporting numerical utilities, agents only report a ranking, or (strict) preference ordering, over all of their neighbors. If nothing could be assumed about the utilities, or edge weights, then little could be said about how to compute a utility maximizing solution given constraints on what connections can be formed. However, if these utilities are symmetric, so both agents receive the same utility for forming an undirected edge between themselves, approximations to the maximum possible total utility can be guaranteed for a wide variety of problems using only ordinal information. Furthermore, the solutions our mechanism constructs are stable in the

sense that no two agents can mutually increase their utility if one or both of them give up some of their connections in order to form a connection with each other.

The model in Chapter 4 is unlike those for voting, as we consider graph formation problems where agent preferences are expressed over whether to connect with one another. Using a graph model, if one considers the set of all permissible graphs as the set of alternatives, then a centralized mechanism that outputs a graph is a social choice function. In this view, agents are expressing preferences over aspects of the outcome, rather than all of the possible outcomes. In the network formation setting, any pair of agents can agree to add a connection between them, possibly giving up others in the process. In a sense every pair of agents agent is "pivotal", and we care about the pairwise-stability of any solution.

In the context of matchings, [27], [28] develop various matching algorithms as a black-box to provide approximations for a variety of matching and clustering problems under the implicit utilitarian view. Additionally, [30], [31] provide results one-sided matchings and [29] consider bipartite matchings. However, *all previous work* on approximation for utility maximization mentioned above either assumes the underlying weights form a metric space [23], [27]–[29], [31], [36], [40], [42], [43] or are normalized [24], [30], [37]–[39], with only two exceptions. The first exception is that Maximum Traveling Salesperson yields a 2-approximation without the metric assumption [28]. We prove this result as part of a much more general theorem using much more general techniques. The second is the result developed in the full-information setting by [61] and affirmed in the ordinal setting by [27], that Max Weight Matching yields a 2-approximation without any assumptions on the weights. We generalize this result to all b-Matchings instead of only $b = 1$.

The work most similar to Chapter 4 is [27] which bounds the distortion of ordinal mechanisms for several problems, including Maximum Traveling Salesperson, but relies heavily on the assumption that the weights obey the triangle inequality. For the Ordinal Greedy algorithm featured in Chapter 4, [27] shows the metric assumption implies that any two edges which are both most preferred by their respective endpoints at some iteration must be within a factor of 2 of one another, even if they are not adjacent. By contrast, this non-local information is unavailable in the model of Chapter 4. Chapter 4 is remarkable because it identifies a large class of problems for which assumptions on the weights are unnecessary to achieve good approximations to optimum with only ordinal information.

The Omniscient Greedy algorithm, which has full information is referred to this way to

differentiate it from the Ordinal Greedy algorithm. The Omniscient Greedy algorithm has been studied extensively. In fact, it is known to be optimal on exactly the set of independence systems which are matroids, including Maximum Spanning Trees [62], [63]. Independence systems are introduced formally in Chapter 4 according to the heredity property that defines them. [64] showed that Omniscient Greedy provides good approximations for many independence systems, including matching and symmetric TSP. Dyer et. al. [65] further demonstrated that Omniscient Greedy provides a tight 3-approximation for Max Weight Planar Subgraph. These results were later reformulated as k-extendibility by [66], who applies this idea to a diverse set of problems, including b-Matching. Unfortunately, the proofs for all results just mentioned rely critically on Omniscient Greedy selecting edges in strictly non-increasing order, making them untenable in the ordinal setting. No ordinal algorithm can yield optimum solutions, even for matroids. However, as shown in Table 4.1, the results obtained here compete well with the best known polynomial-time algorithms for ABC Systems.

Randomized and Deterministic Mechanisms In Chapter 4, attention is restricted to deterministic social choice rules, instead of randomized ones as in e.g., [36], [38], [42], [50], for several reasons. First, consider looking at our mechanisms from a social choice perspective, i.e., as voting rules that need to be adopted by organizations and used in practice. People are far more resistant to adopting randomized voting protocols. This is because an election with a non-trivial probability of producing a terrible outcome is often considered undesirable, even if the *expected* outcomes are good. This may be due to people's intuitive grasp of non-ergodicity [67] or risk aversion. There are many exceptions to this, of course, but nevertheless deterministic mechanisms appear easier to convince people to adopt. Second, consider looking at our mechanisms from the point of view of approximation algorithms, i.e., as algorithms which attempt to produce an approximately-optimal solution given a limited amount of information. For traditional randomized approximation algorithms with guarantees on the quality of the expected outcome it is possible to run the algorithm several times, take the best of the results, and be relatively sure that you have achieved an outcome close to the expectation. In this setting of limited information, however, we cannot know the "true" cost of a candidate even after a randomized mechanism chooses it, and thus cannot take the best outcome after several runs. Therefore, unless stronger approximation

guarantees are given than simply bounds on the expectation, it is quite likely that the outcome of a randomized algorithm in our setting would be far from the expected value. While randomized algorithms are certainly worthy of study even in our setting, and many interesting questions about them exist, we choose to focus only on deterministic algorithms in this chapter.

Strategic Behavior In Chapter 5 we assume that the preference strengths given to our algorithms are truthful, i.e., that the agents are not strategic in reporting their preferences. While it would certainly be interesting and important to consider the case where agents are not truthful (as in e.g., [50], [68]), for many settings with preference strengths it is actually more reasonable to expect agents to be truthful than for settings with only ordinal votes. This is because preference strengths are often signaled passively (e.g., average response times to surveys) or expressing this intensity comes at a cost (e.g., time commitments, activism, or monetary contributions and payments). Even in debates and committees where a member signals their strong preference for A over B, this member is putting their reputation on the line in doing so, and so may not want to do this unless their preference is actually that strong, in order to not look foolish or inconsistent in the future.

Contributions Chapter 4 identifies a surprisingly broad class of graph formation problems for which a simple greedy mechanism using only ordinal information can provide small, constant-factor approximations to maximum social welfare in the worst case. The solutions created by this mechanism are shown to be pairwise stable. Finally, these bounds are shown to be tight, as no other mechanism can give a better worst-case approximation using the same information.

3.2 Metric Distortion

Whereas Chapter 4 considers agents' preferences over one another, Chapter 5 concerns agents' preferences over an inert set of alternatives. The set of agents is tasked with selecting a single winner (via voting) from a set of available alternatives. Once again, if we assume nothing about agents' utility functions, or cost functions, then little can be said about how to optimize social welfare.

Chapter 5 follows the literature on spatial voting, in which the agents and alternatives

are represented by points embedded in a common metric space [69]–[71]. The distance between an agent and an alternative represents their cost if that alternative is selected. Agents prefer alternatives with lower costs. The metric property requires that any three points in space satisfy triangle-inequality. For example, any d-dimensional Euclidean space is a metric space, so this model encompasses problems like voting on a facility location [50], [72]. Note, however, that not all metric spaces are Euclidean.

The metric assumption is a fairly general and modest assumption because all ordinal profiles are consistent with some metric space [73]. Consequently, the general metric assumption is not testable when eliciting only ordinal preferences. Either a subset of metrics must be hypothesized or more information needs to be reported.

In Chapter 5 agents report either their exact preference strengths or which of a set of ranges their preference strength falls in for each pair of alternatives. With preference strengths, the metric assumption is immediately falsifiable with three or more alternatives, as even a single agent's preferences could be shown not to correspond to any metric. A one should expect, the more information we have about agents' preferences, the better approximations to optimal we can likely give for whatever objective we choose.

When agents and alternatives are embedded in a metric, the worst-case ratio over all instances between the social cost of the solution returned by a mechanism and the optimal social cost for that instance is called the metric distortion [22]. In Chapter 5, we introduce a new concept called *ideal candidate distortion*. Metric distortion is based on the comparison between the cost of our solution and the cost of the best available alternative over all instances. By contrast, the ideal candidate distortion is based on comparison to the cost of the best conceivable alternative, even if it is not available as an alternative to choose or over which to express preferences. The key insight here is the trade-off between distortion and ideal candidate distortion for every instance. Namely, for certain mechanisms they cannot both be large. Using the mechanisms we propose, if our solution has high distortion in some instance, it can only be because the best alternative was close to the best conceivable alternative for that instance.

As discussed with graph formation, two main approaches have emerged for analyzing the distortion of various voting mechanisms. The first approach assumes that the underlying unknown utilities or costs are normalized in some way, as in e.g., [24], [38], [39], [68], [74], [75]. The second approach, taken in Chapter 5, assumes all agents and candidates are

points in a metric space, as in e.g., [33], [40], [42], [43], [46]–[53], [76], [77]. In particular, when the hidden numerical costs that induce agent preferences over a set of candidates obey triangle inequality, simple deterministic voting rules yield distortion which is always at most a small constant. It was shown by [23] that no deterministic mechanism can always produce distortion better than 3. For many years it was an open question whether any mechanism could achieve this lower bound. An elegant mechanism that achieves this lower bound was developed by [53], closing the long-standing gap. It is also worth noting that the well-known Copeland rule, Ranked Pairs, and the Schulze rule all achieve a distortion of 5 [23], [40]. Another recent work worth noting explicitly is a recent one which considers a representative voting mechanism where agents are all located on the real-line [72].

Preference Strengths Attempts to exploit preference strength information have led to various approaches for modeling, eliciting, measuring, and aggregating people's preference intensities in a variety of fields, including Likert scales, semantic differential scales, sliders, constant sum paired comparisons, graded pair comparisons, response times, willingness to pay, vote buying, and many others (see [78]–[80] for summaries). In our work we specifically consider only a small amount of coarse information about preference strengths, since obtaining detailed information is extremely difficult. Intuitively, any rule used to aggregate preference strengths must ask under what circumstances an 'apathetic majority' should win over a more passionate minority [81], and in Chapter 5 we provide a partial answer to this question when the objective is to minimize distortion.

Perhaps most related to our work is that of [36] which introduced the concept of *decisiveness*. Using our notation, [36] proves bounds on distortion under the assumption that *every* agent has a preference strength at least α between their top and second-favorite candidates. We, on the other hand, do not require that agents have any specific preference strength between any of their alternatives, and provide general mechanisms and distortion bounds based on knowing a bit more about agents (arbitrary) preference strengths. In other words, while [36] limits the possible space of agent preferences and locations in the metric space, we instead allow those to be completely arbitrary, but assume that we are given slightly more information about them.

Contributions Chapter 5 introduces the concept of ideal candidate distortion. Ideal candidate distortion is shown to be bounded as a function of the distortion. Moreover, the ideal candidate distortion for each instance is shown to be bounded as a function of the distortion for that instance. For the mechanism we propose the distortion and ideal candidate distortion cannot both be large, because the distortion can only be large when the best available candidate is close to the ideal conceivable candidate.

CHAPTER 4
GRAPH FORMATION USING ONLY ORDINAL INFORMATION

4.1 Introduction

We consider a relatively general network formation setting. All problems considered herein are modeled by an undirected complete graph $G = (N, \mathcal{E})$, where the (symmetric) weight $w(x, y)$ of each edge $(x, y) \in \mathcal{E}$ represents the hidden utility of connecting agents $x, y \in N$. The goal is to form a maximum-weight (i.e., maximum utility) graph which obeys some given constraints. For example, constraints may include bounds on the maximum degree, on component size, and many others. This framework includes such problems as matching, group formation, TSP, and far more as special cases; see Section 4.2 for example constraints and how they lead to different important settings.

In many settings we may not have access to the true edge weights $w(x, y)$. Instead, each agent $x \in N$ reports a strict *preference ordering* over the other agents $N - \{x\}$ that it can be connected with. We assume this ordering to be consistent with the latent weights, so that $w(x, y) > w(x, z)$ implies that x prefers y to z. While it is clearly impossible to form an optimal (i.e., maximum-utility) solution without any direct knowledge of the edge weights, we show how to design good approximation algorithms for selecting a maximum weight subgraph of G, subject to a large set of constraints. As usual in this line of work, the measure of performance is simply the sum of the agent utilities. This chapter provides good approximations for a broad class of ordinal analogues to graph optimization problems representing utility maximization for multi-agent systems. Note that unlike *all* previous work mentioned here, we do not make additional assumptions on the structure of the edge weights.

This chapter previously appeared as: B. Abramowitz and E. Anshelevich, "Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences," in *Proc. 32nd AAAI Conf. Artif. Intell.*, S. A. McIlraith and K. Q. Weinberger, Eds., Feb. 2018, pp. 894–901.

This chapter previously appeared as: B. Abramowitz and E. Anshelevich, "Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences - Full Version," Nov. 2017, arXiv:abs/1711.10588. (Retrieved on 7/18/2021).

We *do not* assume either that the agent utilities are normalized (as in [24], [30], [37]–[39]), nor that they form a metric space (as in [23], [27], [28], [31], [36], [40], [42], [43]). Thus our results demonstrate how well one can perform using only ordinal information, without additional assumptions.

ABC Systems We define an expansive class of constraints called *ABC Systems*. The maximization problem for an ABC System is to compute a maximum weight subgraph $S \subseteq \mathcal{E}$ of an undirected, complete graph $G = (\mathcal{N}, \mathcal{E})$ such that in S every node has degree at most b, every connected component has size (number of nodes) at most c, and $S \in \mathcal{A}$ for some arbitrary **attachment set** $\mathcal{A}$. A collection of subgraphs $\mathcal{A}$ is an attachment set of G if the following properties hold for all subgraphs $F \subseteq \mathcal{E}$:
1) *Heredity:* If $F \in \mathcal{A}$ and $F' \subseteq F$ then $F' \in \mathcal{A}$.
2) *Attachment:* If $F \in \mathcal{A}$ and $F + e \notin \mathcal{A}$ for some $e = (u, v) \notin F$, then there is a (u, v)-path in F. [1]

The utility maximization objective for ABC Systems encompasses a wide variety of well-known problems central to algorithm design. The examples we address in this chapter include Max Weight b-Matching, Maximum Weight Spanning Tree, Maximum Traveling Salesperson, and Max Weight Planar Subgraph. Our results also encapsulate many other interesting optimization problems for ABC Systems which we will not discuss directly, like finding the maximum weight subgraph with minimum girth k, maximum cycle length l, or which excludes a variety of graph minors (including all 2-edge-connected minors). As we show, all such problems are amenable to knowing only secondhand ordinal information, instead of the true numerical utilities.

4.2 Model and Problem Statements

The input for all problems in this chapter is a set N of agents (nodes) of size n, and a strict preference ordering for each $x \in N$ over the edges adjacent to x. The preference orderings reported by each agent are induced by a set of hidden symmetric weights $w(x, y) = w(y, x)$ for all $x, y \in N$. The set of hidden weights corresponds to an undirected, complete graph $G = (N, \mathcal{E})$ with non-negative weight function $w : \mathcal{E} \to \mathbb{R}^+$. The transitive relation of the individual preference orderings for all agents determines a partial ordering σ over all

[1] We use the "+" and "-" notation when adding or removing a single edge to or from a set.

edges; note that some pairs of edges may end up being incomparable in σ (see Example 1). The preference ordering σ is said to be **consistent** with the hidden weights if $\forall x, y, z \in N$, if x prefers y to z then it must be that $w(x, y) \geq w(x, z)$. If an edge e_1 is known to be at least as large as edge e_2 according to this partial ordering σ, then we will say that e_1 *dominates* e_2 in σ. The problems we consider are optimization problems where for an instance (w, σ) the objective is to compute the subgraph of G with maximum total edge weight, subject to a set of constraints, knowing only σ.

For weight function w, we let OPT_w be the optimal solution for the weights prescribed by w, and we let $w(OPT_w)$ be the total weight of the optimal solution evaluated by w. Likewise, we use S to denote our constructed solution and $w(S)$ its weight. Our approximation factor for a problem is therefore $\alpha = \max_{(w,\sigma)} \frac{w(OPT_w)}{w(S)}$, where S is any solution returned by our algorithm for (w, σ).

Recall the definition of ABC Systems. Given constraints $\mathcal{A}, b, c$ we can say without loss of generality that $b \leq c-1$, because if any node has c or more neighbors in a component, this component would have to be of size greater than c. Therefore, $1 \leq b < c \leq n$. When $b = c-1$ this effectively removes the node degree constraint. Similarly, when $c = n$, the component size is effectively unbounded. Likewise, when $\mathcal{A} =$ all subgraphs of G, this effectively annuls the attachment set constraint. Some specific problems which we consider in this chapter are as follows.

Max ABC: $c \leq n$, $b < c$, $\mathcal{A} =$ any attachment set of G

Max AB: $c = n$, $b < c$, $\mathcal{A} =$ any attachment set of G

Maximum Spanning Tree: $c = n$, $b = c - 1$, $\mathcal{A} =$ all acyclic subgraphs of G

Maximum Traveling Salesperson: $c = n$, $b = 2$, $\mathcal{A} =$ all subgraphs of G without non-Hamiltonian cycles

Max Weight Planar Subgraph: $c = n$, $b = c - 1$, $\mathcal{A} =$ all planar subgraphs of G

Max Weight b-Matching: $c = n$, $b < c$, $\mathcal{A} =$ all subgraphs of G

Max Weight Matching: $c = n$, $b = 1$, $\mathcal{A} =$ all subgraphs of G

Table 4.1: New Ordinal Greedy results compared to known results for Omniscient Greedy and the best known polynomial-time algorithm with full-information. Ordinal Greedy bounds are tight except for Max Planar Subgraph.

Maximization Problem	Ordinal Greedy	Omniscient Greedy	Best Known
ABC System	$b+1$	$b+1$	-
AB System	$d+1$	$d+1$	-
Spanning Tree	2	1	1
Planar Subgraph	4	3	72/25 [82]
Traveling Salesperson	2	2	9/7 [83]
b-Matching	2	2	1
Matching	2	2	1

4.3 Contributions

Most algorithmic techniques for maximizing utility from the full-information setting do not translate to the ordinal information setting [27]. These typically rely on non-local information, like comparisons between weights of non-adjacent edges, or comparing the total weights of sets of edges. This is not possible using only ordinal information. Even the important (and well-studied [84]) *Omniscient Greedy* algorithm, which simply adds edges in strictly decreasing order of their weight, is not possible to execute using only ordinal information. Instead, we mostly consider the natural *Ordinal Greedy* algorithm (defined in Section 4.4.1), which adds edges iteratively as long as the edge (x, y) being added is the most preferred edge for both x and y out of all the possible edges which could be added at that time. Ordinal Greedy has some very nice properties: in addition to being natural and providing high-utility solutions (as we prove in this chapter), it also always creates *pairwise stable* solutions: no pair of agents would have incentive to destroy some of their links and form a new link connecting them. Note that the performance of Ordinal Greedy can be very different from Omniscient Greedy: see Example 1 in Section 4.4.1 for intuition of why this must be.

In this chapter, we analyze the performance of Ordinal Greedy for many ABC Systems (see Table 4.1). We first prove that for general ABC systems, Ordinal Greedy always produces a solution with weight at most factor $b+1$ away from optimum, and that this factor is tight. In other words, for general problems including all the ones in Table 4.1, as long as the number of connections for each node must be bounded by some small b, then using only

ordinal information it is possible to compete with the best possible solution, and thus with any algorithm which knows the true numerical utilities. Such results tell us, for example, that when b is small, there is no need to attempt to find out the hidden edge weights/utilities; just knowing the ordinal preferences is good enough.

Second, we show that by relaxing the component size constraint (i.e., setting c to be unbounded) we can achieve significant improvements. For convenience, we call ABC systems with the component size constraint relaxed *AB Systems*. We prove that as long as the *sparsity* of any solution formed by such an AB system is guaranteed to be at most d, then Ordinal Greedy always forms a solution within a factor of $d + 1$ from optimum. Since the sparsity of a graph is at most half the average degree in any subgraph, we know that $d \leq \frac{b}{2}$. Therefore an AB System is at worst $(\frac{b}{2} + 1)$-approximable, giving us a factor of 2 improvement over general ABC systems. This result is more powerful than it may first appear, as many important constraints yield solutions with small sparsity. For example, since all tours and trees have sparsity at most 1, Ordinal Greedy provides a 2-approximation for both Maximum Traveling Salesperson and Maximum Weight Spanning Tree. And since all planar graphs have sparsity less than 3, we obtain a 4-approximation for Maximum Weight Planar Subgraph which uses only ordinal information. Note that the result for TSP was previously shown in [28]; we prove this result as part of a much more general theorem using much more general techniques.

Lastly, we consider Maximum b-matching, in which the only constraint is that each agent can be matched with at most b others. This is simply an ABC system with unbounded c and $\mathcal{A}$ being all possible subgraphs. We prove that our approximation factor drops to a constant 2, regardless of the value of b. This generalizes the result provided by [27] for the case of Maximum Weight Matching, where $b = 1$.

To prove the above ordinal approximations, we first demonstrate that for any ABC system (and any system with the independence property defined earlier) Ordinal Greedy achieves its worst approximation on an instance with weight function $w : \mathcal{E} \to \{0, 1\}$. We use this fact heavily to establish our approximation bounds, and believe it to be of independent interest. Note that similar results for Omniscient Greedy have relied critically on the fact that it selects edges in strictly non-increasing order by weight. Clearly this does not and cannot hold for the ordinal setting, as it is even possible for the minimum weight edge of the graph to be selected before the maximum weight edge. Because of this, our proofs require

completely new approaches and techniques.

4.4 Algorithmic Framework

In this section we define the Ordinal Greedy algorithm and reveal some of its salient properties. Rather than limit ourselves only to ABC Systems, in this section we consider general graphic *independence systems*. An independence system for our setting is a pair $(\mathcal{E}, \mathcal{L})$ where $\mathcal{E}$ corresponds to the set of edges in some graph and $\mathcal{L}$ is a collection of subsets of $\mathcal{E}$ such that if $F \in \mathcal{L}$ and $F' \subseteq F$ then $F' \in \mathcal{L}$. The sets in $\mathcal{L}$ are called independent. It is easy to see that all ABC Systems are independence systems. Let $\mathcal{B}$ denote the set of all subgraphs in which all nodes have degree at most b and let $\mathcal{C}$ denote the set of all subgraphs in which all connected components have size at most c. Our Max ABC problem can be restated as: Given a graph $G = (N, \mathcal{E})$, attachment set $\mathcal{A}$, degree limit b and component size limit c, compute the maximum weight subgraph in $\mathcal{L} = \mathcal{A} \cap \mathcal{B} \cap \mathcal{C}$.

4.4.1 The Ordinal Greedy Algorithm

In an ordinal setting, algorithms only have access to a partial ordering, or set of preference orderings, which provide strictly local information about the preferences of each agent. This precludes the use of algorithms which require comparisons between the weights of non-adjacent edges. In fact, it is not difficult to see that no ordinal algorithm can be guaranteed to compute the optimal solution for even simple settings, e.g., forming a matching [27]. However, the Ordinal Greedy algorithm defined below performs well in this setting because it relies on strictly local information. Ordinal Greedy starts from the empty set and builds up a sequence of intermediate solutions by adding locally optimal edges at each iteration which do not violate a set of constraints, i.e., preserve independence. To understand how this heuristic is applied to the ordinal setting, we must formalize what it means for an edge to be locally optimal.

Definition 1. Undominated Edge
Given a set E of edges, $(u, v) \in E$ is undominated if for all (u, x) and (v, y) in E, $w(u, v) \geq w(u, x)$ and $w(u, v) \geq w(v, y)$.

At this point it is important to make several observations. First, every edge set E has at least one undominated edge, because its maximum weight edge must be undominated.

However, there may be undominated edges which are not globally maximum. Second, for any edge set E it is straightforward to find at least one undominated edge using only the partial ordering σ (see [27] for details). Undominated edges are either of the form (u, v) where u and v are each other's most preferred neighbor, or form cycles in which each subsequent node is the first choice of the previous one, and thus all edges in the cycle have the same weight.

What follows is a general purpose Ordinal Greedy algorithm, which starts from the empty set and iteratively selects undominated edges from the set of remaining edges which do not violate the constraints in question. The algorithm uses the partial ordering σ to determine which edges are undominated at each iteration. The algorithm concludes when there are no edges left which can be added to the subgraph without violating the constraints, so the final solution S is maximal in this sense.

Algorithm 1: Ordinal Greedy

Input: Edge set $\mathcal{E}$, partial ordering σ, collection of valid subgraphs $\mathcal{L}$
Initialize $S = \emptyset$, $E = \mathcal{E}$;
while $E \neq \emptyset$ **do**
  Pick an undominated edge $e = (u, v) \in E$ and add it to the intermediate solution: $S \leftarrow S + e$;
  Remove e from E ;
  Remove all edges f from E such that $S + f \notin \mathcal{L}$;
end
Output: Return S

We refer to the iteration at which an edge $e = (u, v)$ is removed from E as the **critical iteration** of e. When the inputs to our algorithm $(\mathcal{E}, \mathcal{L})$ characterize an ABC System, there are exactly four cases which may occur at the critical iteration of edge e:

1) e is added to the ordinal greedy solution S

2) e is removed from E because $S + e \notin \mathcal{A}$

3) e is removed from E because $S + e \notin \mathcal{B}$ (where $\mathcal{B}$ = sets of edges with any degree $\leq b$)

4) e is removed from E because $S + e \notin \mathcal{C}$ (where $\mathcal{C}$ = sets of edges with any component size $\leq c$)

For cases 2-4 we say e was **eliminated due to $\mathcal{A}, \mathcal{B}$, or $\mathcal{C}$**. If an edge $e = (u, v)$ was eliminated due to $\mathcal{A}$, the attachment property implies there must be a (u, v)-path in the intermediate solution at its critical iteration. In other words, u and v are already in the same connected component in S at this iteration. If $e = (u, v)$ was eliminated due to $\mathcal{B}$, either u

or v must already have degree exactly b at this iteration. If $e = (u, v)$ was eliminated due to $\mathcal{C}$, u and v must already be in disjoint connected components whose cumulative size is greater than c. Note that in these three cases, an edge can only be eliminated if at least one adjacent edge of equal or greater weight has already been added to the intermediate solution, and all adjacent edges already added to the intermediate solution must be of equal or greater weight (since only undominated edges are added to our solution).

There are limiting values of $\mathcal{A}$, b, and c for which elimination due to these constraints cannot occur. When $c = n$, no edge can be eliminated due to $\mathcal{C}$ because all nodes can be in the same connected component. Additionally, when $\mathcal{A}$ is the set of all subgraphs of G, no edge can be eliminated due to $\mathcal{A}$. If adding an edge would violate more than one constraint, we say that it was eliminated in order of priority $\mathcal{C} \to \mathcal{B} \to \mathcal{A}$. For example, when $b = c - 1$, no edge can be eliminated due to $\mathcal{B}$, because for a node to reach degree b the size of its component must be exactly c and we say that any incident edge would be eliminated due to $\mathcal{C}$. As the following sections show, the approximation factor of Ordinal Greedy for an ABC System depends on which cases of elimination can occur.

Notice that the performance of the Ordinal Greedy algorithm can deviate significantly from the Omniscient Greedy algorithm in the full-information setting (which we call "Omniscient Greedy" because it knows the underlying edge weights and can choose the edge with maximum weight at each iteration). Consider the following example.

Example 1. Suppose the graph $G = (N, \mathcal{E})$ is constructed as follows.
Let $N = \{u_1, ..., u_k, v_1, ..., v_k\}$. Let $w(u_i, v_i) = 1 + \epsilon$ for $i \leq k$ for some infinitesimal ϵ. Let $w(u_i, u_j) = 1$ for all $i \neq j$. Let $w(v_i, v_j) = \epsilon$ for all $i \neq j$. Let all other edges have weight 0. Consider the ABC System corresponding to finding a Maximum-Weight Spanning Tree. It is clear that Omniscient Greedy will find the optimum solution with weight $w(OPT_w) = k(1 + \epsilon) + (k - 1)$.

Now consider Ordinal Greedy. Suppose Ordinal Greedy begins by selecting (u_i, v_i) for $i \leq k$, which are all undominated at the beginning of the algorithm. Once these edges have been selected, edges of the form (u_i, u_j) and (v_i, v_j) become undominated for $i \neq j$. Now, if an edge (u_i, u_j) or (v_i, v_j) is selected, the other must be eliminated at that iteration, since taking it would form a cycle. Notice that since we only have access to ordinal information, there is *no possible way* for Ordinal Greedy to tell which of these edges is better: they are both edges which are most preferred by their endpoints, even though one secretly has weight

1 and the other only ϵ. In other words, these edges are incomparable in the partial preference order σ. Suppose Ordinal Greedy proceeds by selecting (v_i, v_{i+1}) for $i < k$. Then the Ordinal Greedy solution formed has weight $w(S) = k(1+\epsilon) + (k-1)\epsilon$. This example shows that (in the limit) it is not possible for Ordinal Greedy to always result in solutions better than a factor of 2 away from optimum, even though Omniscient Greedy can easily compute the true optimum solution. As we show in this chapter, however, despite its knowledge handicap, Ordinal Greedy can often produce surprisingly good results.

4.4.2 Properties of Ordinal Greedy

For any independence system in the full-information setting, the Omniscient Greedy algorithm has been shown to achieve its worst approximation on an instance with a **binary weight function** $\bar{w} : \mathcal{E} \to \{0, 1\}$ [64]. However, previous proofs have relied crucially on the fact that Omniscient Greedy selects edges in strictly non-increasing order by weight, which is not possible with only ordinal information. We offer a new proof to show that even in the ordinal setting, Ordinal Greedy always achieves its worst approximation factor on an instance with a binary weight function $\bar{w} : \mathcal{E} \to \{0, 1\}$ for any graphic independence system. This theorem will allow us to prove approximation bounds for ABC and AB Systems later in this chapter.

Theorem 1. *For any graphic independence system $(\mathcal{E}, \mathcal{L})$, for any instance (w, σ) with weight function $w : \mathcal{E} \to \mathbb{R}^+$ and partial ordering σ consistent with w, there exists an instance $(\bar{w}, \sigma)$ with weight function $\bar{w} : \mathcal{E} \to \{0, 1\}$ such that σ is consistent with $\bar{w}$ and the worst-case ratio of the optimal solution to an Ordinal Greedy solution is at least as large as for (w, σ).*

Proof Sketch 1. We begin by providing a short proof sketch. Suppose on instance (w, σ) the ratio between the optimal solution OPT_w and solution S constructed by Ordinal Greedy is $\frac{w(OPT_w)}{w(S)} = \delta$. Our goal is to construct a binary weight function $\bar{w}$ such that $\frac{\bar{w}(OPT_{\bar{w}})}{\bar{w}(S)} \geq \delta$. When δ is infinite constructing $\bar{w}$ is straightforward, so we only consider finite values of δ. First we create a weight function $\hat{w}$ by raising the weights of all edges not in S as much as possible without altering the weights of the edges of S, such that σ remains consistent with $\hat{w}$. Since Ordinal Greedy selects S and none of its edge weights have changed, and the edge weights of OPT_w cannot have decreased, then $\frac{\hat{w}(OPT_w)}{\hat{w}(S)} \geq \delta$. From $\hat{w}$ we carefully create $\bar{w}$ by proving that there must exist a subset of edges to which we can assign weight 1 and let all other edges have weight 0, such that σ is consistent with $\bar{w}$ and $\frac{\bar{w}(OPT_{\bar{w}})}{\bar{w}(S)} \geq \frac{\bar{w}(OPT_w)}{\bar{w}(S)} \geq \delta$.

Proof. Recall that a partial ordering σ is *consistent* with weight function w if for all $x, y, z \in N$, if x prefers y to z in σ then $w(x, y) \geq w(x, z)$. We will now show that for any instance (w, σ) where $w : \mathcal{E} \to \mathbb{R}^+$ for which Ordinal Greedy provides a δ-approximation for $\delta > 0$ in the worst case, there exists an instance $(\bar{w}, \sigma)$ where $\bar{w} : \mathcal{E} \to \{0, 1\}$ for which Ordinal Greedy provides no better than a δ-approximation in the worst case. Given any weight function w we now construct a binary weight function $\bar{w}$ such that the approximation factor is at least as large and σ is still consistent with $\bar{w}$.

Observation 1. *For an independence system $(\mathcal{E}, \mathcal{L})$ the solution S computed by Ordinal Greedy depends only on σ, not the edge weights. Therefore, if σ is consistent with w and $\hat{w}$, the possible solutions S are the same for instances (w, σ) and $(\hat{w}, \sigma)$. However, $\hat{w}(S)$ and $w(S)$ may differ.*

Therefore, given a worst possible solution S constructed by Ordinal Greedy for (w, σ), our goal is to take the weight function w and construct a binary weight function $\bar{w}$ such that $\frac{w(OPT_w)}{w(S)} \leq \frac{\bar{w}(OPT_{\bar{w}})}{\bar{w}(S)}$, and σ is still consistent with $\bar{w}$ (and thus S can still be produced by Ordinal Greedy for the instance with weights $\bar{w}$). Recall that OPT_w is the optimum (maximum-weight) solution for weights w. Since by definition $\bar{w}(OPT_{\bar{w}}) \geq \bar{w}(OPT_w)$, it is enough to show that $\delta = \frac{w(OPT_w)}{w(S)} \leq \frac{\bar{w}(OPT_w)}{\bar{w}(S)}$.

Suppose there is some edge $e \notin S$, such that no edge in S is known to be greater than or equal to it in the partial ordering σ. If such an edge exists, we can let $\bar{w}$ be the weight function such that e and all edges known to be greater than or equal to e in σ have weight 1, and all other edges have weight 0. Clearly, σ remains consistent with $\bar{w}$. To see this, consider any $x, y, z \in N$ such that x prefers y to z. Then either both $w(x, y)$ and $w(x, z)$ dominate e in σ, so $\bar{w}(x, y) = \bar{w}(x, z) = 1$, or neither do, in which case $\bar{w}(x, y) = \bar{w}(x, z) = 0$. The only other case is that $w(x, y)$ dominates e and $w(x, z)$ does not (since we know that $w(x, y)$ dominates everything that $w(x, z)$ does), and then $1 = \bar{w}(x, y) > \bar{w}(x, z) = 0$. In all cases, $\bar{w}(x, y) \geq \bar{w}(x, z)$, so σ is consistent with $\bar{w}$. This means S remains the same set of edges, and since $\bar{w}(S) = 0$, the approximation factor becomes unbounded. This means for any instance where there is some edge $e \notin S$ for which no edge in S dominates it in σ, we can always create a weight function $\bar{w}$ with an approximation factor at least as large. Therefore, for the rest of the proof we assume that for every edge $e \notin S$ there exists some edge in S known to be at least as great by the partial ordering σ.

For our greedy solution S, fix an ordering $\{s_1, s_2, ..., s_m\}$ over the edges of S in non-increasing order by weight so that $w(s_1) \geq w(s_2) \geq ... \geq w(s_m)$. Construct weight function $\hat{w}$ by increasing the weight of each edge not in S to be equal to the weight of the smallest-weight edge $s_i \in S$ known to be greater than or equal to it in the partial ordering σ. Note that by our assumption above, such an edge s_i always exists.

Claim 1. *σ is consistent with $\hat{w}$.*

Proof. Consider any two adjacent edges (x, y) and (x, z) where x prefers y to z. Let $(u, v) \in S$ be a smallest edge known by σ to have weight at least $w(x, y)$. Then (u, v) is also known to have weight at least $w(x, z)$ since x prefers y to z. Therefore, the smallest edge of S known to have weight at least $w(x, z)$ is either (u, v) or has weight smaller than $w(u, v)$. After increasing the weights, $\hat{w}(x, y) = w(u, v) \geq \hat{w}(x, z)$. Therefore if x prefers y to z in σ, then $\hat{w}(x, y) \geq \hat{w}(x, z)$. By definition, σ remains consistent with $\hat{w}$. □

The above process of forming $\hat{w}$ forms an assignment of edges: consider every edge e to be assigned to the smallest edge in $s_i \in S$ known to be larger than or equal to it in the partial ordering, where $\hat{w}(e) = \hat{w}(s_i)$. Now alter these assignments so that if $\hat{w}(s_i) = \hat{w}(s_{i+1})$ then all edges with this weight, including s_i, are assigned to s_{i+1}. Let r_i be the number of edges of OPT_w assigned to s_i. Note that if $\hat{w}(s_i) = \hat{w}(s_{i+1})$ then $r_i = 0$.

Lemma 1. $\sum_{i=1}^{m} r_i \cdot \hat{w}(s_i) \geq \delta \sum_{i=1}^{m} \hat{w}(s_i)$

Proof. On the left side of the inequality, the product $r_i \cdot \hat{w}(s_i)$ denotes the total weight of the edges of OPT_w assigned to s_i after having their weight increased. This sum over $i \leq m$ computes $\hat{w}(OPT_w)$, the total weight of the optimal solution over w, evaluated by $\hat{w}$. On the right hand side, the summation yields the total weight of the edges in the greedy solution, $\hat{w}(S)$ multiplied by δ. Since σ is consistent with $\hat{w}$ by Claim 1 the greedy solution S remains the same, and since none of the weights of edges of S were altered $\hat{w}(S) = w(S)$. By construction, $\hat{w}(OPT_w) \geq w(OPT_w)$ because the weights of edges of OPT_w could only have been increased. And so $\hat{w}(OPT_w) \geq \frac{w(OPT_w)}{w(S)} \hat{w}(S) = \delta \cdot \hat{w}(S)$. □

We now demonstrate that we can alter the weights of $\hat{w}$ to create a binary weight function $\bar{w} : \mathcal{E} \to \{0, 1\}$ such that $\bar{w}(OPT_w) \geq \delta \cdot \bar{w}(S)$. All changes to the weights keep σ consistent with $\bar{w}$, so that S remains a solution of Ordinal Greedy.

Lemma 2. *There exists some $k \leq m$ such that $\sum_{i=1}^{k} r_i \geq \delta k$.*

Proof. Suppose to the contrary that $\sum_{i=1}^{k} r_i < \delta k$ for all $k \leq m$. We show by induction that this implies $\sum_{i=1}^{m} r_i \hat{w}(s_i) < \delta \sum_{i=1}^{m} \hat{w}(s_i)$, which yields a contradiction to Lemma 1. Specifically, we will show that for every $j \leq m$,

$$\sum_{i=1}^{j} r_i \hat{w}(s_i) \leq \delta \sum_{i=1}^{j-1} \hat{w}(s_i) + [\sum_{i=1}^{j} r_i - \delta(j-1)]\hat{w}(s_j). \tag{4.1}$$

When applied to $j = m$, this gives us the result that $\sum_{i=1}^{m} r_i \hat{w}(s_i) \leq \delta \sum_{i=1}^{m-1} \hat{w}(s_i) + [\sum_{i=1}^{m} r_i - \delta(m-1)]\hat{w}(s_m)$. Since $\sum_{i=1}^{m} r_i < \delta m$, then $\sum_{i=1}^{m} r_i - \delta(m-1) < \delta$, and thus the right hand side of the above inequality is strictly less than $\delta \sum_{i=1}^{m} \hat{w}(s_i)$, which gives us a contradiction with Lemma 1, as desired. Note that here we use the fact that $\hat{w}(s_m) > 0$ without loss of generality; if this were not the case then we can make the same argument for j being the largest integer such that $\hat{w}(s_j) > 0$. Thus all that is left is to prove Inequality (4.1).

We proceed by induction. The base case for $j = 1$ is trivially true. Now assume that Inequality (4.1) holds for j, and we will prove it for $j + 1$. Then,

$$\sum_{i=1}^{j+1} r_i \hat{w}(s_i) \leq \delta \sum_{i=1}^{j-1} \hat{w}(s_i) + [\sum_{i=1}^{j} r_i - \delta(j-1)]\hat{w}(s_j) + r_{j+1}\hat{w}(s_{j+1}) \tag{4.2}$$

Let $\xi = \delta j - \sum_{i=1}^{j} r_i$. Since by our assumption $\sum_{i=1}^{j} r_i < \delta j$, we know that $\xi > 0$. Suppose in the right-hand side of the above inequality, we increase the coefficient of $\hat{w}(s_j)$ by ξ, and decrease the coefficient of $\hat{w}(s_{j+1})$ by ξ. Since $\hat{w}(s_j) \geq \hat{w}(s_{j+1})$, this only makes the quantity larger. Thus, we obtain that:

$$\sum_{i=1}^{j+1} r_i \hat{w}(s_i) \leq \delta \sum_{i=1}^{j} \hat{w}(s_i) + [\sum_{i=1}^{j+1} r_i - \delta j]\hat{w}(s_{j+1}) \tag{4.3}$$

as desired. This proves Inequality (4.1) for every $j \leq m$, and thus completes the proof of this Lemma. □

We now use this value of k to construct the binary weight function $\bar{w}$. Take the

smallest k such that $\sum_{i=1}^{k} r_i \geq \delta k$ and let $\bar{w}$ be the weight function where for all $i \leq k$, $\bar{w}(s_i) = 1$, $\bar{w}(e) = 1$ for all e assigned to s_i, and all other edges have weight 0. Note that $r_k > 0$ since if $r_k = 0$ and $\sum_{i=1}^{k} r_i \geq \delta k$, then $\sum_{i=1}^{k-1} r_i \geq \delta(k-1)$.

We now argue that σ is still consistent with $\bar{w}$. Consider any adjacent edges $(x, y), (x, z)$ such that x prefers y to z, and suppose that (x, y) is assigned to edge s_i, while (x, z) is assigned to edge s_j. First consider the case when $w(s_i) \neq w(s_j)$. It must be that $i < j$, since the set of edges dominating (x, y) is a subset of edges dominating (x, z), and edges are assigned to a smallest edge of S dominating them. Then $\bar{w}(x, y) \geq \bar{w}(x, z)$, as desired. If instead $w(s_i) = w(s_j)$, then by construction of our assignment we have that $i = j$, so $\bar{w}(x, y) = \bar{w}(x, z)$. Therefore σ is consistent with $\bar{w}$ because for any adjacent edges $(x, y), (x, z)$ such that x prefers y to z it must be that $\bar{w}(x, y) \geq \bar{w}(x, z)$.

Thus we now have a binary weight function $\bar{w}$ such that S is a possible solution of Ordinal Greedy, since σ is consistent with $\bar{w}$. By definition of $\bar{w}$, we know that $\bar{w}(OPT_w) = \sum_{i=1}^{k} r_i$, and $\bar{w}(S) = k$. Due to our choice of k, we thus have that $\bar{w}(OPT_w) \geq \delta \cdot \bar{w}(S)$.

This concludes our proof that for any instance (w, σ) where $w : \mathcal{E} \to \mathbb{R}^+$ for which Ordinal Greedy provides a δ-approximation for $\delta > 0$ in the worst case, there exists an instance $(\bar{w}, \sigma)$ where $\bar{w} : \mathcal{E} \to \{0, 1\}$ for which Ordinal Greedy provides no better than a δ-approximation in the worst case. □

Another nice property is that every solution constructed by Ordinal Greedy is pairwise stable. Pairwise stability means that no pair of agents x, y has incentive to collude to add edge (x, y) by each giving up some of their edges in the Ordinal Greedy solution S. Either this exchange would decrease the total utility of one of the agents, or adding (x, y) is infeasible even after sacrificing the other edges. Here we assume that the utility of a node x in solution S is simply the total weight of edges in S incident on x.

Theorem 2. *Any solution S constructed by Ordinal Greedy on an independence system is pairwise stable.*

Proof. Let $(x, y) \notin S$ and let F_x and F_y be any set of edges in S adjacent to x and y excluding (x, y). If x and y can improve their individual utilities by adding (x, y) at the expense of removing all of the edges in $F_x \cup F_y$, this means $w(x, y) > w(F_x)$ and $w(x, y) > w(F_y)$. However, this clearly implies $w(x, y)$ is larger than the weight of each individual edge in F_x

and F_y. If $(x, y) \notin S$, then its critical iteration must have occurred before any of the adjacent edges in F_x and F_y were added to the Ordinal Greedy solution. Therefore $S - F_x - F_y + (x, y)$ and cannot be a feasible solution. □

4.5 Ordinal Approximation for ABC Systems

In this section we bound the worst-case performance of Ordinal Greedy compared to the optimal solution for any ABC System. We use $\alpha_{_{ABC}}$ to denote the approximation factor, or the ratio of the optimal solution to the worst possible Ordinal Greedy solution for any ABC System.

Unlike Example 1 in Section 4.4.1 for the maximum spanning tree problem, Ordinal Greedy does not provide a constant approximation factor for all ABC Systems. However, it does always provide a finite approximation which depends on the degree limit b. To simplify notation, since the optimal solution here is only evaluated using the same weight function used to generate it, we refer to the total weight of the optimal solution $w(OPT_w)$ as $w(OPT)$. Here we show that $\alpha_{_{ABC}} \leq b + 1$ for any ABC System and provide a family of examples where $\frac{w(OPT)}{w(S)} = b + 1$ to show that $b + 1$ is a tight bound on the approximation factor. In later sections, we explore classes of ABC Systems in which Ordinal Greedy achieves a better worst-case approximation.

Note that this result is quite general. As we discussed, ABC Systems include many varied constraints, some quite difficult to approximate. Our result in this section states that, even for extremely complex $\mathcal{A}$ and constraints on component size c, as long as the maximum allowed degree of any node is small, then it is possible to form a good approximation to the true optimum solution while only knowing ordinal information instead of the true edge weights.

Theorem 3. *For any ABC System, the Ordinal Greedy algorithm always produces a solution within a factor of (b+1) of the optimal solution, and this bound is tight.*

Proof Sketch 2. We provide a short proof sketch before the full proof. We proceed via a charging argument. We wish to charge the weight of the edges of OPT to the edges of S such that all edges of OPT are fully charged somewhere, and no edge of S receives a charge greater than $b + 1$ times its weight. However, unlike Omniscient Greedy in the full-information setting, we cannot assume that any eliminated edge of OPT has weight smaller

than all edges of the Ordinal Greedy solution S which were added before its critical iteration. This prohibits us from using the methods in previous work. Thankfully, due to Theorem 1 we know that if Ordinal Greedy produces a solution within a factor of $(b+1)$ of optimal for all instances with binary weight functions, then this holds for all instances. We therefore assume that all weights are $\{0,1\}$, and can now charge any edge of OPT to any weight 1 edge of S, but must ensure that no weight is charged to any edges of S with weight 0.

To ensure no edge of S is charged more than $b+1$ times its weight, we look at the connected components of S with only weight 1 edges, and charge all edges of OPT to these components. Specifically, we design the following charging scheme. Let (u,v) be an edge of OPT where $w(u,v)=1$. Let P_u and P_v be the connected components containing u and v in the subgraph of S containing only weight 1 edges. We charge the weight of $w(u,v)$ between P_u and P_v based on what occurs at the critical iteration of (u,v). If $(u,v) \in S$ then $P_u = P_v$, so charge its full weight to this component. If (u,v) was eliminated due to $\mathcal{A}$, charge its full weight to either P_u or P_v arbitrarily. Note that while the attachment property of $\mathcal{A}$ ensures that u and v are in the same connected component in S at this iteration, this does *not* imply that $P_u = P_v$ because all (u,v)-paths in S may contain a weight 0 edge. If (u,v) was eliminated due to $\mathcal{B}$, one of its endpoints must have a degree of exactly b in S at this iteration, so charge its full weight to the component containing this endpoint. If (u,v) was eliminated due to $\mathcal{C}$, we split the charge between P_u and P_v based on the size of the connected components in S containing u and v at this iteration. Let q_u and q_v be the sizes of the connected components in S containing u and v at the critical iteration of (u,v). Charge $\frac{q_u-1}{q_u+q_v-2}$ to P_u and $\frac{q_v-1}{q_u+q_v-2}$ to P_v. In all four cases we have ensured the full weight of the edge of OPT has been charged between the components containing its endpoints.

The rest of the proof involves arguing that each such component P_u of size p is charged a total of at most $(b+1)(p-1)$ using the above charging scheme. Since such a component must contain at least $p-1$ edges with weight 1, this completes the proof of the upper bound because it shows that the total number of edges in OPT with weight 1 is at most $b+1$ times the number of such edges in S.

We then provide a family of examples to show this bound is tight. Omniscient Greedy has the same worst-case solution as Ordinal Greedy on this family of examples, so our ordinal approximation competes well despite its knowledge handicap. Note that in the example yielding the lower bound of $\alpha_{_{ABC}} \geq b+1$, all edges are eliminated due to $\mathcal{C}$. The following

section demonstrates that when the component size constraint is relaxed by allowing $c = n$, the approximation factor improves significantly.

Proof. First, we prove the upper bound $\alpha_{ABC} \leq b + 1$. Since all ABC Systems are independence systems, by Theorem 1 it is sufficient to show that $\alpha \leq b + 1$ for all instances with integral weight functions $w : \mathcal{E} \to \{0, 1\}$ to provide an upper bound on α for all instances.

We proceed via a charging argument. Let w be a binary weight function $w : \mathcal{E} \to \{0, 1\}$, so all edges in our graph are either weight 1 or weight 0. Given the Ordinal Greedy solution S constructed for some instance (w, σ) we consider the subgraph $\bar{S} \subseteq S$ with only weight 1 edges and partition $\bar{S}$ into connected components. We then construct a charging scheme which charges the full weight of all the edges in OPT to these components. The total charge over all these components represents the total weight of OPT, while the total weight of these components is equal to the weight of the greedy solution. If our charging scheme guarantees that the full weight of all edges of OPT has been charged to the components, and no component of $\bar{w}$ is charged more than $(b+1)$ times the total weight of its edges, then we can sum over these components to show that $w(OPT) \leq (b + 1)w(\bar{S}) = (b + 1)w(S)$.

Let $\mathcal{P}$ be the set of connected components in $\bar{S}$. Let $P \in \mathcal{P}$ be any connected component in this subgraph of size $p \geq 1$ and total weight $w(P)$. We make two obvious, but critical observations about each component P.

Observation 2. *For any component P, its total weight $w(P)$ is at least $p - 1$, the weight of its maximum spanning tree.*

Observation 3. *There is a path between any two nodes in S with only weight 1 edges if and only if the nodes are in the same component P.*

Claim 2. There exists a charging scheme which charges the full weight of all edges of OPT to the components of $\mathcal{P}$ such that the total charge to any component P of size p is at most $(b + 1)(p - 1)$.

Proof. We need only charge the edges of OPT such that $w(u, v) = 1$ because the weight 0 edges do not contribute to the value of $w(OPT)$.

Charging Scheme: Let (u, v) be an edge of OPT where $w(u, v) = 1$. Let P_u and P_v be the connected components containing u and v in the subgraph $\bar{S} \subseteq S$ containing only weight 1 edges. We charge the weight of $w(u, v)$ between P_u and P_v based on what occurs

at the critical iteration of (u, v). If $(u, v) \in S$ then $P_u = P_v$, so charge its full weight to this component. If (u, v) was eliminated due to $\mathcal{A}$, charge its full weight to either P_u or P_v arbitrarily. Note that while the attachment property of $\mathcal{A}$ ensures that u and v are in the same connected component in S at this iteration, this does *not* imply that they are in the same connected component in $\mathcal{P}$ because all (u, v)-paths in S may contain a weight 0 edge. If (u, v) was eliminated due to $\mathcal{B}$, one of its endpoints must have a degree of exactly b in S at this iteration, so charge its full weight to the component containing this endpoint. If (u, v) was eliminated due to $\mathcal{C}$, we split the charge between P_u and P_v based on the size of the connected components in S containing u and v at this iteration. Let q_u and q_v be the sizes of the connected components in S containing u and v at the critical iteration of (u, v). Charge $\frac{q_u-1}{q_u+q_v-2}$ to P_u and $\frac{q_v-1}{q_u+q_v-2}$ to P_v. In all four cases we have ensured the full weight of the edge of OPT has been charged between the components containing its endpoints.

The weight of each edge $(u, v) \in OPT$ has been charged exclusively to the components containing its endpoints, P_u and P_v. Therefore, to determine the maximum possible charge to any component P, we bound the charge from edges of OPT with one or both endpoints in P and show this is at most $(b + 1)(p - 1)$.

Case 1) $p > b$

All nodes have at most b adjacent edges in OPT, so the maximum charge to any component P with p nodes is $b \cdot p$. If $p > b$ then the total charge on P is at most $b \cdot p \leq b \cdot p + (p - (b+1)) = (b + 1)(p - 1)$.

Case 2) $p \leq b$

Let u_0 be the first node in P to have an edge of weight 0 added adjacent to it in S at some iteration of Ordinal Greedy. By definition, at any time before the critical iteration of this weight 0 edge u_0 cannot be in the same component in S as any edge of weight 0. If there is no node in P with an adjacent weight 0 edge in S, let u_0 be any arbitrary node in P. Let a **Type 2** edge be an edge of OPT with weight 1 which is incident to u_0, but not to any other node in P. Let a **Type 1** edge be all edges of OPT of weight 1 which are not Type 1, including all edges with both endpoints in P and those with a single endpoint in P which is not u_0.

Since all nodes have degree at most b in OPT and there are $(p - 1)$ nodes other than u_0, it is clear that Type 1 edges cumulatively contribute a charge of at most $b \cdot (p - 1)$ to P. Here we show that Type 2 edges contribute a total charge of at most $p - 1$, limiting the

total charge to any component P to at most $b \cdot (p-1) + (p-1) = (b+1)(p-1)$.

Let (u_0, v) be an edge of OPT where $w(u_0, v) = 1$. The critical iteration of (u_0, v) must be before the critical iteration of the first weight 0 edge incident to u_0 in S. This is because at the iteration the weight 0 edge was added to S it had to be undominated, so (u_0, v) could not still have been adjacent to it in the set of available edges E.

Let P_v denote the component containing v. We look at the four cases of charging based on the critical iteration of (u_0, v) to show that Type 2 edges contribute a charge of at most $(p-1)$ to P.

If $(u_0, v) \in S$, then clearly this is Type 1 because $P = P_v$ and so it has already been charged to P. Likewise, if (u_0, v) was eliminated due to $\mathcal{A}$ then its endpoints must be in the same component $P = P_v$. This is because the attachment property ensures that u and v are in the same connected component in S at this iteration, and by the definition of u_0 it cannot yet have a path to any edge of weight 0. Therefore any path from u_0 to v in S must contain only weight 1 edges, meaning $P = P_v$. If (u_0, v) was eliminated due to $\mathcal{B}$ then its full weight is either charged to P or to P_v. For its weight to be charged to P, u_0 must have degree b at this iteration. For (u_0, v) to have weight 1 this means all b of the edges incident to u_0 in S must have weight 1 because otherwise they could not have been undominated before the critical iteration of (u_0, v). However, this would mean that all neighbors of u_0 are in P, so $p > b$. Therefore, if $p \leq b$ there can be no Type 2 edges charged to P which were eliminated due to $\mathcal{B}$.

We can now see that the only Type 2 edges charged to P are those which are eliminated due to $\mathcal{C}$. For each of these edges P is charged $\frac{q_{u_0}-1}{q_{u_0}+q_v-2}$. For (u_0, v) to be eliminated due to $\mathcal{C}$ this means the combined sizes of the disjoint components in S containing u_0 and v must be at least c, otherwise the edge (u_0, v) would still be a valid edge to add. In other words, $q_{u_0} + q_v > c$ so $q_{u_0} + q_v \geq c - 1$. And since no weight 0 edge may yet be adjacent to the component in S containing u_0 at this iteration, we know that $q_{u_0} \leq p$. Therefore, $\frac{q_{u_0}-1}{q_{u_0}+q_v-2} \leq \frac{p-1}{c-1}$, so the charge from each Type 2 edge eliminated due to $\mathcal{C}$ is at most $\frac{p-1}{c-1}$. And since there can be at most b Type 2 edges, the total charge they contribute to P is at most $b\,\frac{p-1}{c-1} = (p-1)\frac{b}{c-1} \leq p-1$ because $b \leq c-1$ (the maximum degree can never be more than the component size).

This leaves us with a total charge to P of at most $b(p-1) + (p-1) = (b+1)(p-1)$ when $p \leq b$. Together with Case 1, we have shown that the total charge to each component

$P \in \mathcal{P}$ of size p at most $(b+1)(p-1)$. This concludes Claim 2. □

By summing the charge over all components $P \in \mathcal{P}$ we get $w(OPT) \leq (b+1) \sum_{P \in \mathcal{P}} (p-1) \leq (b+1) \sum_{P \in \mathcal{P}} w(P) = (b+1)w(S)$ from the above claim. In other words, $\alpha_{ABC} = \max_{w,\sigma} \frac{w(OPT)}{w(S)} \leq b+1$.

To show that the above result is tight, consider the following ABC System. This system represents the problem of hedonic coalition formation with additive separable symmetric preferences ($b = c-1$ and $\mathcal{A}$ = all subgraphs of G), where agents are partitioned into coalitions (cliques) and each agent's total utility is the sum of its utility for being matched with all other agents in its coalition.

Proof. Suppose that $n = c^2$, $b = c-1$, and $\mathcal{A}$ = all subgraphs of G. In other words, the only constraint is that all components must be of size at most $c = \sqrt{n}$. Label the nodes u_{ij} for $i \in [1, c]$ and $j \in [1, c]$. Let $w(u_{i1}, u_{ij}) = 1$ for all $j > 1$. Let $w(u_{i1}, u_{k1}) = 1 + \epsilon$ for all $k = i + 1$ for some infinitesimal ϵ. Let all other edges have weight 0.

For appropriate choices of preferences σ, Ordinal Greedy may select each of the edges (u_{i1}, u_{k1}) for all $k = i + 1$ before selecting any others, creating a path of length $c - 1$. Thus, all of the weight 1 edges are eliminated due to $\mathcal{C}$. The optimal solution is to select each of the edges $(u_{i1}, u_{ij}) = 1$ for all $j > 1$. This yields $w(S) = (c-1)(1+\epsilon)$ while $w(OPT) = b \cdot c$ because OPT consists of c stars with b edges each. Therefore as $\epsilon \to 0$, $\frac{w(OPT)}{w(S)} = b \cdot \frac{c}{c-1} = b+1$ because $b = c - 1$. As α is an upper bound on the ratio between the optimal solution and the greedy solution for any instance, we have $\alpha \geq b + 1$. □

This concludes our proof of Theorem 3. □

4.6 AB Systems and Important Special Cases

In this section, we bound the performance of Ordinal Greedy on ABC Systems where $c = n$, effectively removing the component size constraint. We then discuss some common examples of maximization problems on AB Systems, including Max Spanning Tree, Max TSP, and Max Planar Subgraph. To improve our bound from $b + 1$ we invoke the notion of sparsity.

Definition 2. Sparsity

A graph S is d-sparse if for all subgraphs $F \subseteq S$ containing $V(F)$ nodes and $E(F)$ edges $\frac{E(F)}{V(F)} \leq d$ and for any $\hat{d} < d$ there exists a subgraph $\hat{F} \subseteq S$ such that $\frac{E(\hat{F})}{V(\hat{F})} > \hat{d}$.

Suppose our attachment set $\mathcal{A}$ and degree limit b imply that any feasible solution must be d-sparse. Note that this sparsity is implied by our constraints, and is not a separate constraint. Our main result in this section is that, for any graph collection which is guaranteed to be d-sparse, ordinal information is enough to produce good approximations. Specifically, we prove a bound of $d + 1$ for such settings. Since the sparsity corresponds to an upper bound on average degree of the nodes, it is always true that $d \leq \frac{b}{2}$, and so when $c = n$, this immediately reduces the approximation factor from $b + 1$ to $\frac{b}{2} + 1$. Even for large b, however, there are many natural classes of graphs that are always sparse, including planar graphs, scale-free graphs, graphs of small arboricity or treewidth, and many others. As we discuss in the next section, this result allows us to provide extremely strong guarantees for many important problems.

Theorem 4. *For any ABC System where the components can be of any size and the constraints imply that any feasible solution must be d-sparse, the Ordinal Greedy algorithm always produces a solution within a factor of $\max\{2, (d + 1)\}$ of the optimal solution, and this bound is tight.*

Proof Sketch 3. As with our proof of Theorem 3 for general ABC Systems, we only need to consider instances with weights $\{0, 1\}$ due to Theorem 1. However, the charging schemes and proofs for ABC Systems and AB Systems differ significantly. To lower the approximation factor from $b+1$ to $\max\{2, d+1\}$, we have to be more selective about where we charge the edges of OPT. For simplicity, we first assign the edges of OPT to their endpoints, before considering the total charge to all the nodes in any component. Since OPT is d-sparse, the edges of OPT can be assigned to their endpoints such that each node is assigned at most d edges. We then take such an assignment and for each edge of OPT eliminated due to $\mathcal{B}$ we change its assignment, if necessary, to the node which caused its elimination. Let P be a component of the subgraph of S containing only weight 1 edges, and suppose $p = |P|$. Then, similarly to the proof of Theorem 3, we must show that this component will be charged at most $\max\{2, d+1\}w(P)$, but unlike before, components may be charged more than $(d + 1)(p - 1)$ if $w(P) > p - 1$.

Now we consider two cases based on whether any node in a component was charged an edge of OPT eliminated due to $\mathcal{B}$. If there such a node in a component, then it must be possible to distribute the charge on the nodes over the edges of the component directly so that no edge is charged more than $\max\{2, d+1\}$ times its weight. If there is no such node, then we show that at least one node in the component must be charged at most $p-1$ and the rest are charged at most $d(p-1)$, cumulatively providing a charge at most $(d+1)(p-1)$ which can then be distributed over the edges in the component. Once again, we provide a family of examples to show that this bound is tight.

Proof. The theorem states that $\alpha_{AB} = \max\{2, d+1\}$. For AB Systems, the fact that $c = n$ means no edge is eliminated due to $\mathcal{C}$. This leaves us with only 3 cases which can occur at the critical iteration of an edge. We begin by creating an assignment of edges of OPT to the nodes to construct our charging scheme. We use the same notation as in the proof of Theorem 3.

As before, we know that the weight of a component P is at least that of its maximum spanning tree, $w(P) \geq p-1$. However, in certain cases, our strategy is different than the ABC Systems proof. We ensure that each component P is charged at most $\max\{2, d+1\}w(P)$, but a component may be charged more than $\max\{2, d+1\}(p-1)$.

As in the proof of Theorem 3, for each edge (u, v) of OPT, we charge its weight between the components containing its endpoints. We need only to charge the edges of OPT such that $w(u, v) = 1$, because the weight 0 edges do not contribute to the value of $w(OPT)$.

Claim 3. *There exists an assignment of edges of OPT to their endpoints such that each node is assigned at most d edges.*

Proof. Let N_d be a set of dn nodes containing d duplicates of each node in N. Let O be a set of nodes where each node corresponds to an edge of OPT. Construct a bipartite graph by building an edge from each node in O to each of the $2d$ nodes in N_d corresponding to the endpoints of the edge of OPT it represents.

Consider any subset of nodes $R \subseteq O$ corresponding to a subset of edges in OPT. Let $V(R)$ be the set of endpoints in N of all edges of OPT represented in R. Let R' be the set of all edges of OPT with both endpoints in $V(R)$, so that $R \subseteq R'$. Since R' can have sparsity at most d, it follows that $|R'| \leq d \cdot |V(R)|$. Therefore, if we consider the $d \cdot |V(R)|$ nodes in N_d corresponding to the duplicates of $V(R)$, which all have at least one edge to a node in R,

we have that $|R| \leq |R'| \leq d \cdot |V(R)|$. By Hall's Condition, we can create a perfect matching between O and N_d.

Take this perfect matching as described above, and assign each edge of OPT to the endpoint in N corresponding to the node in N_d to which they are matched. Since there d duplicates of each node in N_d, each node in N can be assigned at most d edges.

Note that while the proof above is written assuming that d is an integer to duplicate the nodes, when d is fractional the same result holds using a similar argument by duplicating the nodes in O. □

However, this assignment is not sufficient for our charging scheme. There may be too much assigned to small components P, and even to nodes which have no adjacent weight 1 edges in S at all would be charged d, when they should not be charged at all. Therefore, we take this assignment and alter it to create our charging scheme so that every component P is charged at most $(d+1)w(P)$.

Take the (possibly fractional) assignment of edges of OPT to their endpoints from Claim 3, and for every edge of OPT which was eliminated due to $\mathcal{B}$ change its assignment, if necessary, to be entirely to the endpoint which had degree b at its critical iteration.

Given this new assignment, our charging scheme is now simple. For each node in P charge the weight of the edges of OPT assigned to it to P. We now show that each component P has been charged at most $\max\{2, d+1\}w(P)$.

Let a *b-node* be defined as a node with exactly b adjacent edges of weight 1 in S, and therefore b adjacent edges in P. By construction, only *b-nodes* can receive a charge greater than d and at most b, while all other nodes are charged at most d. If a node has degree b in S but is not a *b-node* because one or more of its adjacent edges in S has weight 0, then any edges of OPT assigned to it which were eliminated due to $\mathcal{B}$ must have weight 0. The is because if any of the eliminated edges had weight 1, an edge of weight 0 could not be undominated at an iteration when the weight 1 edge is still adjacent to it in E. Therefore, a node which is not a *b-node* can have at most d edges of weight 1 assigned to it.

Case 1) P contains at least one *b-node*

Instead of showing that the total charge on P is bounded by $\max\{2, d+1\}w(P)$, in this case it is simpler to think of the charge from the edges of OPT as assigned to each particular node. The charge on P is the cumulative charge on all of the nodes in P, which we distribute over the edges of P. For any component P, if the total charge to P can be distributed over

its edges so that each edge is charged at most $\max\{2, d+1\}$, then the total charge to P is at most $\max\{2, d+1\}w(P)$. By summing over all $P \in \mathcal{P}$ we get $OPT(w) \leq \max\{2, d+1\}w(S)$.

All *b-nodes* have charge at most b, which we can distribute so that each of its b adjacent edges is charged at most 1. Now select one *b-node* and consider the maximum spanning tree of P. For all nodes which are not *b-nodes* distribute all d of their charge to their adjacent edge on the path to the selected *b-node* in this maximum spanning tree.

If any edge is between two *b-nodes*, it is charged at most 2. If any edge has exactly one *b-node* as an endpoint, it may be charged at most 1 from this endpoint and d from the other for a total of $d+1$. For any edge between two nodes which are not *b-nodes*, it is only charged from one of its endpoints, which has charge at most d. Since every edge of P has at most $\max\{2, d+1\}$ charge, the total charge over P is at most $\max\{2, d+1\}w(P)$.

Case 2) P does not contain any *b-nodes*

Let u_0 be the first node in P to have an edge of weight 0 added adjacent to it in S at some iteration of the Ordinal Greedy algorithm. If there is no node in P with an adjacent weight 0 edge in S, let u_0 be any node in P.

We show that the total charge to any component P is at most $\max\{2, d+1\}(p-1)$ by showing that there are at most $d(p-1)$ edges of OPT adjacent to the nodes of P, excluding those that have u_0 as their only endpoint in P. And there are at most $p-1$ edges of OPT which have u_0 as their only endpoint in P.

Since there are no *b-nodes*, all nodes in P must be charged at most d. Clearly, u_0 could only be charged by edges which were eliminated due to $\mathcal{A}$ or are included in P. This means that for all edges of OPT of weight 1 charged to u_0, their other endpoint must be in P. Therefore the total charge on u_0 is at most $min\{d, p-1\}$. Note that we cannot assume these edges have already been charged to P.

The sum of the total charge to the nodes of P is at most $d(p-1)+min\{d, p-1\}$ which we can distribute over at least $(p-1)$ edges of P. We have that $\frac{d(p-1)+min\{d,p-1\}}{p-1} \leq d+1$. Our charging scheme ensures that each component P is charged at most $(d+1)(p-1) \leq (d+1)w(P) \leq \max\{2, d+1\}w(P)$.

Together with Case 1, we sum over $P \in \mathcal{P}$ and get $w(OPT) \leq \max\{2, d+1\}w(S)$. □

Proof. Here we show that the above bound is tight. Let $N = \{u_1, ..., u_k, v_1, ..., v_k\}$. Let $w(u_i, u_j) = 1$ for all $i \neq j$, $w(u_i, v_i) = 1 + \epsilon$ for all $i \leq k$ for some infinitesimal ϵ, and let all other edges have weight 0. Let $c = n$, $b = c - 1$ and $\mathcal{A}$ = all planar subgraphs such that no

cycle may contain an edge (u_i, v_i) for any $i \leq k$. This implies $d \leq \frac{6k-6}{2k} = \frac{3k-3}{k}$ because any planar graph on n nodes has at most $3n - 6$ edges.

For appropriate preference orderings σ, Ordinal Greedy selects the edges (u_i, v_i) for all $i \leq k$ first. Now edges (v_i, v_j) with weight 0 may be undominated. If any edge (v_i, v_j) or (u_i, u_j) is selected or eliminated the other must be eliminated at that iteration, but no ordinal algorithm can decide optimally between these edges because they are not adjacent.

Suppose Ordinal Greedy selects all edges (v_i, v_j) where $j = i + 1$, causing all edges (u_i, u_j) with weight 1 to be eliminated due to $\mathcal{A}$, and proceeds by selecting additional (v_i, v_j) edges until the solution is maximally planar. This yield $w(S) = k(1 + \epsilon)$, whereas the optimal solution selects a maximal planar subgraph of edges (u_i, u_j). The edges (u_i, u_j) have a combined weight of $3k - 6$ because any clique on k nodes has a maximally planar triangulation of size $3k - 6$. The edges (u_i, v_i) have combined weight $k(1 + \epsilon)$, for a total of $w(OPT) = 4k - 6 + \epsilon k$. Therefore as $\epsilon \to 0$, $\alpha_{AB} \geq \frac{w(OPT)}{w(S)} = \frac{4k-6}{k} = \frac{3k-6}{k} + 1$, which asymptotically approaches $\frac{3k-3}{k} + 1 = d + 1$ as $k \to \infty$. □

The example for Max Weight Matching in Section 4.7 provides a lower bound example where $d = \frac{1}{2}$, so $d < 1$ and $\alpha \geq 2$.

4.6.1 Important Cases of AB Systems

Theorem 4 establishes that for AB Systems in which solutions are always sparse, ordinal algorithms don't perform much worse than ones which know the true underlying edge weights. While our result in the previous section is quite general, it is worth noting how it applies to many important problems which happen to be special cases of AB Systems. Since all tours and cycles are 1-sparse, and all planar graphs are at most 3-sparse, we immediately arrive at the following corollaries:

Corollary 1. *Ordinal Greedy always computes a 2-approximation for Maximum Weight Spanning Tree, and this bound is tight.*

See Example 1 stated earlier for tight lower bound examples for Maximum Spanning.

Corollary 2. *Ordinal Greedy always computes a 2-approximation for Maximum Traveling Salesperson, and this bound is tight.*

Proof. Let the set of agents be $N = \{u_1, ..., u_k, v_1, ..., v_{k-3}\}$. Let $w(u_i, u_{i+1}) = 1 + \epsilon$ for all $i \leq k-1$ for some infinitesimal ϵ. Let $w(v_i, u_{i+1}) = 1$ and $w(v_i, u_{i+2}) = 1$ for all $i \leq k-3$. Let $w(u_1, u_{k-1}) = 1$ and $w(u_2, u_k) = 1$ and $w(u_1, u_k) = 1$. Let all other edges have weight 0.

Suppose the Ordinal Greedy algorithm begins by selecting $(u_i, u_{i+1}) = 1 + \epsilon$ for all $i < k$. This creates a path of length $k-1$ which causes all weight 1 edges in the graph to be eliminated due to $\mathcal{A}$ or $\mathcal{B}$. Ordinal Greedy may then proceed by selecting some set of weight 0 edges, such as (u_1, v_1), (u_k, v_{k-3}), and (v_i, v_{i+1}) for $i < k-3$. Therefore, $w(S) = (k-1)(1+\epsilon)$. Meanwhile, the optimal solution is to select all of the weight 1 edges of the graph: (v_i, u_{i+1}) and (v_i, u_{i+2}) for all $i \leq k-3$ as well as (u_1, u_{k-1}), (u_2, u_k), and (u_1, u_k). This yields $w(OPT) = 2k-3$. Therefore, $\alpha_{TSP} \geq \frac{w(OPT)}{w(S)} = \frac{2k-3}{(k-1)(1+\epsilon)}$ which asymptotically approaches 2 as $k \to \infty$ and $\epsilon \to 0$. □

Corollary 3. *Ordinal Greedy always computes a 4-approximation for Max Weight Planar Subgraph.*

Generally, the same arguments can be applied to any problem where the goal is to find maximum-weight subgraphs with some excluded minor, finding maximum-weight graphs with small treewidth or arboricity, as well as a variety of other graph problems.

4.7 *b*-Matching

For any ABC System where $c = n$ and $\mathcal{A}$ = all subgraphs of G, the only constraint is that each node must have degree at most b. This is equivalent to the well-known problem of Max Weight b-Matching. In this case, the approximation provided by Ordinal Greedy improves greatly over general AB Systems. In fact, it provides a strict 2-approximation regardless of the value of b.

Theorem 5. *For any ABC System on graph G, where $c = n$ and $\mathcal{A}$ = all subgraphs of G, Ordinal Greedy always constructs a solution within a factor of 2 of the optimal solution. This bound is tight.*

Proof. Assume $\bar{w} : \mathcal{E} \to \{0, 1\}$. We proceed in the same way as the proof of Theorem 4. For each weight 1 edge $(u, v) \in OPT$ eliminated due to $\mathcal{B}$, charge its full weight to its endpoint which has degree b. All b adjacent edges to this node in the greedy solution must have weight 1 to have been undominated before the critical iteration of (u, v), so it is a *b-node*. Since no

node can have more than b adjacent edges in the optimal solution, this means that the total charge on any node is bounded by the number of edges of weight 1 adjacent to this node in the greedy solution. Therefore, all nodes can distribute their charge over their adjacent edges in S such that no edge in the greedy solution is charged more than 2.

Consider the following minimal example to show that the factor of 2 is tight. Suppose $b = 1$. Let $N = \{u_1, u_2, v_1, v_2\}$. Let $w(u_1, u_2) = w(v_1, v_2) = w(u_1, v_1) = 1$. Let $w(u_2, v_2) = 0$. For suitable σ, Ordinal Greedy selects (u_1, v_1) first because it is preferred by both u_1 and v_1, and the only other edge it can subsequently add is (u_2, v_2). This yields $w(S) = 1$ and $w(OPT) = 2$. □

Note that $b = 1$ is the problem Max Weight Matching, and our result generalizes the results from [27] and [61].

4.8 Conclusion and Further Directions

In this chapter we identify a large class of problems we call ABC Systems for which ordinal preference information is sufficient for algorithms to provide good approximations to optimal, even without access to cardinal utilities. Previous work has shown that if agent preferences form a metric space, approximations for TSP and matching can improve in expectation [27]. It remains to be seen how Ordinal Greedy performs on ABC Systems in expectation and how much the approximation factors for general ABC or AB Systems improve when this metric assumption holds. Also along the lines of previous work, it would be interesting to investigate whether truthful ordinal algorithms for ABC and AB Systems can compete with our non-truthful algorithm, much as [28] did for the problems first approached by [27]. Lastly, we have seen that all solutions produced by Ordinal Greedy are pairwise stable, but it is unknown for our problems whether all pairwise stable solutions produce a good approximation to optimum (although it is easily seen to be true for MST and TSP).

CHAPTER 5
IDEAL CANDIDATE DISTORTION WITH PREFERENCE STRENGTHS

5.1 Model and Notation

As in previous work on metric distortion, we have a set of agents $N = \{1, 2, \ldots, n\}$ and a set of candidates, or alternatives, X. These agents and candidates correspond to points in an arbitrary (unknown) metric space d. The agent preferences over the candidates are induced by the underlying metric, i.e., agents prefer candidates who are closer to them. Agent i prefers candidate p over candidate q (i.e., $p \succ_i q$) only if $d(i,p) \leq d(i,q)$. Moreover, we assume that the strengths of agent preferences are induced by these latent distances. If i prefers p over q, then the strength of this preference is $\alpha_i^{pq} = \frac{d(i,q)}{d(i,p)}$. The cost to agent i if candidate p is elected is $d(i,p)$, and the goal is to select the candidate minimizing the Social Cost: $SC(p) = \sum_{i \in N} d(i,p)$.

In previous work on metric distortion only the ordinal preferences were known, i.e., whether $(p \succ_i q)$ or $(q \succ_i p)$. In this chapter we assume that we are given some information about the preference strengths $\alpha_i^{pq} = \frac{d(i,q)}{d(i,p)}$ as well. Note that knowing these values still does not tell us how $d(i,p)$ compares with $d(j,p)$ for $i \neq j$, only how strongly each agent feels when comparing different candidates. While knowing even the *exact* preference strengths of all the agents is not enough to be able to select the optimum candidate (as shown in [2]), knowing just one bit of information about α_i^{pq} (such as whether $\alpha_i^{pq} \geq \tau$ for a threshold τ) is enough to create mechanisms with much better distortion than with only ordinal information.

Distortion For a given voting rule F and instance $I = \{N, X, d\}$, let y^* be the winning candidate selected by F and let x^* be the best available candidate (the one minimizing the

This chapter previously appeared as: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," in *Web and Internet Econ. - 15th Int. Conf. Proc.*, vol. 11920, I. Caragiannis, V. S. Mirrokni, and E. Nikolova, Eds., Dec. 2019, pp. 3–16, doi: 10.1007/978-3-030-35389-6_1.

This chapter previously appeared as: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," Jun. 2019, arXiv:1906.10562. (Retrieved on 7/18/2021).

Social Cost). Then, the *distortion of winning candidate* y^* is defined as

$$\delta_I = \frac{SC(y^*)}{SC(x^*)} \tag{5.1}$$

The *distortion of a voting rule* F is defined by its behavior on a worst-case instance:

$$\delta = \max_I \delta_I = \max_I \frac{SC(y^*)}{SC(x^*)} \tag{5.2}$$

5.2 Preliminaries

In our model we have a set of agents $N = \{1, 2, \ldots, n\}$ and a set of candidates X. These agents and candidates correspond to points in an arbitrary metric space d such that for any three points x, y, z triangle inequality holds: $d(x, y) \leq d(x, z) + d(y, z)$. We assume that agents' preferences over the candidates are induced by the underlying metric, and that agents are truthful (e.g. non-strategic). That is, agents prefer candidates who are closer to them. Agent i prefers candidate p over candidate q ($p \succ_i q$) only if $d(i, p) \leq d(i, q)$. Moreover, we assume that the strength of agents' preferences are induced by these latent distances. If i prefers p over q, then the strength of this preferences is $\alpha_i^{pq} = \frac{d(i,q)}{d(i,p)}$. When it is clear we are referring to 2 candidates p and q, we will drop the superscript.

Given a set of preference strength thresholds $\{1 \leq \tau_1 < \tau_2 < \ldots < \tau_m\}$, agents report the largest threshold which their preference strength exceeds for each pair of candidates. We let $A_l = \{i \in N : d(i, p) \leq d(i, q)$ and $\tau_l \leq \alpha_i < \tau_{l+1}\}$ and $B_l = \{j \in N : d(j, q) \leq d(j, p)$ and $\tau_l \leq \beta_j < \tau_{l+1}\}$. For convenience, we say $\tau_{m+1} = \infty$ and $\tau_0 = 1/\tau_1$. When $\tau_1 = 1$ we know the preferred candidate of every agent. When $\tau_1 > 1$ we let C denote the set of agents with preference strength strictly less than τ_1 whose preferred candidate is unknown.

The cost to agent i if candidate y^* is elected is $d(i, y^*)$ and the Social Cost is the sum of the costs to all of the individual agents, $SC(y^*) = \sum_{i \in N} d(i, y^*)$. We would like to select the candidate with the minimum social cost. However, preference strength information is insufficient for any mechanism to guarantee selection of the best available candidate. Therefore, the *distortion* (δ) measures the worst-case approximation ratio between the social cost of the candidate we select (y^*) and the best available candidate ($x^* \in X$) over all possible instances I, as defined in the Introduction.

5.2.1 Known Distortion Results

Table 5.1 summarizes the distortion bounds obtained in [2], [85]. These distortion bounds are achieved using the weighted majority rules listed below to compare all pairs of candidates. When there are only two candidates, selecting the winner according to the weighted majority rule gives the distortion bounds in the first column of Table 5.1. When there are more than two candidates, selecting a winner from the uncovered set of the tournament created using the weighted majority rule guarantees the distortion bounds in the second column.

Table 5.1: Distortion with preference strengths in different settings. Reprinted by permission from Springer Nature from: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," in *Web and Internet Econ. - 15th Int. Conf. Proc.*, vol. 11920, I. Caragiannis, V. S. Mirrokni, and E. Nikolova, Eds., Dec. 2019, pp.3–16, doi: 10.1007/978-3-030-35389-6_1. Copyright 2019.

Distortion	Two Candidates	More than Two Candidates
Preference Orders and a threshold τ	$\max\{\frac{\tau+2}{\tau}, \frac{3\tau-1}{\tau+1}\}$	$\min\{\max\{\frac{3\tau-1}{\tau+1}, \frac{\tau+2}{\tau}\}+2,$ $\max\{(\frac{3\tau-1}{\tau+1})^2, (\frac{\tau+2}{\tau})^2\}\}$
m thresholds $\tau_1, \ldots, \tau_m$	$\max_{1\le l\le m}\{\frac{\tau_l\tau_{l+1}+2\tau_{l+1}-1}{\tau_l\tau_{l+1}+1}\}$	$\max_{1\le l\le m}\{(\frac{\tau_l\tau_{l+1}+2\tau_{l+1}-1}{\tau_l\tau_{l+1}+1})^2\}$
Exact preference strengths	$\sqrt{2}$	2

Weighted Majority Rule 1. *Given voters' preferences and a threshold τ for two candidates, if $\tau \ge \sqrt{2}+1$, assign weight $\frac{\tau+1}{\tau-1}$ to all the voters with preference strengths $> \tau$ and weight 1 to all the voters with preference strengths $\le \tau$. If $\tau < \sqrt{2}+1$, assign weight τ to all the voters with preference strengths $> \tau$ and weight 1 to all the voters with preference strengths $\le \tau$. Choose the candidate by a weighted majority vote.*

Weighted Majority Rule 2. *Given candidates p and q and any single threshold $\tau \ge 1$, give all voters with preference strength at least τ a weight of 1 and all other voters a weight of 0. Select the candidate by weighted majority rule.*

Weighted Majority Rule 3. *For all $l < k$, assign to all voters in A_l and B_l a weight of $\frac{(\delta+1)(\tau_l\tau_{l+1}-1)}{(\tau_l+1)(\tau_{l+1}+1)}$. For all $l \ge k$, assign voters in A_l and B_l a weight of $\big((\frac{\tau_{l+1}-\delta}{\tau_{l+1}-1}) + (\frac{\delta\tau_l-1}{\tau_l+1})\big)$. Lastly, assign all voters in C a weight of 0. Choose the candidate by a weighted majority vote.*

Weighted Majority Rule 4. *Given the exact preference strength of every voter for two candidates, assign weight* $\frac{\sqrt{2}\alpha_i - 1}{\alpha_i + 1}$ *to each voter* $i \in A$ *such that* $\alpha_i > \sqrt{2}$*, and weight* $\alpha_i - 1$ *to each voter* $i \in A$ *such that* $\alpha_i \leq \sqrt{2}$*. Assign weight* $\frac{\sqrt{2}\beta_j - 1}{\beta_j + 1}$ *to each voter* $j \in B$ *such that* $\beta_j > \sqrt{2}$ *and weight* $\beta_j - 1$ *to each voter* $j \in B$ *such that* $\beta_j \leq \sqrt{2}$*.*

The following two established lemmas from [40] and [2] are useful for our proofs.

Lemma 3. *If a majority of voters prefer* P *to* Q*, then* $SC(P) \leq 2 \cdot SC(Z) + SC(Q)$ *for any other possible candidate* Z *[40].*

Lemma 4. *If Weighted Majority Rule 1 selects* P *over* Q*, then* $SC(P) \leq 2 \cdot SC(Z) + SC(Q)$ *where* Z *can be any point in the metric space [2].*

5.3 Ideal Candidate Distortion

In this section, we study the trade-off between the winning candidate distortion δ_I and the ideal candidate distortion Δ. For an instance $I = \{N, X, d\}$, suppose the winner is y^*, and the optimal candidate is x^*. Denote the distortion of y^* as $\delta_I = \frac{SC(y^*)}{SC(x^*)}$. Recall the ideal candidate is the best possible point in the metric that minimizes the total social cost (this point may or may not be in X); we denote this point as z^*. Then the ideal candidate distortion Δ of a candidate y^* is $\Delta = \frac{SC(y^*)}{SC(z^*)}$.

We show that for *any instance*, we have that *either* the distortion of our mechanism is small, *or* the ideal candidate distortion of our winning candidate is small. In other words, we establish that the only time when the selected alternative is not similar to the absolutely best possible alternative (which may be even better than any of the candidates in the running), is when it is similar to the best candidate from the ones up for consideration. So, for cases when distortion is large, at least we have a "consolation prize" that the chosen candidate is not too far from all possible alternatives, even the ones which the agents don't know about and do not express their preferences over.

To prove our results, we first need the following definition of a λ-bounded rule.

Definition 3. A majority rule is λ-bounded if p beating q directly in pairwise comparison implies $SC(p) \leq SC(q) + \lambda \cdot SC(z)$ for any point z in the metric space.

Theorem 6. *If* y^* *wins under a* λ*-bounded majority rule with two candidates, then* $\Delta = \frac{SC(y^*)}{SC(z^*)} \leq \frac{\lambda \delta_I}{\delta_I - 1}$*. If* y^* *is in the uncovered set under a* λ*-bounded majority rule with multiple candidates, then* $\Delta = \frac{SC(y^*)}{SC(z^*)} \leq \frac{2\lambda \delta_I}{\delta_I - 1}$*.*

Proof. First consider the two candidates setting with candidates p and q, and suppose p is the winner. We know that $SC(q) = \frac{1}{\delta_I} SC(p)$, and by the definition of λ-bounded majority rule,

$$\begin{aligned} SC(p) &\leq SC(q) + \lambda SC(z^*) \\ SC(p) &\leq \frac{1}{\delta_I} SC(p) + \lambda SC(z^*) \\ SC(p) &\leq \frac{\lambda \delta_I}{\delta_I - 1} SC(z^*) \end{aligned}$$

Then for the multiple candidates setting, if p is in the uncovered set, we know that either p beats q directly (and we get the same bound as in the two candidates setting), or there exists a candidate r, that p beats r and r beats q. Then by the definition of λ-bounded majority rule,

$$\begin{aligned} SC(p) &\leq SC(r) + \lambda SC(z^*) \\ SC(p) &\leq SC(q) + \lambda SC(z^*) + \lambda SC(z^*) \\ SC(p) &\leq \frac{1}{\delta_I} SC(p) + 2\lambda SC(z^*) \\ SC(p) &\leq \frac{2\lambda \delta_I}{\delta_I - 1} SC(z^*) \end{aligned}$$

□

Corollary 4. *With only agents' ordinal preferences, in the two candidates setting, the majority winner has an ideal candidate distortion of* $\Delta \leq \frac{2\delta_I}{\delta_I - 1}$. *And in the multiple candidate setting, any candidate in the uncovered set has an ideal candidate distortion of* $\Delta \leq \frac{4\delta_I}{\delta_I - 1}$.

Proof. By Lemma 3, the majority rule is 2-bounded. Then we get the conclusion directly from Theorem 6. □

Thus, in the usual "ordinal preference" setting of [23] and [40], either distortion of Copeland (or any candidate in the uncovered set) is actually bounded by $\delta_I \leq 3$ (instead of the worst-case of 5), or the ideal candidate distortion $\Delta \leq 6$, which may not seem like a great bound, but is impressive because it means that the selected candidate is a factor of 6

away from all possibilities, ones that are not known to anyone, ones that no one expresses their preferences over, and ones that may arise sometime in the future. The only assumption required is that all the possible alternatives and agents lie in some arbitrary, possibly very high-dimensional, metric space.

The same trade-off between δ_I and Δ occurs if we have are given agents' preferences and a single threshold on preference strength.

Corollary 5. *With agent preferences and a single threshold τ, we use Weighted Majority Rule 1 to decide pairwise winners. Then in the two candidate setting, the winner has an ideal candidate distortion of $\Delta \leq \frac{2\delta_I}{\delta_I - 1}$ (Figure 5.1). And in the multiple candidate setting, any candidate in the uncovered set has an ideal candidate distortion of $\Delta \leq \frac{4\delta_I}{\delta_I - 1}$ (Figure 5.2).*

Proof. By Theorem 4, Weighted Majority Rule 1 is 2-bounded. Then we get the conclusion directly from Theorem 6. □

5.3.1 Ideal Candidate Distortion With Exact Preference Strengths

In this section, we discuss the ideal candidate distortion when we know the agents' exact preference strength. We first show Weighted Majority Rule 4 is $(1 + \sqrt{2})$-bounded, then get the ideal candidate distortion by Theorem 6.

Suppose there are two candidates p and q, and we are given the preference strength of every agent. Denote A as the set of agents that prefer p to q, and B as the set of agents that prefer q to p. The preference strength of any $i \in A$ is denoted as α_i, and the preference strength of any $j \in B$ is denoted as β_j,

We first present a lemma which allows us to charge agents in B to agents in A; this lemma has not appeared previously and may be useful as a technique for proving other results as well.

Lemma 5. *Given any agent $i \in A$, and agent $j \in B$, we have that $d(j,p) \leq \frac{\beta_j(\alpha_i+1)}{\alpha_i\beta_j - 1} d(i,j)$.*

Proof. By the triangle inequality, $d(i,j) \geq d(j,p) - d(i,p)$, and $d(i,j) \geq d(i,q) - d(j,q) = \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)$. Thus, $d(i,j) \geq \max\{d(j,p) - d(i,p), \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)\}$.

Case 1. $d(j,p) - d(i,p) \geq \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)$:

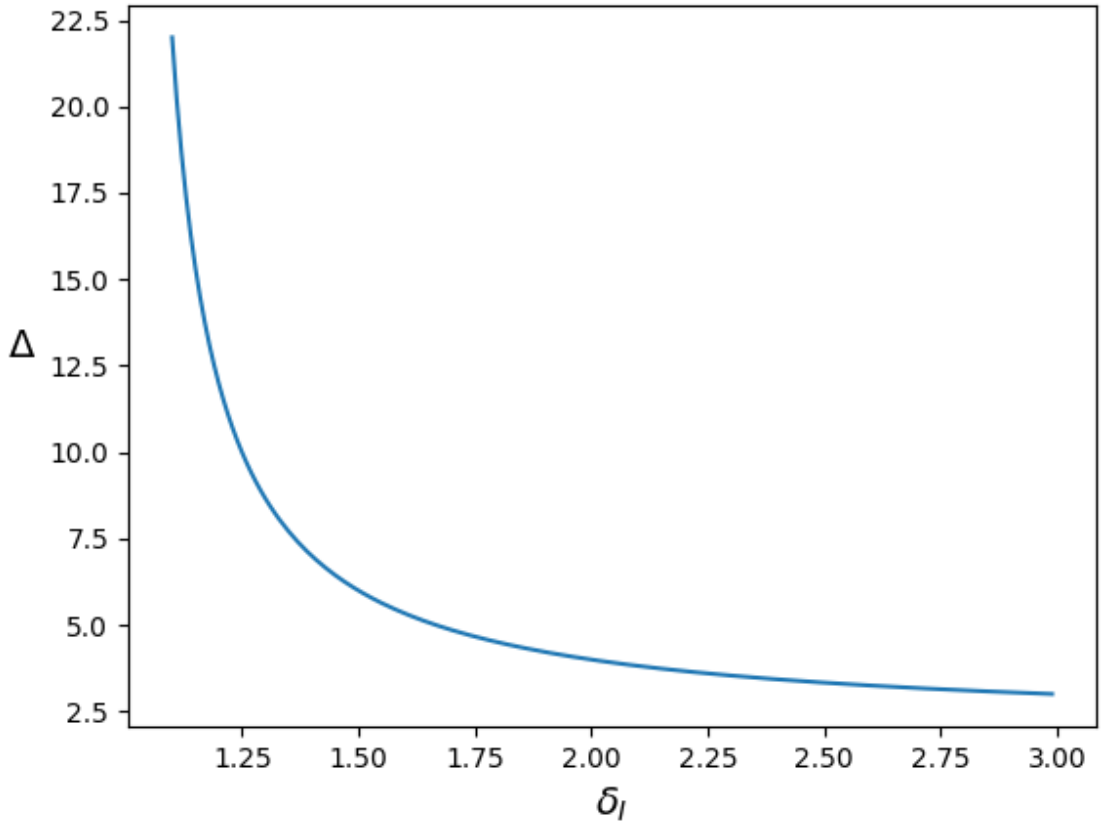

Figure 5.1: **Trade-off between δ_I and Δ with agent preferences and a single threshold τ in the two candidates setting. Reprinted by permission from Springer Nature from: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," in *Web and Internet Econ. - 15th Int. Conf. Proc.*, vol. 11920, I. Caragiannis, V. S. Mirrokni, and E. Nikolova, Eds., Dec. 2019, pp.3–16, doi: 10.1007/978-3-030-35389-6_1. Copyright 2019.**

$$d(j,p) - d(i,p) \geq \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)$$

$$(1 + \frac{1}{\beta_j}) d(j,p) \geq (\alpha + 1) d(i,p)$$

$$d(i,p) \leq \frac{\beta_j + 1}{\beta_j(\alpha_i + 1)} d(j,p)$$

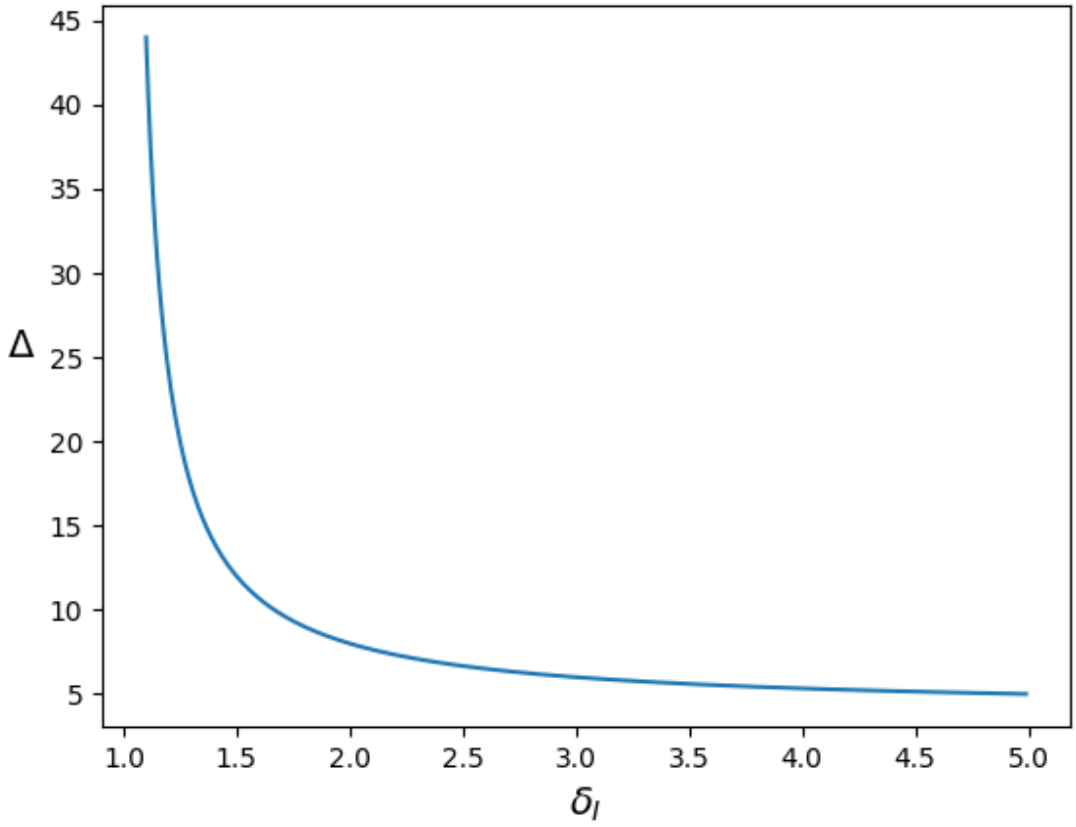

Figure 5.2: **Trade-off between δ_I and Δ with agent preferences and a single threshold τ in the multiple candidates setting. Reprinted by permission from Springer Nature from: B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," in *Web and Internet Econ. - 15th Int. Conf. Proc.*, vol. 11920, I. Caragiannis, V. S. Mirrokni, and E. Nikolova, Eds., Dec. 2019, pp.3–16, doi: 10.1007/978-3-030-35389-6_1. Copyright 2019.**

$$\begin{aligned} d(i,j) &\geq d(j,p) - d(i,p) \\ &\geq d(j,p) - \frac{\beta_j + 1}{\beta_j(\alpha_i + 1)} d(j,p) \\ &= \frac{\alpha_i \beta_j - 1}{\beta_j(\alpha_i + 1)} d(j,p) \end{aligned}$$

Case 2. $d(j,p) - d(i,p) < \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)$:

$$d(j,p) - d(i,p) < \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)$$
$$(1 + \frac{1}{\beta_j}) d(j,p) < (\alpha + 1) d(i,p)$$
$$d(i,p) > \frac{\beta_j + 1}{\beta_j(\alpha_i + 1)} d(j,p)$$

$$d(i,j) \geq \alpha_i d(i,p) - \frac{1}{\beta_j} d(j,p)$$
$$> \frac{\alpha_i \beta_j - 1}{\beta_j(\alpha_i + 1)} d(j,p)$$

□

Lemma 6. *If Weighted Majority Rule 4 selects p over q, for any $i \in A$, $j \in B$, let $r_{ij} = \frac{(\beta_j - 1)(\alpha_i + 1)}{\alpha_i \beta_j - 1}$. Then $\sum_{j \in B} \frac{1}{\sum_{i \in A} \frac{1}{r_{ij}}} \leq 1 + \sqrt{2}$.*

Proof.

$$\sum_{j \in B} \frac{1}{\sum_{i \in A} \frac{1}{r_{ij}}} = \sum_{j \in B} \frac{1}{\sum_{i \in A} \frac{\alpha_i \beta_j - 1}{(\beta_j - 1)(\alpha_i + 1)}}$$
$$= \sum_{j \in B} \frac{1}{\frac{\beta_j}{\beta_j - 1} \sum_{i \in A} \frac{\alpha_i - \frac{1}{\beta_j}}{\alpha_i + 1}}$$
$$\leq \sum_{j \in B} \frac{1}{\frac{\beta_j}{\beta_j - 1} \sum_{i \in A} \frac{\alpha_i - 1}{\alpha_i + 1}}$$
$$= \frac{\sum_{j \in B} \frac{\beta_j - 1}{\beta_j}}{\sum_{i \in A} \frac{\alpha_i - 1}{\alpha_i + 1}}$$

If p is selected as the winner, it means

$$\sum_{i \in A | \alpha_i > \sqrt{2}} \frac{\sqrt{2}\alpha_i - 1}{\alpha_i + 1} + \sum_{i \in A | \alpha_i \leq \sqrt{2}} (\alpha_i - 1) \geq \sum_{j \in B | \beta_j > \sqrt{2}} \frac{\sqrt{2}\beta_j - 1}{\beta_j + 1} + \sum_{j \in B | \beta_j \leq \sqrt{2}} (\beta_j - 1)$$

Claim 4. $\forall \alpha_i \geq \sqrt{2},\ \frac{\frac{\sqrt{2}\alpha_i - 1}{\alpha_i + 1}}{\frac{\alpha_i - 1}{\alpha_i + 1}} = \frac{\sqrt{2}\alpha_i - 1}{\alpha_i - 1} \leq \sqrt{2}$

Claim 5. $\forall \alpha_i \leq \sqrt{2},\ \frac{\alpha_i - 1}{\frac{\alpha_i - 1}{\alpha_i + 1}} = \alpha_i + 1 \leq 1 + \sqrt{2}$

Claim 6. $\forall \beta_j \geq \sqrt{2},\ \frac{\frac{\beta_j - 1}{\beta_j}}{\frac{\sqrt{2}\beta_j - 1}{\beta_j + 1}} \leq 2(\sqrt{2} - 1) < 1$

Proof.

$$
\begin{aligned}
((\sqrt{2} - 1)\beta_j - 1)^2 &\geq 0 \\
(3 - 2\sqrt{2})\beta_j^2 - 2(\sqrt{2} - 1)\beta_j + 1 &\geq 0 \\
(2 - 2\sqrt{2})\beta_j^2 - 2(\sqrt{2} - 1)\beta_j &\geq \beta_j^2 - 1 \\
2(\sqrt{2} - 1)(\sqrt{2}\beta_j^2 - \beta_j) &\geq \beta_j^2 - 1 \\
2(\sqrt{2} - 1)\beta_j(\sqrt{2}\beta_j - 1) &\geq (\beta_j + 1)(\beta_j - 1) \\
2(\sqrt{2} - 1)\frac{\sqrt{2}\beta_j - 1}{\beta_j + 1} &\geq \frac{\beta_j - 1}{\beta_j}
\end{aligned}
$$

□

Claim 7. $\forall \beta_j \leq \sqrt{2},\ \frac{\frac{\beta_j - 1}{\beta_j}}{\beta_j - 1} = \frac{1}{\beta_j} \leq 1$

By the four claims above,

$$
\begin{aligned}
\sum_{j \in B} \frac{\beta_j - 1}{\beta_j} &\leq \sum_{j \in B | \beta_j > \sqrt{2}} \frac{\sqrt{2}\beta_j - 1}{\beta_j + 1} + \sum_{j \in B | \beta_j \leq \sqrt{2}} (\beta_j - 1) \\
&\leq \sum_{i \in A | \alpha_i > \sqrt{2}} \frac{\sqrt{2}\alpha_i - 1}{\alpha_i + 1} + \sum_{i \in A | \alpha_i \leq \sqrt{2}} (\alpha_i - 1) \\
&\leq (1 + \sqrt{2}) \sum_{i \in A} \frac{\alpha_i - 1}{\alpha_i + 1}
\end{aligned}
$$

Thus,

$$\sum_{j\in B}\frac{1}{\sum_{i\in A}\frac{1}{r_{ij}}} \le \frac{\sum_{j\in B}\frac{\beta_j-1}{\beta_j}}{\sum_{i\in A}\frac{\alpha_i-1}{\alpha_i+1}}$$
$$\le (1+\sqrt{2})$$

□

Theorem 7. *If Weighted Majority Rule 4 selects p over q, then $SC(p) \le (1+\sqrt{2})SC(z) + SC(q)$ where z can be any point in the metric space.*

Proof. Let A be the set of agents prefer p to q, and B be the set of agents prefer q to p. If p is selected as the winner, it means

$$\sum_{i\in A|\alpha_i>\sqrt{2}}\frac{\sqrt{2}\alpha_i-1}{\alpha_i+1}+\sum_{i\in A|\alpha_i\le\sqrt{2}}(\alpha_i-1) \ge \sum_{j\in B|\beta_j>\sqrt{2}}\frac{\sqrt{2}\beta_j-1}{\beta_j+1}+\sum_{j\in B|\beta_j\le\sqrt{2}}(\beta_j-1)$$

Select an arbitrary agent $i \in A$, and agent $j \in B$.

By Lemma 5,

$$\begin{aligned}
d(j,p) &\le \frac{\beta_j(\alpha_i+1)}{\alpha_i\beta_j-1}d(i,j)\\
\frac{\beta_j-1}{\beta_j}d(j,p) &\le \frac{\beta_j-1}{\beta_j}\times\frac{\beta_j(\alpha_i+1)}{\alpha_i\beta_j-1}d(i,j)\\
\frac{\beta_j-1}{\beta_j}d(j,p) &\le \frac{(\beta_j-1)(\alpha_i+1)}{\alpha_i\beta_j-1}d(i,j)\\
\frac{\beta_j-1}{\beta_j}d(j,p) &\le \frac{(\beta_j-1)(\alpha_i+1)}{\alpha_i\beta_j-1}(d(i,z)+d(j,z))
\end{aligned}$$

Let $r_{ij} = \frac{(\beta_j-1)(\alpha_i+1)}{\alpha_i\beta_j-1}$,

$$\frac{1}{r_{ij}}\frac{\beta_j-1}{\beta_j}d(j,p) \le (d(i,z)+d(j,z))$$

Claim 8. *For any $\alpha_i \ge 1$, $\beta_j \ge 1$, $r_{ij} \le 2$.*

Proof.

$$\begin{aligned}
\alpha_i &\geq 1 \\
\alpha_i(\beta_j + 1) &\geq \beta_j + 1 \\
\alpha_i\beta_j + \alpha_i - \beta_j - 1 &\geq 0 \\
2\alpha_i\beta_j - \alpha_i\beta_j + \alpha_i - \beta_j - 2 + 1 &\geq 0 \\
2\alpha_i\beta_j - 2 &\geq \alpha_i\beta_j - \alpha_i + \beta_j - 1 \\
2(\alpha_i\beta_j - 1) &\geq (\alpha_i + 1)(\beta_j - 1) \\
\frac{(\beta_j - 1)(\alpha_i + 1)}{\alpha_i\beta_j - 1} &\leq 2
\end{aligned}$$

□

Sum up for all $i \in A$,

$$\begin{aligned}
\sum_{i\in A} \frac{1}{r_{ij}} \frac{\beta_j - 1}{\beta_j} d(j,p) &\leq \sum_{i\in A} (d(i,z) + d(j,z)) \\
\frac{\beta_j - 1}{\beta_j} d(j,p) &\leq \frac{1}{\sum_{i\in A} \frac{1}{r_{ij}}} \sum_{i\in A} d(i,z) + \frac{1}{\sum_{i\in A} \frac{1}{r_{ij}}} |A| d(j,z)
\end{aligned}$$

Then sum up for all $j \in B$,

$$\begin{aligned}
\sum_{j\in B} \frac{\beta_j - 1}{\beta_j} d(j,p) &\leq \sum_{j\in B} \frac{1}{\sum_{i\in A} \frac{1}{r_{ij}}} \sum_{i\in A} d(i,z) + |A| \sum_{j\in B} \frac{1}{\sum_{i\in A} \frac{1}{r_{ij}}} d(j,z) \\
&\leq \sum_{j\in B} \frac{1}{\sum_{i\in A} \frac{1}{r_{ij}}} \sum_{i\in A} d(i,z) + |A| \sum_{j\in B} \frac{1}{\sum_{i\in A} \frac{1}{2}} d(j,z) \\
&\leq \sum_{j\in B} \frac{1}{\sum_{i\in A} \frac{1}{r_{ij}}} \sum_{i\in A} d(i,z) + 2 \sum_{j\in B} d(j,z) \\
&\leq (1 + \sqrt{2}) \sum_{i\in A} d(i,z) + 2 \sum_{j\in B} d(j,z)
\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{i\in A} d(i,p) + \sum_{j\in B} d(j,p) &= \sum_{i\in A} d(i,p) + \sum_{j\in B} (\frac{1}{\beta_j} + \frac{\beta_j - 1}{\beta_j}) d(j,p) \\ &\leq \sum_{i\in A} d(i,q) + \sum_{j\in B} d(i,q) + \sum_{j\in B} \frac{\beta_j - 1}{\beta_j} d(j,p) \\ &\leq \sum_{i\in A} d(i,q) + \sum_{j\in B} d(i,q) + (1+\sqrt{2}) \sum_{i\in A} d(i,z) + 2\sum_{j\in B} d(j,z)\end{aligned}$$

□

Corollary 6. *With every agent's exact preference strength, we use Weighted Majority Rule 4 to decide pairwise winners. In the two candidates setting, the winner has an ideal candidate distortion of* $\Delta \leq \frac{(\sqrt{2}+1)\delta_I}{\delta_I - 1}$. *And in the multiple candidate setting, any candidate in the uncovered set has an ideal candidate distortion of* $\Delta \leq \frac{2(\sqrt{2}+1)\delta_I}{\delta_I - 1}$.

Proof. By Theorem 7, Weighted Majority Rule 4 is $(\sqrt{2}+1)$-bounded. Then we get the conclusion directly from Theorem 6. □

5.3.2 Ideal Candidate Distortion Without Agent Preference Orders

With only one threshold τ, Weighted Majority Rule 2 is not λ-bounded for any constant λ [2]. However, we can still get some trade-off between the distortion δ_I of the winning candidate and the ideal candidate distortion Δ for δ_I in a certain range. We will first show the relationship among $SC(p)$, $SC(q)$, and $SC(z)$ for any point z in the metric space by the following lemma.

Lemma 7. *Consider the setting with two candidates* p, q, *and a single threshold* τ. *If* p *pairwise beats* q *by Weighted Majority Rule 2, then* $SC(p) \leq \tau SC(q) + 2SC(z)$ *for any point* z *in the metric space.*

Proof. Let A denote the set of agents who prefer p to q and have preference strength $\geq \tau$, let B denote the set of agents who prefer q to p and have preference strength $\geq \tau$, and let C denote the rest of the agents. Because p pairwise beats q by Weighted Majority Rule 2, we know that $|A| \geq |B|$.

For agents in A and B, by Lemma 3,

$$\sum_{i \in A+B} d(i,p) \leq \sum_{i \in A+B} d(i,q) + 2 \sum_{i \in A+B} d(i,z)$$

For agents in C, because we know that they do not strongly prefer q to p, it must be that $\sum_{i \in C} d(i,p) \leq \tau \sum_{i \in C} d(i,q)$. Summing up all the agents in A, B, C, we get $SC(p) \leq \tau SC(q) + 2SC(z)$. □

Theorem 8. *Consider the setting with two candidates p, q, and a single threshold τ. Suppose z^* is the ideal possible candidate. If p pairwise beats q by Weighted Majority Rule 2, when the distortion $\delta_I \geq \tau$, then $\Delta = \frac{SC(p)}{SC(z^*)} \leq \frac{2\delta_I}{\delta_I - \tau}$.*

Proof. By Lemma 7, we know that:

$$\begin{aligned} SC(p) &\leq \tau SC(q) + 2SC(z^*) \\ SC(p) &\leq \tau \times \frac{1}{\delta_I} SC(p) + 2SC(z^*) \\ (1 - \frac{\tau}{\delta_I}) SC(p) &\leq 2SC(z^*) \\ SC(p) &\leq \frac{2\delta_I}{\delta_I - \tau} SC(z^*) \end{aligned}$$

□

Rewriting the bound of Δ in terms of $\frac{\delta_I}{\tau}$, $\Delta \leq \frac{2\frac{\delta_I}{\tau}}{\frac{\delta_I}{\tau} - 1}$ when $\frac{\delta_I}{\tau} \geq 1$. Thus the tradeoff is the same as in the case with only ordinal preferences being known (Figure 5.1), except replacing δ_I with $\frac{\delta_I}{\tau}$. This makes sense since the case with only ordinal preferences is exactly the special case with a single threshold $\tau = 1$.

5.3.3 Ideal Candidate Distortion With General Thresholds

In Section 5.3.2, we discussed that with only one threshold τ, there is a tradeoff between δ_I and Δ when the distortion $\delta_I \geq \tau$. Similarly, in the general setting when we are given m thresholds $\{1 \leq \tau_1 < \tau_2 < \ldots < \tau_m\}$, there is also a tradeoff between δ_I and Δ when the distortion $\delta_I \geq \tau_m$.

Lemma 8. *Consider the setting with two candidates p, q, and thresholds $\{1 \leq \tau_1 < \tau_2 < \ldots < \tau_m\}$. If p pairwise beats q by Weighted Majority Rule 3, then $SC(p) \leq \tau_m SC(q) + 2SC(z)$ for any point z in the metric space.*

Proof. Let A denote the set of all the agents that prefer p to q, and have preference strength $\geq \tau_1$, i.e., $A = A_1 \cup A_2 \cup \cdots \cup A_m$. Recall A_l denotes the set of agents have preference strength α_i such that $\tau_l \leq \alpha_i < \tau_{l+1}$. Similarly, define $B = B_1 \cup B_2 \cup \cdots \cup B_m$.

First we prove the size of A is at least the size of B_m. Weighted Majority Rule 3 assigns heavier weights to agents with stronger preference strengths. Thus, the agents in B_m and A_m are assigned the heaviest weight. Remember agents in C are assigned weight 0, so the winner is decided by agents in A and B. If p wins over q, it must be the case that $|B_m| \leq |A|$, because otherwise the total weight of agents in B_m must be higher than the total weight of agents in A, and q would be the winner instead.

For agents in A and B_m, by Lemma 3 we have that

$$\sum_{i \in A+B_m} d(i,p) \leq \sum_{i \in A+B_m} d(i,q) + 2 \sum_{i \in A+B_m} d(i,z)$$

For any other agent i in C or B_l ($l < m$), we know that $d(i,p) \leq \tau_m d(i,q)$. Summing up for all agents, we get $SC(p) \leq \tau_m SC(q) + 2SC(z)$. □

Theorem 9. *Consider the setting with two candidates p, q, and thresholds $\{1 \leq \tau_1 < \tau_2 < \ldots < \tau_m\}$. Suppose z^* is the ideal possible candidate. If p pairwise beats q by Weighted Majority Rule 3, when the distortion $\delta_I \geq \tau_m$, then $\Delta = \frac{SC(p)}{SC(z^*)} \leq \frac{2\delta_I}{\delta_I - \tau_m}$.*

Proof. The proof is exactly the same as for Theorem 8, by using Lemma 8 instead of Lemma 7. □

5.4 Conclusion

Unfortunately, one of the drawbacks to distortion as a measure efficiency is that it is not robust in practice. For example, with multiple candidates, the addition or subtraction of a single candidate or agent can cause the actual approximation achieved by Copeland in an instance to swing between 1 (optimal) and 5 (worst-case distortion). Our notion of *ideal candidate distortion* partly addresses this issue by showing that the distortion of Copeland (and other λ-bounded rules) can only be high when the winning candidate is

within a constant factor of the ideal conceivable candidate, even if they are not a candidate and nothing about them is known. However, when distortion is low the ideal candidate distortion is unbounded in general. Therefore we observe a general tradeoff between the quality of the available candidates and how poorly we might possibly choose from among the candidates.

Acknowledgements

This work was partially supported by NSF award CCF-1527497.

PART II

PREFERENCES OVER BOTH AGENTS AND ALTERNATIVES

CHAPTER 6
CONTRIBUTIONS, BACKGROUND, AND RELATED WORK

Chapter 4 and Chapter 5 concerned different preference objects. In the former, agents had reciprocal preferences over other agents, while in the latter, agents had preferences over non-agent alternatives. In Chapter 7 and Chapter 8, agents will have preferences over both agents and alternatives, though in different ways.

6.1 Forking

In Chapter 7, our agents are faced with two options, A and B. Unlike Chapter 5, in which all agents had to collectively select a single winning alternative, now each agent will select an alternative to which they will be assigned. For example, if A and B are locations, each agent must go to one and only one location. However, agents also care about how many other agents are assigned to the same location. Agents want there to be as many other agents with them as possible. Thus, the full set of candidates over which agents express their preferences is the set of tuples (option, number of agents). An agent might prefer to be at location A, but only if there are at least 10 other agents there with them, otherwise they prefer to be at location B.

A prime consideration of this chapter is constructing and characterizing stable assignments. In particular, it is desirable to know when only a single stable assignment exists and

Portions of this chapter previously appeared as: B. Abramowitz and N. Mattei, "Flexible representative democracy: An introduction with binary issues," in *Proc. 28th Int. Joint Conf. Artif. Intell., Macao, China*, S. Kraus, Ed., Aug. 2019, pp. 3–10, doi: 10.24963/ijcai.2019/1.

Portions of this chapter previously appeared as: B. Abramowitz and N. Mattei, "Flexible representative democracy: An introduction with binary issues," Apr. 2019, arXiv:1811.02921 (Retrieved on 7/18/2021).

Portions of this chapter previously appeared as: B. Abramowitz, E. Elkind, D. Grossi, E. Shapiro, and N. Talmon, "Democratic forking: Choosing sides with social choice," Mar. 2021, arXiv:abs/2103.03652 (Retrieved on 7/18/2021).

Portions of this chapter are to appear as: B. Abramowitz, E. Elkind, D. Grossi, E. Shapiro, and N. Talmon, "Democratic forking: Choosing sides with social choice," in *7th Int. Conf. Algorithmic Decision Theory.,* Nov. 2021.

when giving all agents the same assignment is a stable solution. In addition, it is desirable that the assignment mechanism be strategy-proof, meaning that no subset of agents can all benefit by misreporting their preferences in a coordinated manner.

Chapter 7 is positioned at the interface of the research on blockchain technology and Computational Social Choice. A blockchain is a replicated data structure designed to guarantee the integrity of data (e.g., monetary transactions in cryptocurrency applications such as Bitcoin [86]) and computations on them, combined with consensus protocols, which allow peers to agree on their content (e.g., who has been paid) and to ensure that no double spending of currency has occurred (see [87] for recent overview). By now, blockchain is an established technology, and cryptocurrency applications are attracting considerable attention [88].

When the community of a specific blockchain protocol—such as the Bitcoin protocol—is not satisfied with it, it may break into several subcommunities. The community typically consists of developers, who build the software, miners, who operate the protocol, and users, e.g., account holders. When some of the developers of the current protocol decide to modify it, they create an alternative branch that obeys their new protocol, and if it attracts a sufficient number of miners and users, the result is a so-called *hard fork*. Several such hard forks have been documented, including among the most influential cryptocurrencies. Bitcoin (cf. [89]), despite its relatively short history, has already undergone seven hard forks. At the moment, a key feature of these hard forks is that they happen through an informal social process, and, crucially, in ways that are completely exogenous to the protocol underpinning the blockchain. In blockchain terminology, they are said to happen 'off-chain'. This points to a lack of governance in most current blockchain systems, for better or worse.

Against this backdrop there have been attempts at incorporating protocol amendment procedures within blockchain protocols themselves so as to avoid 'off-chain' forks (so-called 'on-chain' governance, cf. [90]); more generally, the issue of governance is attracting increasing attention [91], [92]. We are not aware, however, of research that approaches forking as a social choice problem, and aims for an algorithmic solution. We lay the foundations for this approach here.

Social choice theory studies preference aggregation methods for various settings [37]. Our social choice setting is closely related to assignment problems, as the result of a fork is that each agent is "assigned" to a community. In this context, we mention works on judge-

ment aggregation [37, Ch. 17, pp. 319–426], [93] (which, formally, can capture assignments as well) and on partition aggregation [94]. To the best of our knowledge, the specific social choice setting we consider is novel. In a broader context, we mention work on stability in coalition formation games [95] [37, Ch. 15, pp. 356-376] as well as the recent paper on deliberative majorities [96], which studies coalition formation in a general voting setting. Our model is also related to the group activity selection problem with (increasing) ordinal preferences (o-GASP) [97] (see also [98], [99]), where our notion of stability corresponds to core stability in o-GASP. However, due to our focus on strategy-proofness, and the fact that our setting does not admit for a no-choice option (void activity in o-GASP), most of the existing results for o-GASP are not directly relevant to our study, so we chose not to use the o-GASP formalism in our presentation.

Forking is not limited to the cryptocurrency setting only; in particular, forking is relevant to projects of open-source code, in which a community jointly writes a piece of code and may experience different opinions regarding the code that is being written. Indeed, there is some work on using social choice mechanisms (and, in particular liquid democracy) for revision control systems [100]. Others have been studying the phenomena of forking in open source projects; see, e.g., the work of Zhou et al. [101].

Contributions Chapter 7 introduces a simple, efficient mechanism for assigning agents to two options, or locations, such that the assignment is stable, meaning that no subset of agents can move in a coordinated fashion such that everyone who moves benefits. It is shown that this mechanism is strategy-proof if and only if there is only a single stable assignment. Various preference structures are explored that can be elicited efficiently, as compared to ranking all possible combinations of (option, number of agents). Lastly, computing stable assignments with three or more options is NP-Hard.

6.2 Flexible Representative Voting

In Chapters 4-7, there is a single decision to be made with a single outcome, and all agents are able to express their preferences directly over some set of alternatives that can be used to directly compute the outcome. But in many real-world situations, groups of agents have to make many complex decisions. This is why large groups of agents often employ methods of representation, to reduce the burden of attention and effort on the society of

agents at large. The problem with using representatives to make decisions on behalf of a larger body of agents is that the representatives may make different decisions than the agents as a whole would if the larger body of agents were reporting their preferences directly.

One special form of representative decision making is that of proxy voting, in which every agent can select a single proxy to represent them individually [76]. In Chapter 8, we propose a particular mechanism of proxy voting that can be used to interpolate between direct voting and representative voting. The performance of our voting system is compared to that of simpler mechanisms in terms of the fraction of decisions on which it returns the same outcome that would occur under direct voting with full participation. When voters do not use the proxy voting function at all, the voting system is identical to representative voting. When the voters all use the proxy voting function on every decision, thereby shifting power to the representatives who agree with them on each decision, the system can emulate representative voting. When only a fraction of voters use the proxy feature, the system lands somewhere in between.

Miller [102], inspired by Tullock [103] and shareholder proxy voting, suggested an interactive democratic system for legislation that could operate at scale using computers. Miller lamented the lack of "flexibility" in Representative Democracy and sought to remedy this using a dynamic system of proxies, although he admitted this was not conducive to creating legislation. Soon after, [104] warned that electronic systems may accelerate the legislative process in undesirable ways and suggested holding every referendum twice to guarantee time for sufficient public deliberation. In Chapter 8, our use of a fixed, elected set of representatives answers Miller's question of how to produce legislation, and rather than holding redundant referenda we make representatives' votes public and then give the agents sufficient time to deliberate and alter their choice of proxies (or delegates).

Before the dawn of the internet, Tullock [105] revisited his ideas from 1967 [103] in a proposal that motivates the default and delegation mechanisms in flexible representative voting (or flexible representative democracy). The notion of the default distribution is similar to that proposed by [106], which suggests that the weights of representatives be based on the preferences of agents expressed in the election, but these weights are fixed during their term. By contrast, in flexible representative voting the weight of each representative on each issue is not strictly determined by the election. [107] took an analytical approach to studying a Proxy Voting model close to that of [106] for decision making with no election, infinite

agents, spatial preferences, and that agents lie in a metric space.

The hallmark of interactive democracies like flexible representative voting and Liquid Democracy is that rather than adjudicating whether a direct or representative system is better for achieving some objective and asserting it by fiat, the extent to which the system is direct or representative is itself a function of the "will of the voters". Currently, Liquid Democracy is the most well-known and well-studied form of Interactive Democracy, and has been studied from an algorithmic perspective as a decision-making process in the AI and COMSOC literature and elsewhere [108]–[129] (see [130] for summary). Unlike Liquid Democracy, flexible representative voting does not allow transitive delegations nor delegations to another agent, thereby violating the second axiom proposed by [114].

Fractional delegations in flexible representative voting serve a similar function to that of the virtual committees proposed by [114], although in theory flexible representative voting could incorporate virtual committees as well as other mechanisms for delegating voting power.

The design of flexible representative voting is largely based on work in probabilistic voting, binary aggregation, statistical decision theory, and computational social choice, as in e.g, [131]–[140]. In particular, the largest influences were work on the optimal weighting of experts [132], [139], [141]–[144], the Condorcet Jury Theorem [145]–[148], variable electorates [149]–[151], optimal committee sizes [152]–[154], and power in weighted voting games [122], [155]. In flexible representative voting, one can view the agent delegations as a pseudo-tie breaking mechanism for the representatives or, conversely, see the default distribution as a way to dampen the variance in the outcome in that may occur a direct democracy when the set of participating agents is small or biased. Another view is that electing representatives is analogous to a compression algorithm [156], which is the algorithmic version of John Adams's alleged intuition that the representatives should be a microcosm of the population (taken from [106]). In this view, the delegations in flexible representative voting are analogous to a decompression mechanism where a higher delegation rate reduces the "loss" of representation. Our evaluations are similar to those of [157], however, in their model the quality of the committee is measured as the sum of the agent proportion being represented for each issue, while we focus only on the total number of issues correct according to direct voting.

The work most similar following ours is that of [158], which considers the efficiency of

sortition (random selection), with and without additional delegation. Another adaptation of our mechanism demonstrated that the larger the number of agents, the higher the probability that a majority of representatives weighted with delegations will agree with the agent majority [159], [160].

Contributions In Chapter 8 we introduce the general concept of Flexible Representative Voting as a means of interpolating between direct voting and representative voting using a special form of proxy voting. When voters are making collective decisions on a set of dichotomous choices, our voting system is show to have a significant advantage compared to other simpler systems, when voters take advantage of the proxy functionality. As part of our mechanism, agents begin by electing a set of representatives, who will then be weighted. We show that without weighting, computing the set of representatives that maximizes the number of issues on which their majority agrees with the voter majority is NP-Hard. In fact, even finding a set of representatives such that at least one representative agrees with the voter majority on every issue is NP-Hard; assuming we even know what the voter majority prefers. This motivates the addition of the proxy functionality to standard representative voting.

CHAPTER 7
DEMOCRATIC FORKING

7.1 Formal Model

Setting We have a set of agents $V = \{v_1, \ldots, v_n\}$. This community will vote on a set of two alternatives $\{A, B\}$ (say, cryptocurrency protocols or locations). However, unlike in most voting scenarios, the agents are not all bound to accept the same winning alternative. Agents have the ability to *fork*, or forge a new community centered around the "losing" alternative. Ultimately, either all of the agents will remain in a single community or they will split into two communities that have accepted opposite alternatives.

Agent Preferences Agents care about what alternative their community adopts and how many people are in their community, but not the identities of the other agents in their community. We can represent agent preferences as total orders over the possible tuples (S, j), where $S \in \{A, B\}$ is the alternative to which they are assigned and $j \in [1, n]$ is the number of agents in their community (including themselves and $j - 1$ other agents). We denote the set of all such tuples by T The preference relation $(A, j) \succ_i (B, k)$ means that agent v_i would prefer to be in a community of size j that accepts alternative A rather than a community of size k that accepts alternative B. Agent preferences are *monotonic* in the size of their community, so given a fixed alternative, they would always prefer to be in a larger community. Formally, for each agent $v_i \in V$ we have $(S, j) \succ_i (S, k)$ for all $1 \leq k < j \leq n$ and $S \in \{A, B\}$. We denote the set of all monotonic total orders over T by $\mathcal{L}(T)$. For $S \in \{A, B\}$, let V_S^* denote the set of agents who prefer (S, n) to (S', n). We will overload notation and use v_i to represent both an agent $v_i \in V$ and their preference ordering $v_i \in \mathcal{L}(T)$. In a similar fashion, V is the set of agents and also the preference *profile*, or collection of the voters' total orders, $V \in \mathcal{L}(T)^n$. We refer to the pair $(V, \{A, B\})$, with

This chapter previously appeared as: B. Abramowitz, E. Elkind, D. Grossi, E. Shapiro, and N. Talmon, "Democratic forking: Choosing sides with social choice," Mar. 2021, arXiv:abs/2103.03652 (Retrieved on 7/18/2021).

This chapter is to appear as: B. Abramowitz, E. Elkind, D. Grossi, E. Shapiro, and N. Talmon, "Democratic forking: Choosing sides with social choice," in *7th Int. Conf. Algorithmic Decision Theory.,* Nov. 2021.

$V \in \mathcal{L}(T)^n$, as a *forking problem*.

Example 2. Suppose we have $n = 3$ agents, $V = \{v_1, v_2, v_3\}$. Consider the preferences of a single agent v_i. By monotonicity, $(A,3) \succ_i (A,2) \succ_i (A,1)$ and $(B,3) \succ_i (B,2) \succ_i (B,1)$ must hold for all agents $v_i \in V$. However, these two orders may be interleaved differently for different agents.

Assignments We refer to the community that accepts alternative A (resp., B) as community A (resp., B). An assignment $f : V \to \{A, B\}$ assigns agents to one of the two communities, and we denote by $f(v_i) \in \{A, B\}$ the community into which agent v_i is placed. The set $\mathcal{F}$ is the set of all 2^n possible assignments, or partitions, of the agents. Voters' preferences over T induce preferences over assignments in $\mathcal{F}$: a voter v_i prefers an assignment f to an assignment g if $(f(v_i), |f^{-1}(f(v_i)|) \succ_i (g(v_i), |g^{-1}(g(v_i)|)$. Given a forking problem $(V, \{A, B\})$ a voting rule $R : \mathcal{L}(T)^n \to \mathcal{F}$ selects an assignment $R(V) = f \in \mathcal{F}$. We let $a = |f^{-1}(A)|$ be the size of community A, and similarly for $b = |f^{-1}(B)|$.

7.2 Stability

Our primary goal is to construct stable assignments. An assignment is stable if no subset of agents has an incentive to move simultaneously to a new community.

Definition 4 (Stability)**.** An assignment $f : V \to \{A, B\}$ is *stable* if there is no assignment $f' : V \to \{A, B\}$ such that each voter v_i with $f'(v_i) \neq f(v_i)$ prefers f' over f.

A voting rule R is *stable* if it returns a stable assignment whenever one exists.

Example 3. Consider two agents, $V = \{v_1, v_2\}$, where $v_1 : (A,2) \succ_1 (A,1) \succ_1 (B,2) \succ_1 (B,1)$ and $v_2 : (B,2) \succ_2 (B,1) \succ_2 (A,2) \succ_2 (A,1)$. Each agent would prefer to be alone at their preferred alternative to being together with the other agent at their less preferred alternative. Thus, the only stable assignment f has $f(v_1) = A$ and $f(v_2) = B$.

7.2.1 Finding Stable Solutions

When preferences are monotonic, there must be at least one stable assignment, and it can be computed in polynomial time.

Theorem 10. *There is a polynomial time assignment rule (Algorithm 2) that finds a stable assignment for any monotonic profile.*

Algorithm 2: General Stable Assignment Rule

$V_A = V$, $V_B = \emptyset$, $a \leftarrow |V_A|$, $b \leftarrow |V_B|$
while `true` **do**
  $k \leftarrow \max\{j : 0 \leq j \leq a, |\{v_i \in V : (B, b+j) \succ_i (A, a)\}| \geq j\}$
  if $k = 0$ **then**
    return $\{V_A, V_B\}$
  else
    Let $X = \{v_i \in V : (B, b+k) \succ_i (A, a)\}$
    $V_B \leftarrow V_B \cup X$, $V_A \leftarrow V_A \setminus X$
    $a \leftarrow |V_A|$, $b \leftarrow |V_B|$
  end if
end while

Proof. Consider Algorithm 2. Let $a = |V_A|$ and $b = |V_B|$. Initially, we place all agents in V_A, so $a = n$ and $b = 0$. If this assignment is not stable, then there exists a subset X of some $k > 0$ agents that all prefer (B, k) to (A, n). We move all these agents to V_B. Monotonicity implies that moving additional agents from V_A to V_B will never cause agents in V_B to want to move back to V_A; thus, as long as we are not in a stable state, there must be a subset of agents at V_A who would prefer to move together to V_B. As long as such a set of agents exists, we continue to move them over together. This procedure halts in at most n steps, and when it halts, the result must be stable, as there is no subset of agents who will move together. A naive implementation of the algorithm loops at most n times, and each computation of the set X of agents to move takes $O(n^2)$ time. □

7.2.2 Elicitation

Algorithm 2 does not use all of the information in agents' preferences. An iterative version of the algorithm can ask only for the information it needs. Instead of assuming that the total orders of all agents are given explicitly in the input, we place all of the agents at A and at each iteration we ask the a remaining agents at A for the minimum value j such that if j agents could be moved to B, they would now prefer the new community $(B, b+j)$ over their current community (A, a). Once an agent has been moved to B there is no need to ask them for any more information about their preferences. In Algorithm 3 we repeatedly query agents about the conditions under which they are willing to leave their current community. Agents indicate their preferences with a single integer that says how many agents would have to move with them for them to prefer the new assignment to the status quo.

Algorithm 3: General Stable Assignment Rule with Iterative Elicitation

$V_A = V$, $V_B = \emptyset$, $a \leftarrow |V_A|$, $b \leftarrow |V_B|$
while `true` **do**
  Ask each agent $v_i \in V_A$ for the smallest value $j \leq a$ such that $(B, b+j) \succ_i (A, a)$
  $k \leftarrow \min\{j : j \in [0, a], |\{v_i \in V : (B, b+j) \succ_i (A, a)\}| \geq j\}$
  if $k = 0$ **then**
    return $\{V_A, V_B\}$
  else
    Let $X = \{v_i \in V : (B, b+k) \succ_i (A, a)\}$
    $V_B \leftarrow V_B \cup X$, $V_A \leftarrow V_A \setminus X$
    $a \leftarrow |V_A|$, $b \leftarrow |V_B|$
  end if
end while

Ideally, we would like to only have to query each agent a small number of times. If agents' preferences are structured, it becomes possible to compute stable assignments with little information. To capture this intuition, we introduce the concept of non-critically-interleaving preferences.

Definition 5 (Non-critically-interleaving)**.** A preference is *non-critically-interleaving* if it is monotonic and $(A, j) \succ (B, n) \succ (B, n-j) \succ (A, j-1)$ or $(B, j) \succ (A, n) \succ (A, n-j) \succ (B, j-1)$ for some $j \in [1, n]$. A profile is *non-critically-interleaving* if it contains only non-critically-interleaving preferences.

When preferences are non-critically-interleaving, we only need to ask each agent if they prefer (A, n) or (B, n), and the minimum value of j such that they would rather be at their preferred alternative in a coalition of size j than at the other alternative in a coalition of size n. From this information the relevant part of the preference order of each agent can be inferred, and so Algorithm 3 will compute a stable assignment.

7.2.3 Uniqueness

While at least one stable assignment must exist for all monotonic profiles (Theorem 10), it is not necessarily unique.

Example 4. Let $V = \{v_1, v_2, v_3, v_4\}$ be a set of four agents with preferences that contain the following prefixes, respectively:

- $v_1 : (B, 4) \succ_1 (B, 3) \succ_1 (A, 4) \succ_1 (B, 2) \succ_1 (A, 3) \succ_1 \cdots$

- $v_2 : (B,4) \succ_2 (B,3) \succ_2 (B,2) \succ_2 (A,4) \succ_2 (B,1) \succ_2 \cdots$
- $v_3 : (A,4) \succ_3 (A,3) \succ_3 (A,2) \succ_3 (B,4) \succ_3 (A,1) \succ_3 \cdots$
- $v_4 : (A,4) \succ_4 (A,3) \succ_4 (B,4) \succ_4 (A,2) \succ_4 (B,3) \succ_4 \cdots$

Regardless of how the remainder of the preference profile is filled, as long as monotonicity is maintained, there are at least three stable assignments: (1) all agents at A; (2) all agents at B; or (3) v_1 and v_2 at B and v_3 and v_4 at A.

We would like to identify conditions under which a profile admits a unique stable assignment. One extreme case is when preferences are non-interleaving.

Definition 6 (Non-interleaving)**.** A preference order is *non-interleaving* if it is monotonic and either $(A,1) \succ (B,n)$ or $(B,1) \succ (A,n)$. A profile is *non-interleaving* if it contains only non-interleaving preference orders.

The profile in Example 3 is an instance of a non-interleaving profile. If an agent's preference is non-interleaving, then their choice of community is independent of the other agents: they would rather be alone at their preferred alternative than with everyone else at the other alternative. Thus, their preference is described by a single bit of information: it suffices to know whether they are in V_A^* or in V_B^*. Non-interleaving preferences can be viewed as a degenerate case of non-critically-interleaving preferences when $j = 1$.

Observation 4. When preferences are non-interleaving, there is a unique stable assignment.

Proof. The only stable assignment assigns to all agents in V_A^* to A and all agents in V_B^* to B. Otherwise, an agent assigned to the opposite community will wish to move, even if on their own. □

Non-interleaving preferences can be generalized to domains of preferences that guarantee unique stable assignments. Informally, we say that an agent is k-loyal to an alternative S if they prefer to be at S with k other agents to being at the other alternative in a coalition of size n.

Definition 7 (k-Loyalty)**.** An agent $v_i \in V$ is *k-loyal* to alternative S, $k \in [n]$, if $v_i \in V_S^*$ and $(S,k) \succ_i (S',n)$ for $S' \neq S$.

When all agents are sufficiently loyal to their preferred alternatives, there is a unique stable assignment.

Proposition 1. *Suppose there exist some k_1, k_2 such that $k_1 \leq |V_A^*|$, $k_2 \leq |V_B^*|$, every agent in V_A^* is k_1-loyal, and every agent in V_B^* is k_2-loyal. Then there is a unique stable assignment.*

Proof. By construction, any stable assignment must have all agents in V_A^* assigned to A, because otherwise those assigned to B would prefer to move together to A, forming a coalition of size $|V_A^*| \geq k_1$ at A. Symmetrically, any stable assignment must have all agents in V_B^* assigned to B, as otherwise those assigned to A would prefer to move together to B, forming a coalition of size at least $|V_B^*| \geq k_2$ at B. □

Proposition 1 holds because all agents must necessarily be assigned to their preferred alternative. We now examine a sub-domain of non-critically-interleaving preferences in which there is always a unique stable assignment, but not all agents are necessarily assigned to their preferred alternative.

Proposition 2. *Suppose agents' preferences are non-critically-interleaving. Let*

$$V_A' = \arg\max_{U \subseteq V} |\{v_i \in U : (A, |U|) \succ (B, n)\}| \,, \tag{7.1}$$

$$V_B' = \arg\max_{U \subseteq V} |\{v_i \in U : (B, |U|) \succ (A, n)\}| \,. \tag{7.2}$$

If none of the agents in $V_A^* \setminus V_A'$ are $(n - |V_B'|)$-loyal and none of the agents in $V_B^* \setminus V_B'$ are $(n - |V_A'|)$-loyal, then there is a unique stable assignment.

Proof. Note first that, by monotonicity, the set arg max in the definition of V_A' and V_B' is a singleton, so V_A' and V_B' are well-defined. As with Proposition 1, for any assignment to be stable it must assign all agents in V_A' to A and those in V_B' to B. For the remaining agents, they must necessarily be assigned to the opposite alternative, because there cannot be enough agents at their most preferred alternative for them to stay there. □

The maximal class of profiles for which there is a unique stable solution is still more general than those we describe above. We can use Algorithm 2 to characterize the set of profiles that admit a unique stable assignment. Let R_A be the assignment rule given by Algorithm 2, and let R_B be the complementary assignment rule that starts with all agents at B and iteratively moves them to A in the same manner.

Theorem 11. *Algorithm 2 (R_A) and the reverse assignment rule (R_B) return the same assignment if and only if the profile admits a unique stable assignment.*

Proof. If the stable assignment is unique, then both R_A and R_B must return this assignment. We now show that if R_A and R_B return the same stable assignment, then it must be the unique stable assignment. Let V_A^1 and V_B^1 be the communities in the stable assignment $f_1 = R_A(V)$. Let V_A^2 and V_B^2 be the communities according to a different stable assignment f_2. By monotonicity and the properties of Algorithm 1 we have $V_A^1 \subsetneq V_A^2$. Consider the set of agents $V_A^2 \setminus V_A^1$, and in particular, the agent(s) in this set that were the first to be moved to B by R_A. At the beginning of the iteration in which they were moved, the number of agents at A had to be at least $|V_A^2|$ (before moving). This contradicts the claim that f_2 is stable, as there are agents in V_A^2 preferring to move together to B. □

7.2.4 Cohesiveness

Not all stable assignments may be equally attractive. In real life, forking comes at a cost, such as the need to replicate infrastructure and to carve out or abandon intellectual property or goodwill, as well as the social and emotional cost of separation. The cost of forking within our framework is implicit in the preferences of the agents. In line with the monotonicity of preferences, it is natural that the community may want to avoid forks when possible. When it is desirable to avoid forking, we prefer stable assignments that place all agents at the same alternative over those that fork. We call these non-forking assignments. A profile is said to be *cohesive* if it admits at least one non-forking stable assignment; otherwise, we say that a profile is *forking*. The profile in Example 4 is cohesive, although it also permits a forked stable assignment.

The assignment in Example 3 is stable, but the input profile is forking, because no stable assignment exists with all agents in one community. For a profile to be cohesive there must be at least one alternative (w.l.o.g, A) such that for all $j \leq n$, there are fewer than j agents who prefer (B, j) over (A, n). The following example shows a cohesive profile with no stable forked assignments.

Example 5. Let $V = \{v_1, v_2\}$, where $(A,2) \succ_1 (B,2) \succ_1 (A,1) \succ_1 (B,1)$ and $(B,2) \succ_2 (A,2) \succ_2 (B,1) \succ_2 (A,1)$. Each agent would prefer to be together with the other agent at their less preferred alternative rather than alone at their preferred alternative. This profile is cohesive, and only admits stable assignments that are non-forking.

7.3 Strategyproofness

So far we have considered the existence and the possibility of efficiently computing stable assignments when agents report their preferences truthfully. Another important question is whether there exist strategyproof stable assignment rules, i.e., rules that output stable assignments and do not incentivize the agents to misreport their true preferences.

Definition 8 (Strategyproofness)**.** A rule R is *strategyproof over domain* $D \subseteq \mathcal{L}(T)^n$ if for all profiles $V \in D$ and assignments $f = R(V)$, there is no agent $v_i \in V$ that can unilaterally change her preference order to v_i', creating a new profile V' such that she prefers $f'(v_i)$ over $f(v_i)$, where $f' = R(V')$.

Similarly, a rule is k-strategyproof in our setting if no subset of agents of size k can simultaneously report false preferences to yield an assignment they all prefer. Naturally, k-strategyproofness implies $(k-1)$-strategyproofness.

Definition 9 (k-Strategyproofness)**.** A rule R is *k-strategyproof over domain* $D \subseteq \mathcal{L}(T)^n$ if for all profiles $V \in D$ and assignments $f = R(V)$, there is no subset of agents $U \subseteq V$ of size $|U| \leq k$ that can simultaneously change their preferences, creating a new profile V' such that each agent $v_i \in U$ prefers $f'(v_i)$ over $f(v_i)$, where $f' = R(V')$.

For the domain of all monotonic profiles, no strategyproof stable rules exist. This can be seen from Example 5. In this example there are two stable assignments, one creating $(A, 2)$ and the other creating $(B, 2)$. Suppose the agents are both assigned to B. If v_1 were to change their reported preferences to $(A, 2) \succ_1 (A, 1) \succ_1 (B, 2) \succ_1 (B, 1)$, then $(A, 2)$ would become the only stable assignment for the new profile, which v_1 clearly prefers over $(B, 2)$. This profile is symmetric, so if the agents were to be assigned to $(A, 2)$ (by some tie-breaking mechanism) then v_2 has the opportunity to be strategic. In fact, no strategyproof stable assignment rule can exist for any domain containing a profile that admits two or more stable assignments.

Theorem 12. *No assignment rule can be strategyproof over a domain that includes a profile that admits more than one stable assignment.*

Proof. Suppose profile V permits at least two stable assignments, and our assignment rule R picks one of them, $f_1 = R(V)$. Let f_2 be the closest stable assignment to f_1, in the sense that

there is no other assignment f_3 such that $|f_3^{-1}(A)|$ is between $|f_1^{-1}(A)|$ and $|f_2^{-1}(A)|$, and consequently no $|f_3^{-1}(B)|$ between $|f_1^{-1}(B)|$ and $|f_2^{-1}(B)|$ (indeed, $f^{-1}(A) + f^{-1}(B) = n$ for any f). For brevity, let $V_A^1 = f_1^{-1}(A)$, $V_A^2 = f_2^{-1}(A)$, $V_B^1 = f_1^{-1}(B)$, $V_B^2 = f_2^{-1}(B)$. Assume that $|V_A^1| > |V_A^2|$ and $|V_B^2| > |V_B^1|$; later we will see that this assumption is without loss of generality.

From monotonicity, we know that $V_A^2 \subset V_A^1$ and $V_B^1 \subset V_B^2$. Consider an agent $v_i \in V_A^1$ with $(B, |V_B^2|) \succ_i (A, |V_A^1|) \succ_i (B, |V_B^1| + 1)$. At least one such agent must exist because otherwise f_2 could not be stable, as all the agents in $V_A^1 \cap V_B^2$ would prefer to move together to A rather than stay in $(B, |V_B^2|)$. If v_i commits to B by falsely reporting that they prefer $(B, 1)$ to (A, n), then they must be assigned to B (as otherwise the assignment would not be stable). By construction, however, no intermediate stable state can exist between f_1 and f_2, so agents will prefer to move from V_A^1 to B until it is of size $|V_B^2|$, which is what v_i preferred. By symmetry, if our rule R had picked assignment f_2 instead of f_1 then we would have the same result; thus, our assumption that $|V_A^1| > |V_A^2|$ and $|V_B^2| > |V_B^1|$ is indeed without loss of generality. □

The above theorem means that the domain consisting of those profiles that admit a unique stable assignment is the maximal domain for which a stable assignment rule can be strategyproof. Next, we show that, whenever a profile admits a unique stable assignment, Algorithm 2 is strategyproof.

Lemma 9. *Algorithm 2 is strategyproof over the domain of all profiles that admit a unique stable assignment.*

Proof. Consider a run of Algorithm 1 in which it assigns some agent v_i to B and outputs the assignment f. First, observe that v_i cannot misreport their preferences so that they would end up in a larger community at B that they prefer. If the agents assigned to A do not move at any iteration, then v_i moving at an earlier to later iteration has no effect on them. And, due to monotonicity, v_i staying at A would not further entice anyone to move to B.

Second, we want to show that agent v_i cannot manipulate the outcome so that it will be assigned to a larger community at A that it prefers. Suppose that at the beginning of the iteration when v_i is moved from A to B, the size of the community at A is a. The agents who were moved from A to B at an iteration before the iteration at which v_i is moved will be assigned to B regardless of what v_i reports. In general, agents moved to B together at

one iteration must end up at B regardless of the preferences of those moved to B at later iterations and those who stay at A. As a consequence, v_i can never induce an assignment with a community at A of size greater than a. Since v_i was moved at the iteration when the size of A was a, it must be to a community B that they prefer over (A, a). Therefore no agent $v_i \in B$ can deviate profitably from their true preferences.

It remains to show that no agent assigned to A can benefit from strategic behavior. By symmetry, if there is a unique stable assignment, then if Algorithm 2 starts with all agents at A and moves them in batches to B, or starts at B and moves them in batches to A, then it must return the same assignment. We can therefore use the same argument as above for agents assigned to A according to the algorithm that initializes A and B in the opposite way. □

The result extends to n-strategyproofness, or group-strategyproofness, since, if we consider any coalition of agents assigned to B by Algorithm 2, and look at only those who were moved first (in the same iteration as one another, but before everyone else in the coalition who was moved), then they have no incentive to misreport their preferences for the same reason as the agent in the proof of Lemma 9. By combining this with Theorem 12, we arrive at our main result.

Theorem 13. *Algorithm 2 is group-strategyproof over the domain of all profiles that admit a unique stable assignment.*

Just how common are profiles that admit strategyproof stable assignment rules? One specific domain restriction that implies group-strategyproofness of Algorithm 2 is the domain of non-interleaving profiles (recall that, for this domain restriction, placing all agents at their preferred alternative is stable). Non-interleaving preferences are indeed very extreme, in that agents ignore each other completely. However, if we relax this extreme constraint on preferences even the slightest bit, we can lose strategyproofness.

Definition 10 (Minimally-interleaving)**.** A preference order is *minimally-interleaving* if it is monotonic and $(A, 2) \succ (B, n) \succ (A, 1) \succ (B, n-1)$ or $(B, 2) \succ (A, n) \succ (B, 1) \succ (A, n-1)$. A profile is a *minimally-interleaving* if it contains only non-interleaving and minimally-interleaving preferences.

Minimally-interleaving preferences can be interpreted as just barely extending non-interleaving preferences to allow that agents may be willing to go with their less preferred

alternative if they would otherwise be alone with their more preferred alternative. Note that the minimally-interleaving domain is still a severely restricted domain. In particular, it allows each agent to specify only one of four possible orders. However, it turns out that if we allow just minimal interleaving, then there is no assignment rule that is both stable and strategyproof.

Observation 5. There is no assignment rule that is both stable and strategyproof for all minimally-interleaving preference profiles.

Proof. Let R be a stable assignment rule and consider the profile from Example 5: $v_1 : (A,2) \succ_1 (B,2) \succ_1 (A,1) \succ_1 (B,1)$; $v_2 : (B,2) \succ_2 (A,2) \succ_2 (B,1) \succ_2 (A,1)$. Note that, indeed, this profile is minimally-interleaving. Observe that the only stable assignments are the two that place both agents in the same community. Since there are two stable assignments, Theorem 12 implies that R cannot be strategyproof for this profile. For illustrative purposes, consider the case in which v_1 votes strategically by reporting $v_1' : (A,2) \succ_1 (A,1) \succ_1 (B,2) \succ_1 (B,1)$. The only stable assignment places both agents at A, creating $(A,2)$, which v_1 prefers to $(B,2)$. So if R placed both agents at $(B,2)$, it cannot be strategyproof. □

We say that a profile is *k-interleaving* if it may contain preference orders in which $(S',n) \succ \ldots \succ (S,n) \succ (S',k)$, but not in which $(S',n) \succ \ldots \succ (S,n) \succ (S',k+1)$, where $S' \neq S$. Hence, non-interleaving preferences are equivalent to 0-interleaving; minimally-interleaving preferences are the same as 1-interleaving; and n-interleaving is the domain of all monotonic preferences. Naturally, the set of all k-interleaving preferences encompasses all $(k-1)$-interleaving preferences, so no strategyproof stable assignment rule can exist for $k \geq 1$. While non-interleaving is a sufficient condition for strategyproofness, the next example demonstrates that it is not a necessary condition.

Example 6. Consider two agents, $V = \{v_1, v_2\}$, where $v_1 : (A,2) \succ_1 (A,1) \succ_1 (B,2) \succ_1 (B,1)$ and $v_2 : (B,2) \succ_2 (A,2) \succ_2 (B,1) \succ_2 (A,1)$. Any stable assignment must have v_1 at A, independent of the preferences of v_2. Agent v_2 would prefer to be at A with v_1 to being alone at B, so the only stable assignment has both agents at A, and neither agent has an incentive to be strategic.

Notice that our interleaving conditions apply to the preference order of each agent individually. Example 6 suggests that we should instead consider restrictions on the profile as

a whole. While we know that the necessary and sufficient conditions for stable strategyproof assignment rules to exist is that there be a unique stable assignment, characterizing the profiles for which this occurs is an interesting challenge.

7.4 Forking With More Than Two Alternatives

So far, we focused on the case of two alternatives. We conclude the paper with two observations about the general case: (1) stable assignments are no longer guaranteed to exist; (2) deciding whether an assignment is stable is NP-complete.

Proposition 3. *There exist monotonic profiles with no stable assignment.*

Proof. Consider the problem with three agents $V = \{v_1, v_2, v_3\}$, three alternatives $\{A, B, C\}$, and a following profile:

- $v_1 : \cdots \succ_1 (B,2) \succ_1 (A,2) \succ_1 (A,1) \succ_1 (B,1) \succ_1 (C,3) \succ_1 \cdots$
- $v_2 : \cdots \succ_2 (C,2) \succ_2 (B,2) \succ_2 (B,1) \succ_2 (C,1) \succ_2 (A,3) \succ_2 \cdots$
- $v_3 : \cdots \succ_3 (A,2) \succ_3 (C,2) \succ_3 (C,1) \succ_3 (A,1) \succ_3 (B,3) \succ_3 \cdots$

Assume for contradiction that this profile admits a stable assignment f. As v_2 prefers $(B,1)$ to $(A,3)$, the assignment f cannot assign v_2 to A. By considering v_1 and v_3, we conclude that $|f^{-1}(s)| < 3$ for every $S \in \{A, B, C\}$. Suppose $f(v_1) = A$, $f(v_3) = A$. Then we have $f(v_2) = B$ because $(B,1) \succ_2 (C,1)$. But in this case, v_1 would prefer to move to B since $(B,2) \succ_1 (A,2)$. By the same reasoning $|f^{-1}(S)| \neq 2$ for each $S \in \{A, B, C\}$. The only remaining option is to have one voter at each alternative. Let v be the voter at A. If $v = v_1$, then v_1 prefers to move to B and if $v = v_3$ then v_3 would prefer to move to C. Finally, if $v = v_2$ then v_1 prefers to join v_2 at A. □

From a complexity-theoretic perspective, it is then natural to ask if there are efficient algorithms for (a) checking whether a given assignment is stable, and (b) deciding if a given profile admits a stable assignment. It turns out that, while the answer to the first question is 'yes', the answer to the second question is likely to be 'no'.

Proposition 4. *We can decide in polynomial time whether a given assignment for a forking problem is stable.*

Proof. Note first that if an assignment f is not stable, then this can be witnessed by a deviation in which all deviating agents move to the same alternative (say, A). Indeed, the agents who deviate from f by moving to A would find this move beneficial even if other agents did not move (in particular, due to monotonicity, they benefit from other agents not moving away from A). Thus, to decide if a given assignment f is stable, it suffices to consider deviations that can be described by a pair (S, n_S), where S is an alternative and $n_S > f^{-1}(S)$. For each such pair, we need to check if there are $n_S - f^{-1}(S)$ agents who are currently not assigned to S, but prefer (S, n_S) to their current circumstances. □

Proposition 5. *Deciding whether a forking problem admits a stable assignment is NP-complete.*

Sketch. By Proposition 4, our problem is in NP. For hardness we adapt the reduction argument of Darmann [97, Th. 3], establishing NP-hardness for the core stability problem in o-GASP with increasing preferences. That construction makes use of so-called void activities, which are available in o-GASP, but not in forking problems. In the profile constructed for [97, Th. 3], the occurrence of void activities in each agent's preference needs to be replaced by $(S^*, 1)$, where S^* denotes the top alternative in the agent's preference. □

Our hardness reduction produces an instance where the number of alternatives is linear in the number of voters. The complexity of finding a stable assignment for a fixed number of alternatives (e.g., $m = 3$) remains open.

7.5 Conclusions and Future Work

In the real world, communities sometimes fracture, or fork. This can happen for many reasons, but can generally be seen as a consequence of the decisions the community has made. If agents associate freely, with the ever-present option of leaving, then we can account for this possibility in collective decision making procedures themselves. This enables minorities to threaten a fork in protest and wield their collective power against the tyranny of the majority, while giving the majority an opportunity to concede to prevent a fork. Such a forking process also facilitates the emergence of new communities, as it may be easier to sprout a community from an existing one rather than to build one from scratch.

We have shown that, while it may not be difficult to find stable partitions of a set of agents, constructing strategyproof rules is only possible in restricted domains. While the

necessary and sufficient conditions for strategyproofness remain an interesting open question, we have identified a range of circumstances that are sufficient for strategyproofness. Lastly, we have shown that efficient preference elicitation is possible and desirable.

The social choice setting we considered is, to the best of our knowledge, novel and our work has only made the first steps towards its analysis. Several directions for future research present themselves: (1) first, as mentioned above, settling the question about the domain restrictions that are necessary and sufficient for the existence of stable and strategy-proof assignment rules is a priority; (2) second, natural generalizations of the setting we propose will be worth investigating—e.g., settings in which agents can be assigned to several alternatives (similarly to how, e.g., large miners can be present in several forks), or settings in which the identities of the agents matter (as agents may wish to fork with other specific agents); and (3) third, studying mechanisms for the converse problem, in which several communities could merge into a new one.

CHAPTER 8
FLEXIBLE REPRESENTATIVE VOTING

8.1 Introduction

With direct voting, agents report their preferences and a voting rule is applied to determine the outcome. Consider a set of agents N voting over a set of alternatives X. A voting rule F aggregates their preferences, collectively profile $\mathcal{V}$, to produce $F(\mathcal{V}) = y^*$. By default, if a voter does not participate by actively casting their vote they are usually not counted. In some cases active votes for abstentions count towards quorum, and these are treated differently than if the individuals do not cast votes at all. Direct voting systems require agents to opt-in for every issue under consideration to be counted for each one.

Representative voting lessens the attention and effort required of voters, making it more practical at large scale than direct voting. Instead of needing to vote on every issue, voters can vote once per election cycle and rely on their elected representatives to make many decisions on their behalf. This is particularly important when a society must make many complex decisions. During an election there are two kinds of agents, the general voting population V and the set of candidates C seeking to get elected. For ease of presentation it is assumed that the sets of voters and candidates are disjoint. An election rule F_C aggregates voter preferences over the candidates to elect a subset of them as representatives: $F_C(\mathcal{V}_\mathcal{C}) = D \subseteq C$. The preferences of the representatives over the alternatives X are then aggregated to yield an outcome $F(\mathcal{D}) = y^*$.

There are many different multi-winner voting rules that can be used to elect representatives. Which subset of candidates gets elected can depend on how voters choose who to vote for, how they express their votes, and what candidates are available to choose from. In this chapter we will assume that voters preferences over candidates are derived from the

This chapter previously appeared as: B. Abramowitz and N. Mattei, "Flexible representative democracy: An introduction with binary issues," in *Proc. 28th Int. Joint Conf. Artif. Intell., Macao, China*, S. Kraus, Ed., Aug. 2019, pp. 3–10, doi: 10.24963/ijcai.2019/1.

This chapter previously appeared as: B. Abramowitz and N. Mattei, "Flexible representative democracy: An introduction with binary issues," Apr. 2019, arXiv:1811.02921 (Retrieved on 7/18/2021).

similarity in their respective preferences regarding possible outcomes. Representatives are seen as a device to reduce the effort required of voters. Our goal is to design voting systems that emulate direct voting by producing the same outcomes that would be achieved if direct voting had been used. If we fix a model of agent behavior and make assumptions about how agent preferences are generated, we can begin to compare the performance of various multi-winner rules according to whatever objectives we specify.

In a traditional representative voting system, all representatives' votes are given equal weight, though there are many situations in which this may not be desirable. For example, if each representative represents a district, it might be desirable to weight them according to the number of voters in their district [161]. The central idea is that representatives should be weighted in some way proportional to the number of people they represent. Taking this to the limit, in proxy voting systems, voters can individually select which representative should represent them, and the representatives can be weighted by the number of voters who choose them as their proxy. Our proposal is a novel voting system that uses a form of proxy voting to interpolate between direct and representative voting.

8.2 Single-Issue Dichotomous Choice

Vote Profiles Let $X = \{0, 1\}$ be a set of two alternatives, let V be a set of n voters, and let C be a set of m candidates. For simplicity of presentation, we assume that the set of voters and candidates are disjoint. The voters' collective profile of ballots over the alternatives is given by $\mathcal{V}$, indexed by the voters such that $\mathcal{V}(i) \in \{0, 1\}$ is the vote of voter $i \in V$. Similarly, $\mathcal{C}$ is the profile of candidate ballots, indexed by the candidates. The number of voters who vote for 1 is denoted by $[\![\mathcal{V}]\!] = \{i \in V : \mathcal{V}(i) = 1\}$, and similarly for $[\![\mathcal{C}]\!] = \{i \in C : \mathcal{C}(i) = 1\}$. Following May's Theorem [162], we presume that a direct vote would use majority voting as the voting rule F. We therefore assume, without loss of generality, that $[\![\mathcal{V}]\!] > \frac{n}{2}$, so $F(\mathcal{V}) = 1$ is the alternative that wins by majority vote under direct voting.

8.2.1 Electing Unweighted Representatives

For any representative system, a primary consideration must be what multi-winner voting rule to use for electing representatives. Even in the simple world of dichotomous choices, this can be tricky, and depends on how voters decide which candidates to support.

One might suspect that simply electing the most popular candidates - by some measure of popularity - would do well at emulating direct voting. Unfortunately, as we will show, the problem is not so simple. For example, if the goal is to recover what would happen under direct voting, plurality rule is a bad choice.

Example 7. Suppose $\mathcal{V} = \{1, 1, 1, 1, 0, 0, 0\}$ and $\mathcal{C} = \{1, 1, 0\}$. The voters will elect a single representative. Each voter votes for a single candidate who agrees with them on which alternative should be selected. If voters 1-2 vote for candidate 1, voters 3-4 vote for candidate 2, and voters 5-7 vote for candidate 3, then candidate 3 will be elected under Plurality, and will vote 0 so the winning alternative will be 0.

Even though a voter's vote for a candidate implicitly captures all of the same information that would be present if they voted directly, not all of this information is maintained within the profile of the elected representatives.

When electing only a single representative from a set of candidates, there is an obvious election rule that works. Let all voters provide a list of all candidates of whom they approve, assuming the approve of a candidate if and only if they vote for the same alternative, and elect the candidate with the greatest number of approvals. However, electing a single representative is an edge case. It is easy to see how the issue with Example 7 generalizes to electing multiple representatives, but the failure of the approval rule is more subtle.

Unfortunately, all methods of electing a subset of candidates as representatives, including electing the top k most highly approved candidates, face a serious issue when the number of candidates is small relative to the number of representatives who must be elected. If the number of representatives that must be selected is a large fraction of the candidates, so there are relatively few options to choose from, the quality of the set of representatives becomes more dependent on the quality of the original set of candidates.

Example 8. Suppose $\mathcal{V} = \{1, 1, 1, 1, 0, 0, 0\}$ and $\mathcal{C} = \{1, 0, 0, 0, 0\}$. The voters will elect $k = 3$ representatives. No matter what election rule is used to elect D, we will have $[\![\mathcal{D}]\!] < \frac{k}{2}$, so majority rule will result in the representatives selecting $y^* = 0$.

In the extreme case, no candidate agrees with the voter majority, so no election can result in a positive outcome.

8.2.2 Electing Weighted Representatives

One way to think of the election of representatives is as a compression algorithm that compresses the profile of voters' ballots into a profile of representatives' ballots. The choice of multi-winner voting rule determines how "lossy" the compression is as it is "read" by the decision rule that aggregates the representatives' votes. Viewed this way, the weighting of representatives is like the reverse process of decompression to recover the relevant aspects of the original voter profile. In Example 7, this decompression is not possible, because the single elected representative will always vote for 0. However, as long as there is at least one representative who votes for the hypothetical winner under direct voting (1), all hope is not lost.

Example 9. Suppose $\mathcal{V} = \{1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0\}$, and $\mathcal{C} = \{1, 1, 1, 1, 0, 0\}$. The voters are going to elect 3 representatives according to their plurality scores. If voters 1-6 vote for candidate 1, voters 7-8 vote for candidate 5, and voters 9-10 vote for candidate 6, then the elected representatives will have profile $\mathcal{D} = \{1, 0, 0\}$. If the representatives use an unweighted majority vote 0 will win. However, if the representatives are weighted by the number of voters who voted for them in the election, then 1 wins as desired.

While Example 9 gives some hope, it rests on the fact that the majority of voters supported the same candidate. However, when voters support candidates who are not elected, we must determine what happens to their votes.

Example 10. Suppose $\mathcal{V} = \{1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0\}$, and $\mathcal{C} = \{1, 1, 1, 1, 0, 0\}$. The voters are going to elect 3 representatives according to their plurality scores. If voters 1-3 vote for candidate 1, voter 4 votes for candidate 2, voter 5 voters for candidate 3, voter 6 votes for candidate 4, voters 7-8 vote for candidate 5, and voters 9-10 vote for candidate 6, then the elected representatives will have profile $\mathcal{D} = \{1, 0, 0\}$. If the representatives use an unweighted majority vote 0 will win and if the representatives are weighted by the number of voters who voted for them in the election, then 0 still wins.

One solution would be to allow unelected candidates to transfer the power of the votes they receive to elected representatives. This is a form of transferable delegation, which is taken to the extreme in Liquid Democracy [108]–[111], [115]–[117], [122], [124], [130], [163]–[167]. When voting on only a single issue with two alternatives, voters' votes over

candidates can easily contain all of the information that their direct votes would contain, assuming that their preferences over candidates is derived from similarity in their preferences over alternatives. However, when there are more alternatives or multiple issues, this is no longer the case.

8.2.3 Flexible Representative Voting

We have just considered what happens when representatives' weights can be determined directly from the election, but this is not always ideal. Some voters may have voted for candidates who were not elected, so it needs to be determined how their voting units are assigned or whether they are not assigned at all by default. There are additional complications; perhaps voters do not know the preferences of the representatives on the issue at hand before the election, or representatives change their minds after the election. A key observation here is that it may be undesirable for representatives to have the same weight for all issues considered until the next election.

We now propose the system of Flexible Representative Voting. First, voters elect a subset of electoral candidates as representatives: $F_C(\mathcal{V}_\mathcal{C}) = D \subseteq C$. After the election, the elected representatives are assigned weights $\bar{W}_D$. These weights signify voting units belonging to the voters being assigned to the representatives. The voters then have the chance to update these weights by *delegation*, thereby re-assigning the voting units belonging to them to representatives as they see fit. After the delegation phase is over, the representatives have updated weights W_D. Finally, a decision rule is applied to the weighted representatives' preference profile to yield a decision outcome $F(\mathcal{D}, W_D)$

While Flexible Representative Voting is meant for settings where many issues are being decided, we analyze what happens with a single dichotomous choice before looking at the multi-issue setting. In the multi-issue setting, voters can update their delegations so that they differ on each issue.

Suppose our n voters V have a set D of k representatives. Instead of distributing voting units according to the electoral ballots, lets suppose each voter receives k voting units which are distributed uniformly over the representatives by default; each representative receives one voting unit from each voter. Now each representative has weight n. If we were to stop here and use weighted majority to select a winning alternative, this is functionally identical to unweighted representative voting. We would like to know if and how the voters' active

delegations can change the outcome, and the circumstances under which they change it in the desirable direction (from 0 to 1).

Deterministic Delegation Recall the the majority of voters prefer 1 over 0, so $[\![\mathcal{V}]\!] > \frac{n}{2}$. However, we are not guaranteed that $[\![\mathcal{D}]\!] > \frac{k}{2}$. How many voters who prefer 1 must use their option to actively delegate to guarantee that the outcome is 1?

Observation 6. Assume that voting units are distributed uniformly by default under $\bar{W_D}$. If a voter delegates, they are assumed to delegate all of their voting units to representatives who agree with them. The outcome selected by the weighted majority vote of the representatives is guaranteed to be 1 if the fraction of voters who prefer 1 that delegate is at least $\max\{\frac{nk-2[\![\mathcal{C}]\!][\![\mathcal{V}]\!]}{2(k-[\![\mathcal{C}]\!])}, 0\}$

Probabilistic Delegation Suppose each voter $i \in V$ chooses to delegate (deviate from the default) with independent probability p_i. Recall that delegations are assumed to be incisive, in that if a voter delegates then all of their voting units go to representatives who agree with them. Assume, without loss of generality, that each voter gets 1 divisible voting unit, originally distributed evenly such that each representative receives $\frac{1}{k}$ from each voter. Let $w_i(1) \in [0, 1]$ be the fraction of voter i's voting units assigned to representatives who agree with the voter majority (vote for 1); assigned either through delegation or by default. If voter i does not delegate, then $w_i(1) = [\![\mathcal{D}]\!]/k$. If i is in the majority and delegates then $w_i(1) = 1$, and if i is in the minority and delegates then $w_i(1) = 0$.

Let $W_D(1) = \sum_{i \in V} w_i(1)$ be the total voting units assigned to the representatives who vote for 1. Let $\mu = E[W_D(1)] = \sum_{i \in V : \mathcal{V}(i)=1} (p_i + (1 - p_i)\frac{[\![\mathcal{D}]\!]}{k}) + \sum_{i \in V : \mathcal{V}(i)=0} (1 - p_i)\frac{[\![\mathcal{D}]\!]}{k}$ be the expected value of the total voting weight assigned to representatives who agree with the voter majority.

Theorem 14. *Suppose there are an odd number of voters n, odd number of representatives k, and voters who delegative do so entirely to representatives who agree with them. Further suppose each voter $i \in V$ delegates with probability p_i such that $\mu > n/2$. Then the probability that the outcome agrees with the voter majority is bounded by $Pr(y^* = 1) \geq 1 - e^{-(n-2\mu)^2/4n}$.*

Proof. Recall that $w_i(1) \in [0, 1]$ is the amount of voting weight voter i gives to representatives who agree with the voter majority and $W_D(1) = \sum_{i \in V} w_i(1)$. Given some tie breaking rule,

we have that $Pr(y^* = 1) = Pr(W_D(1) > n/2) + Pr(y^* = 1|W_D(1) = n/2) \cdot Pr(W_D(1) = n/2)$. First we show that $Pr(W_D(1) = n/2) = 0$, then we give a lower bound for $Pr(W_D(1) > n/2)$.

Lemma 10. *If n is odd, k is odd, and all voters who delegate do so entirely to representatives who agree with them, then no ties are possible.*

Proof. Let $w'_i(1) = k \cdot w_i(1)$ where $w_i(1) \in [0, 1]$ is the amount of weight (voting power) voter v_j assigns to representatives who agree with the voter majority. If i does not delegate then $w'_i(1) = [\![\mathcal{D}]\!]$. If i delegates and is in the voter majority then $w'_i(1) = k$, and if i delegates and is in the voter minority then $w'_i(1) = 0$. Therefore, $\forall i \in V : w'_i(1) \in \{0, [\![\mathcal{D}]\!], k\}$.

Let $W'_D(1) = \sum_{i \in V} w'_i(1)$ and $W'_D(0) = \sum_{i \in V}(k - w'_i(1))$. Then $W'_D(1)$ and $W'_D(0)$ are non-negative integers and $W'_D(1) + W'_D(0) = kn$. Since kn is odd, it must be that $W'_D(1)$ and $W'_D(0)$ have opposite parity and so they cannot be equal. Therefore $W_D(1) = W'_D(1)/k \neq W_D(0) = W'_D(0)/k$, meaning the total amounts of weight delegated to the representatives on either side of the issue cannot be equal, so no ties may occur. □

Given that no ties are possible, we have that $Pr(y^* = 1) = Pr(W_D(1) > n/2)$. Remember that $W_D(1) = \sum_{i \in V} w_i(1)$ where $w_i(1)$ is the total weight that i assigns to representatives who agree with the voter majority. If i prefers 1 then $E[w_i(1)] = p_i + (1 - p_i)\frac{[\![\mathcal{D}]\!]}{k}$, else if i prefers 0 then $E[w_i(1)] = (1-p_i)\frac{[\![\mathcal{D}]\!]}{k}$. Let $\mu = E[W_D(1)]$ be the expected total weight assigned to representatives who agrees with the voter majority, then $\mu = \sum_{i \in V : \mathcal{V}(i) = 1} (p_i + (1 - p_i)\frac{[\![\mathcal{D}]\!]}{k}) + \sum_{i \in V : \mathcal{V}(i) = 0} (1 - p_i)\frac{[\![\mathcal{D}]\!]}{k}$. We now use the fact that $Pr(W_D(1) > n/2) = 1 - Pr(W_D(1) \leq n/2)$. Let $\delta = (2\mu - n)/2\mu$. If $\mu > n/2$, then $\delta > 0$. This allows us to apply a Chernoff bound to derive our lower bound, $Pr(W_D(1) > n/2) = 1 - Pr(W_D(1) \leq n/2) = 1 - Pr(W_D(1) \leq (1 - \delta)\mu) \geq 1 - e^{-\delta^2\mu^2/n} = 1 - e^{-(2\mu - n)^2/4n}$. □

8.3 Multi-Issue Dichotomous Choice

Suppose we have a set S of r independent issues, where each issue is a choice between two alternatives $\{0, 1\}$. Without loss of generality the alternative preferred by the voter majority is 1 for each issue. So the outcome from voting directly on every issue using majority rule would be represented by a string of r ones, $\{1\}^r$. We denote the preference of voter i on issue j as $\mathcal{V}(i, j) \in \{0, 1\}$, and similarly for candidates (before election) and representatives (after the election) as $\mathcal{C}(i, j) \in \{0, 1\}$ and $\mathcal{D}(i, j) \in \{0, 1\}$, respectively.

In a real-world representative voting system, we would not know the alternative preferred by the voter majority, and may not know with certainty the preferences of all electoral candidates on all issues before the election. Unfortunately, even if we have all of this extra information, electing a good set of representatives is still hard.

8.3.1 Complexity of Elections

To emulate direct voting using unweighted representative voting, one wants to elect a set of k representatives that maximizes the number of issues on which the majority of representatives agree with the voter majority by voting for 1. Unfortunately, given any fixed k, this problem is NP-Hard, as we show in Theorem 16.

If all voters use their option to delegate on all issues, then this directly emulates direct democracy, but only as long as there is at least one representative who agrees with the voter majority on each issue. Intuitively, it should be easier to find a set of representatives that maximizes the number of issues for which at least one representative agrees with the voter majority than to find a set where the majority of representatives agrees with the voter majority on the maximum number of issues. Unfortunately, this seemingly easier problem is also NP-Hard, as we show in Theorem 15.

Coverage A set of representatives *covers* an issue if there is at least one representative who votes for 1 on that issue. In other words, the only way for a set of representatives not to cover an issue would be for them to unanimously vote against the alternative preferred by the voter majority.

Theorem 15 (Max k-Coverage)**.** *If candidates' preferences are known for a set of binary issues, and the alternative preferred by the majority is known (1), selecting the subset of candidates that maximizes the number of issues covered is NP-Hard.*

Proof. We prove the hardness of our problem using the hardness of Max k-Cover [168], [169]. For the Max k-Cover problem we are given a set $S = \{s_1 \dots s_r\}$ of r elements, a collection $C = \{C_1 \dots C_m\}$ of subsets of U, and an integer k. The objective of Max k-Cover is to select k subsets from C such that their union has maximum cardinality.

Given an instance (S, C, k) of Max k-Cover we create an instance of our problem as follows. For every element $s_j \in S$ create an issue s_j and for every subset $c_l \in C$ create a

candidate l. For all elements s_j and subsets c_l, if $s_j \in C_l$ then let $\mathcal{C}(l, j) = 1$; otherwise let $\mathcal{C}(l, j) = 0$. And let k be the number of representatives we will elect.

There is a one-to-one correspondence between the number of issues covered by our k representatives and the cardinality of the corresponding subsets in the original Max k-Cover instance. Therefore, any set of k candidates that maximizes coverage corresponds exactly to a collection of k subsets in our Max k-Cover instance whose union has maximum cardinality. □

Majority Agreement

Theorem 16 (Max k-Agreement)**.** *If candidates' preferences are known for a set of binary issues, and the alternative preferred by the majority is known (1), selecting the subset of candidates that maximizes the number of issues on which the majority agrees with the voter majority (votes for 1) is NP-Hard.*

Proof. We prove the hardness of Max k-Agreement by reduction using our problem of Max k-Coverage. We begin by replicating the instance of Max k-Coverage and adding $r + 1$ issues to the original r issues so that the total number of binary issues is $2r + 1$. Then, we augment the candidate set with $k + 1$ additional candidates who must be included in any committee that maximizes agreement between the voter and representative majorities. The objective becomes to select the $2k + 1$ candidates which maximize agreement. The $k + 1$ additional candidates must be in the solution set for Max k-Agreement, and the remaining k candidates selected will correspond exactly to the k candidates in the solution to our original instance of Max k-Coverage.

Given an instance of Max k-Coverage with set $S = \{s_1, \ldots, s_r\}$ of binary issues, candidates $C = \{c_1 \ldots c_m\}$, and an integer k of the number of representatives to elect from C, we construct an instance of Max k-Agreement with set $\tilde{S} = \{\tilde{s}_1, \ldots, \tilde{s}_{\tilde{r}}\}$ of $\tilde{r}$ issues, candidates $\tilde{C} = \{\tilde{c}_1 \ldots \tilde{c}_{\tilde{m}}\}$, and $\tilde{k}$ as the number of representatives to elect from $\tilde{C}$. Create a set of binary issues $\tilde{S} = \{\tilde{s}_1, \ldots, \tilde{s}_{2r+1}\}$ and a set of candidates $\tilde{C} = \{\tilde{c}_1, \ldots, \tilde{c}_{m+k+1}\}$. Let $\tilde{k} = 2k + 1$. Construct the profile of candidate preferences as follows:

- If $l \leq m$ and $j \leq r$, then $\tilde{\mathcal{P}}_C(l, j) = \mathcal{C}(l, j)$.
- If $l \leq m$ and $r < j \leq 2r + 1$, then $\tilde{\mathcal{P}}_C(l, j) = 0$.

- If $m < l < m + k + 1$, then $\tilde{\mathcal{P}}_C(l, j) = 1$ for all issues $\tilde{s}_j$.
- If $l = m + k + 1$ and $j \leq r$, then $\tilde{\mathcal{P}}_C(l, j) = 0$.
- If $l = m + k + 1$ and $r < j \leq 2r + 1$, then $\tilde{\mathcal{P}}_C(l, j) = 1$.

The majority of representatives in $\tilde{D}$ agree with the voter majority on the remaining $r + 1$ issues $\{r + 1, \ldots, 2r + 1\}$ if and only if $\tilde{D}$ contains $\{\tilde{c}_{m+1}, \ldots, \tilde{c}_{m+k+1}\}$, because this is the only way at least $k+1$ out of the $2k+1$ representatives can agree with the voter majority on any of these issues. Therefore any set $\tilde{D} \subseteq \tilde{C}$ of $2k + 1$ representatives which maximizes majority agreement must contain $\{\tilde{c}_{m+1}, \ldots, \tilde{c}_{m+k+1}\}$. □

8.3.2 Approximation and Simulation

We investigate the properties of coverage and majority agreement as functions of the number of candidates (m), number of issues (r), and number of representatives (k).

We assume that voter preferences over candidates ($\mathcal{V}_\mathcal{C}$) are induced by how close each of the candidates' preferences are to the individual voters. Informally, a voter will prefer a candidate whose preferences are closer to theirs to one whose preferences are less similar. Here, the natural notion of similarity is the fraction of issues on which a voter and candidate agree. Hence, a voter will prefer a candidate who votes the same way as them on a greater number of issues. More formally, since each agent's preference can be represented by a binary string of length r, two agent's preferences are more similar if they have a smaller Hamming distance: $d(v_i, c_l) = \frac{1}{r} \sum_{s_j \in S} |\mathcal{V}(i, j) - \mathcal{C}(l, j)|$.

Approximating Agreement Following Chapter 5, one can say the voters and candidates as being embedded in a metric in which every point in the metric is a binary string and the Hamming distance is the distance function. Since Condorcet-consistent rules are known to provide low distortion in that setting, one might suspect that Condorcet-consistent election rules would be good at approximating Max k-Agreement. Unfortunately, this is not at all the case.

Suppose each voter reports a total order over the candidates based on the number of issues on which they agree, breaking ties privately (e.g. randomly).

Theorem 17. *No Condorcet-consistent election rule using induced total orderings to elect one or more representative can approximate* Max k-Agreement.

Proof. The proof follows by example. In this pathological example, built from the profile in [170], the voters will elect one of two candidates. There are eleven columns representing binary issues, the first eleven rows represent the preferences of eleven voters over these issues, and the last two rows are two candidates. Notice that these two candidates are the best and worst conceivable candidates having preferences $\{1\}^r$ and $\{0\}^r$, respectively.

Table 8.1: Pathological example for Condorcet-consistent elections.

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}
v_1	0	0	0	0	0	0	0	1	1	1	1
v_2	1	1	1	1	0	0	0	0	0	0	0
v_3	1	0	0	0	0	0	0	0	1	1	1
v_4	1	1	0	0	0	0	0	0	0	1	1
v_5	1	1	1	0	1	0	0	0	0	0	0
v_6	0	0	0	1	1	1	1	0	0	0	0
v_7	0	0	0	1	1	1	1	1	0	0	0
v_8	1	1	1	0	0	1	1	1	1	1	1
v_9	1	0	1	1	1	1	1	1	1	1	0
v_{10}	0	1	1	1	1	1	1	1	1	0	1
v_{11}	0	1	1	1	1	1	1	1	1	1	1
c_1	1	1	1	1	1	1	1	1	1	1	1
c_2	0	0	0	0	0	0	0	0	0	0	0

By duplicating agents in the example above, one can see that this holds for electing any number of representatives from some set of candidates. □

Seeing as no Condorcet-consistent rule can even approximate Max k-Agreement, and maximizing agreement is NP-Hard, we are moved to consider methods of reweighting the representatives.

Induced Preferences Over Candidates We consider three ways in which voters might report their preferences over candidates: approvals, total orders, and numerical weights. In all three cases, voters' preferences over candidates are assumed to be derived from, and consistent with, their relative distances $d(v_i, c_l) = \frac{1}{r} \sum_{s_j \in S} |\mathcal{V}(i,j) - \mathcal{C}(l,j)|$.

When voters report approvals, we assume that voter i approves of candidate l if and only if $d(v_i, c_l) < \frac{1}{2}$. When voters report total orders, we assume voter i ranks c_l above c_h if $d(v_i, c_l) < d(v_i, c_h)$, and voters break ties privately (e.g. randomly). Lastly, when voters report normalized weights, the weight that v_i givens to c_l is $\frac{1-d(v_j,c_l)}{\sum_{c_h \in C} 1-d(v_j,c_h)}$.

Comparing Election Rules by Simulation We compare several common, anonymous election rules F_C with a fixed, odd committee size k so that the set of elected representatives is $D \subseteq C$ where $|D| = k$. All rules considered are deterministic other than randomized tie-breaking. In the setting where voters submit approval ballots, we consider Approval Voting and Re-weighted Approval Voting (AV, RAV). When voters submit their preferences over candidates as total orderings, we consider Single-Transferrable Vote (STV), Borda, k-Median, and Chamberlin-Courant (CC). When voters submit their ballots as normalized weights over the candidates, we consider the rule which selects the k candidates who receive the largest total weight. Lastly, we compare these rules to sortition, selecting k representatives uniformly at random from the candidates. Formal definitions of AV, RAV, STV, and Borda can be found in [37, ch. 2, pp. 23–56] and definitions of CC and k-Median can be found in [157].

In all our simulations agent preferences are uniformly random, so for all issues s_j we let $\mathcal{V}(i,j) = 1$ (respectively, $\mathcal{C}(i,j) = 1$) with probability $\frac{1}{2}$ for all voters (candidates). All candidates and voters come from the same populations, i.e., they are both drawn from the same distribution. Relaxing this assumption is an interesting direction for future work. In all of our runs, all issues were covered by the elected representatives, hence we omit the consequent graphs. For all simulations we perform 50 iterations at each data point and plot the mean ($\sigma^2 \leq 0.002$).

Bear in mind that in practice the number of representatives on either side of the issue may be known, but it is not known which side corresponds to the voter majority. Therefore, whether the majority of representatives agrees with the voter majority for an issue may be unknown.

For Flexible Representative Voting, we allocated each voter one divisible vote for each independent issue, maintaining the principle of "one person, one vote". Once the representatives have been elected, each voter's unit of voting power is distributed among the representatives on each issue. In this paper we distribute this power uniformly by default, so initially $W_D(l) = {}^{n}/_{k}$ for each of the k candidates on each of the r issues. Various other distributions from the literature on voting power [171], [172] and Proxy Voting [106] are worth consideration in future work. We do not consider abstentions by representatives nor voter abstentions, whereby a voter assigns less than a full vote to the representatives as a whole. The total voting power held by the representatives remains n collectively for all issues.

We refer to the fraction of voters who use their delegation option as the *delegation*

rate α, and the delegations rates may differ for the majority and minority (α_1, α_0). For our purposes, the total weight $W_D(l)$ assigned to representative d_l on issue s_j is the sum of the voting power they receive from default and delegation. In this paper our decision rule is weighted majority with random tie breaking. That is, $y_j^* = 1$ if $W_D(1) > \frac{n}{2}$, $y_j^* = 0$ if $W_D(1) < \frac{n}{2}$, and $y_j^* = 1$ with probability $1/2$ if $W_D(1) = \frac{n}{2}$, for all independent issues $s_j \in S$.

A voter delegates incisively if they only delegate voting power to representatives who agree with their preferred outcome on an issue. This is equivalent to restricting voters to only delegate to a single representative, or to vote directly on the issue when the representatives are not unanimous. We relax the assumption of incisive delegations in our simulations and consider voters who delegate only to their most preferred candidate(s) or divide their delegation evenly across their approved subset of representatives. If the representatives are unanimous, the outcome is determined regardless of delegation since opposing voters have no one to whom they can delegate.

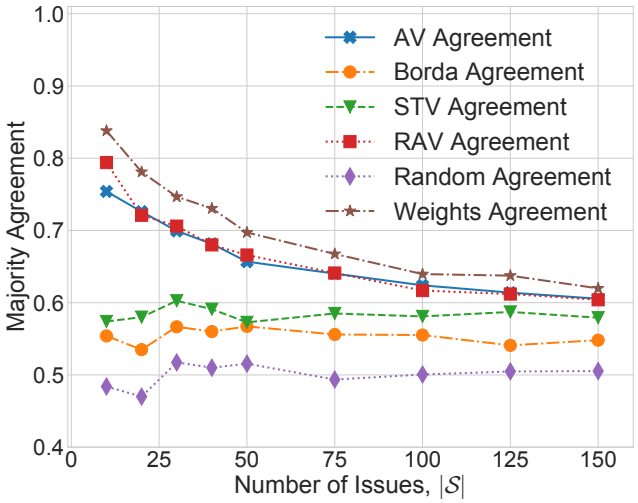

Figure 8.1: Varying number of issues with $k = 21, |C| = 60$.

Benefits of Flexibility Given our theoretical results on how delegation can improve outcomes, we investigate the effect of the overall delegation rate α on recovering the ideal outcomes according to direct voting. For our simulated delegations we create instances with $|V| = 301$, $|C| = 60$, $|S| = 150$, and $m = 21$. We vary $\alpha \in \{0, 1.0\}$ in increments of 0.01 and for each setting of α we run 50 iterations. We plot the means in Figure 8.5 and note again that the variance is ≤ 0.002 at every point.

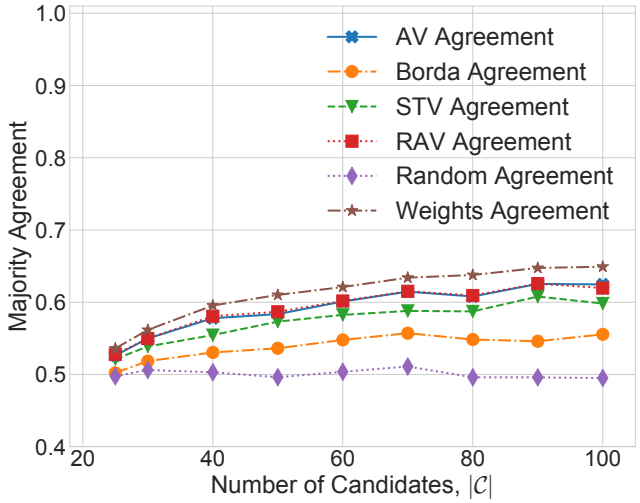

Figure 8.2: Varying number of candidates with $k = 21, |S| = 150$.

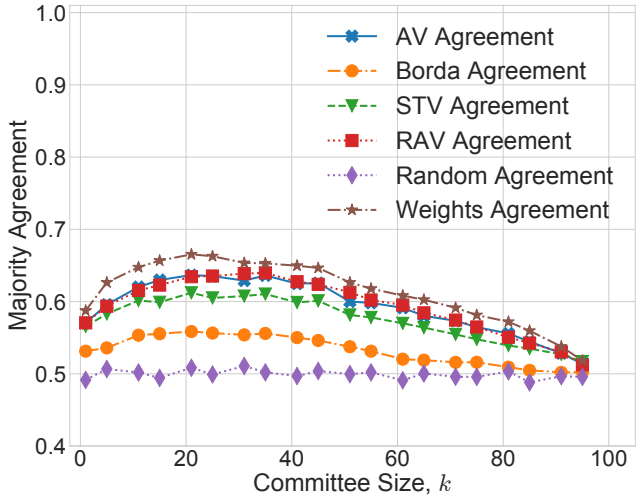

Figure 8.3: Varying the committee size with $|C| = 100, |S| = 150$.

In Figure 8.5 we can see the performance of and unweighted representative voting, weighted representative voting, and flexible representative voting for the weighted voting election rule for several delegation types and delegation rates. For each of the instances we measure the proportion of issues where the outcome is 1. A value of 1.0 means that the outcomes of all issues are the same as they would be under direct voting.

We compare four different delegation schemes: (1) *Approve* where voters delegate evenly across the representatives of whom they approve and do not update; *Best Rep* where voters delegate to their single most preferred representative across all issues and do not update per issue; *Best-3 Rep* where voters delegate equally to their three most preferred

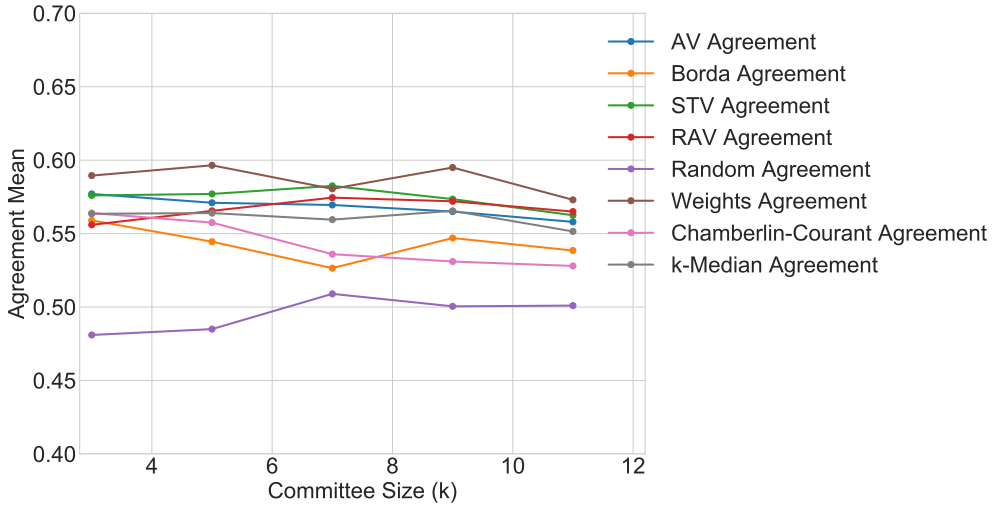

Figure 8.4: Comparison of NP-hard rules with our polynomial rules for $|C| = 17, |S| = 80, |V| = 51$. We cannot scale this graph further due to the high computational cost of computing the winning sets for $k-$Median and CC. However, from this small sample we see that Weights, STV, and AV all strictly dominate both CC and $k-$Median in terms of agreement.

representatives and do not update per issue; and finally *Optimal* where voters delegate to a single representative with whom they agree on each individual issue.

Most surprising is how little delegations help emulate direct voting when they are not on a per-issue basis, or not incisive. The *Approve* delegation system is perhaps closest to the proposal of Proxy Voting espoused by [102] but does not improve representative voting in a meaningful way. Similarly interesting are the *1-Best* and *3-Best* delegations, which also do not move the outcome towards the ideal of direct voting. Hence, we can see that the issue-specific flexibility our proposal allows can be effectively used to improve outcomes of decision making systems. Another striking result in Figure 8.5 is how drastically flexible representative voting can improve agreement over representative voting when voters are highly attentive. With as little as 60% of the population delegating we can improve agreement by 10%, and when the delegation rate reaches 80% we see an almost 20% increase, eventually reaching 100% when everyone delegates, if the representatives are not unanimously opposed to the voter majority.

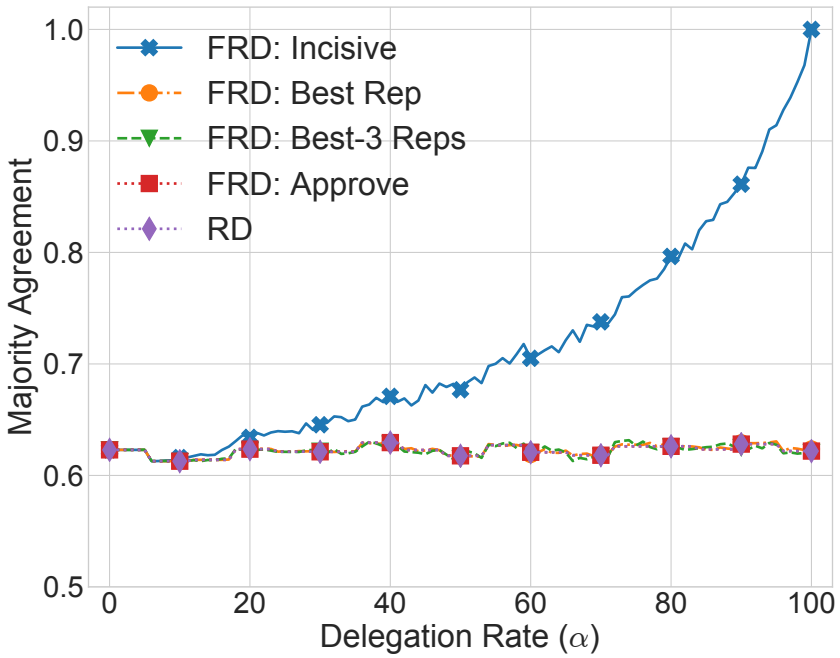

Figure 8.5: Weighted voting representative election.

PART III

AXIOMATIC FOUNDATIONS OF DECIDING HOW TO DECIDE

CHAPTER 9
CONTRIBUTIONS, BACKGROUND, AND RELATED WORK

How decisions are made matters. In the first two chapters of *Proposed Roads to Freedom*, Bertrand Russell compares the Anarchist Communists following Proudhon and the Socialists following Marx [11]. The two groups differed crucially in that the Marxist Socialists accepted concentrated power and majority rule while the Anarchists pursued the pure liberal ideal of unanimity, where everyone has a veto. Unfortunately, as Russell noted, unanimous consent may be highly desirable, but it makes the organization of large groups impossible. At the same time, majoritarianism and concentrated power can enable an oppressive society. One approach at compromise is to use supermajority rules that lie somewhere between unanimity and simple majority, but this middling approach does not satisfy those who believe unanimity should be required or be seen as cumbersome to others.

The intractable nature of achieving unanimous consent at scale combined with adherence to core liberal values has led many scholars to turn to markets as the solution. Market transactions are made only with the consent of the parties involved, and no one is forced to participate in the market. Markets only require the enforcement of contracts that parties enter into knowingly, rather than an implicit social contract in the style of Rousseau [173]. And markets can be decentralized. The allure of markets has generated attempts to turn voting systems into marketplaces by means of vote buying [174], including notable proposals using quadratic voting [175]–[179]. Market-based thinking has influenced Social Choice as well as Mechanism Design for Social Good [180], [181]. However, markets are not appropriate for all collective decisions. Markets may be coercive, concentrate power plutocratically,

Portions of this chapter appeared as: B. Abramowitz, E. Shapiro, and N. Talmon, "How to amend a constitution? Model, axioms, and supermajority rules: (extended abstract)," in *Proc. of the 20th Int. Conf. Autonomous Agents and Multiagent Syst.*, May. 2021, pp. 1443–1445.

Portions of this chapter previously appeared as: B. Abramowitz, E. Shapiro, and N. Talmon, "How to amend a constitution? Model, axioms, and supermajority rules," Mar. 2021, arXiv:2011.03111 (Retrieved on 7/18/2021).

Portions of this chapter are to appear as: B. Abramowitz, E. Shapiro, and N. Talmon, "In the beginning there were n agents: Founding and amending a constitution," in *7th Int. Conf. on Algorithmic Decision Theory.*, Nov. 2021.

produce negative externalities, and can change the nature of decisions [182].

The field of Mechanism Design studies mechanisms without monetary payments [183], [184, ch. 10, pp. 243–266], but generally maintains many of the assumptions of those market models. Consider the sociological gap between human behavior, particularly voting behavior, and the market models in which individuals are assumed to be rationally self-interested utility maximizers. This type of model is frequently adopted in the fields of Game Theory, Public Choice [185] and Computational Social Choice [37], [186], [187]. However, people are not strictly self-interested utility maximizers. Voters in large democracies can't be utility maximizers or else no one would vote, since the probability of being pivotal in large populations is effectively zero [188], [189]. While some scholars claim observed voting behavior is due to the irrationality or bounded rationality of voters, rationality is not predicated on self-interest, utility maximization, or even instrumental reasoning [10].

The properties of preference aggregation procedures (functions) can be important beyond what outcomes (output) they produce. The properties of decision making procedures are generally described by *axioms*. For example, one may put value on using a *non-dictatorial* decision making process non-dictatorial where no single agent can enforce the outcome they prefer without regard for the preferences of others. Some normative principles surrounding preference aggregation are defined by whether certain properties of a preference profile (input) are reflected in the outcome (output). For example, if all agents prefer alternative A over alternative B, then B should not be ranked above A in the social order when the outcome is a ranking over alternatives. The way in which a decision is made, or the axioms a preference aggregation procedure satisfies, can impact on whether agents view a decision as legitimate or justified.

The highly productive axiomatic or Arrovian approach to Social Choice was ignited by Arrow [19], [37]. On one side of the ledger there are possibility results that characterize concrete preference aggregation procedures by the axioms they satisfy. A simple example is May's Theorem, which shows that majority rule is the unique rule that satisfies a particular set of desirable axioms [162]. One the other side of the ledger are impossibility theorems. Arrow's General Impossibility Theorem states that no *social welfare function* can satisfy a small set of desirable axioms at once for all preference profiles [19], [190], [191, ch. 8, pp. 80–87]. Hence, if we want to use a social welfare function as our preference aggregation procedure, we must choose at least one of those axioms to forgo. There are a number of

similar impossibility theorems that create dilemmas when deciding how to decide [6]. Above all, the literature on Social Choice shows us that when it comes to preference aggregation there is no perfect mechanism and we sometimes have to make sacrifices when deciding what normative principles to uphold.

Mechanism designers must make choices about what matters. On some level this is unavoidable. However, just as an agent who single-handedly determines the outcome is a dictator, an agent who single-handedly determines what preference aggregation procedure is used is dictator of norms. It has been suggested that mechanism designers should act principally in an engineering capacity, taking a position of "informed neutrality" between ethical positions, and leave the primary responsibility of deciding how to decide to philosophers and policymakers [17]. But if mechanism designers, or whoever is assumed to be implementing such mechanisms, care about what axioms a procedure should satisfy, then it is odd to always assume that the agents we model do not have the same concerns. In other words, there must be a way to aggregate agent preferences over how their preference should be aggregated, particularly for any process intended to be *democratic*.

Mechanism design is different from economic analysis [192], [193]. When taking the position of designer, on some level there must be a dictator. One cannot vote on the voting rule for voting on the voting rule for voting on the voting rule *ad infinitum* as this leads to a problem of infinite regress [194]. However, instead of dictating outcomes or dictating rules for aggregating agent preferences over outcomes, one can take just one more step back and dictate the means by which agent preferences over rules and axioms lead to the selection of aggregation rules [195], [196].

While one might think that preferences over axioms and rules are relevant only in certain circumstances, this is misleading. The only laws that cannot be changed are the laws of physics. For all decision making procedures, there is always some set of agents with the power to change the rules, even when a change seems unlikely. Whether there is no formal procedure in place (e.g. a group of friends deciding where to go for dinner), or whether there is a formal procedure for changing the rules (e.g. constitutional amendments), the rules of the game are always in play.

9.1 Founding and Amending a Constitution

Kenneth May's seminal paper [162] demonstrates that majority rule is the only voting rule that satisfies a small, intuitive set of axioms. We have taken May's Theorem as inspiration to answer the question of how a group of agents could establish a set of rules for themselves where none existed before. While our axioms lead the agents to use majority rule initially when founding their constitution, they also enable the agents to change this rule according to their collective preferences. While there is a large literature on characterizing social choice rules axiomatically (see, e.g., [6], [197]), we believe that our work is the first to introduce a set of axioms that simultaneously lead to the formulation of an initial rule and the process by which it can amend itself. We consider two alternative axioms for how the agents can determine what proposal should win if multiple incompatible proposals are viable, namely Condorcet-consistency and Conservatism. Extensions of the Condorcet principle have a long history in social choice [198], [199], and the principle of Conservatism comes from reality-aware parameter update rules [200].

Voters can vote directly over what axioms they think a voting rule should satisfy [201], [202]. Alternatively, rather than thinking in terms of properties or axioms directly, one might have preferences over procedures or rules themselves. For example, members of a parliament or congress might decide to use majority rule for some decisions and a supermajority rule for others, depending on the nature of the decision. Just as one might characterize what outcome is "best" one can debate what rule is "best" to use. Again, just as with characterizing outcomes, one can consider the axioms a rule satisfies or its *stability* with respect to a particular profile and set of alternate rules.

Much of the work on constitutional amendments focuses on some notion of stability or self-selectivity [203]–[206]. Here, again, impossibilities arise. In particular, what we might call Houy's Theorem shows that some assumptions must be made about the structure of agent preferences or the set of alternative voting rules in order for a unique rule to be the obvious winner [207].

One line of research on constitutional amendments assumes that voter preferences over possible rules are derived from their preferences over the outcomes that each rule produces [203], [208]. This is the "consequentialist approach." Perhaps the work here most similar to ours is that of Bhattacharya [209] as they consider single-peaked preferences. We take a different approach to modeling preferences over constitutional amendments, more

similar to that of Barbera and Jackson [210]. In their model, a constitution is just a set of voting rules. A constitution may be a single rule used to update itself, or it can consist of a fixed amendment rule used to update the decision rule. In contrast to Barbera and Jackson, who conclude that a rule used to update itself is too simple for their setting, we see this simplicity as a virtue in bootstrapping a constitution.

Contributions Chapter 10 demonstrates that if a set of agents agrees upon a rudimentary set of axioms, they can not only arrive at an initial preference aggregation procedure to use, but also the means by which the procedure can change itself according to the preferences of the agents over the set of procedures that satisfy those axioms. In other words, these axioms constitute a rudimentary constitution. When agent preferences over rules are fixed, the iterative procedure described in Chapter 10 ultimately settles at a unique rule.

9.2 Political Social Choice

Since preference aggregation procedures can be characterized by the axioms they do and do not satisfy, such axioms can be used to generate explanations and justifications for decisions [211]–[216]. Consequently, the generation of explanations and justifications is receiving increased attention in the area of Computational Social Choice [217]. According to [212], a justification consists of an explanation and normative principles. The explanation for why a particular outcome is selected follows from the application of axioms to concrete instances.

Explanations and justifications are particularly important for comparing rules and procedures after a decision has been made or during deliberation. However, our focus on preference aggregation is much narrower than the broader study of abstract argumentation [218]–[220] and is not encapsulated by judgment aggregation [221]–[223]. Chapter 11 also does not consider persuasion [224], and differs greatly from recently proposed spatial models of deliberation [96].

When coming up with explanations or justifications for decisions, it is important to consider what information is available. In particular, if one can only observe the agent preferences (input) and outcome (output), one cannot necessarily say for certain what preference aggregation procedure (function) was used in the process. If one only cares about the outcome, or whether some feature of the profile is reflected in the outcome (e.g. Weak Pareto

axiom), then it doesn't matter. However, problems arise if one cares about other properties of the preference aggregation procedure.

Recall the axiom of *non-dictatoriality* (used in Arrow's Theorem) which requires that no single agent determine the outcome single-handedly, regardless of the preferences of others. A similar axiom is that of *anonymity* (used in May's Theorem), which requires that all agents be treated identically. A preference aggregation procedure is anonymous if reassigning all of the agents to different preferences such that the profile as a whole remained the same except for relabeling does not change the outcome. If one can only observe agent preferences and the outcome, on cannot generally determine whether the function used to produce the outcome satisfied either anonymity or non-dictatoriality.

Preferences for non-dictatoriality and anonymity are limiting cases of preferences over the *a priori* distribution of power. Numeric indices of power have been developed and studied in Game Theory and Social Choice [122], [171], [172], [225]–[227]. Measures of power receive special mention here because (1) preferences over the distribution of power are commonly observed, including in other species [10], and (2) indices of power are a prime example of the counterfactual reasoning behind inter-profile comparisons.

Given a single instance of a preference profile and outcome, determining whether preference aggregation procedure used satisfies the axioms of non-dictatoriality and anonymity, we need to reason about what would have happened if the preference profile had been different. This type of counterfactual reasoning is necessary for measures of power and many other axioms, and it is central to reasoning about causality [228]–[230].

Since justifications for any choice are derived from axioms, if those axioms require counterfactual reasoning and different agents can use different counterfactual reasoning given the same facts, then different agents may be satisfied by separate justifications for the same choice even when no single justification can satisfy all of them. This is the core lesson of Chapter 11.

Contributions Chapter 11 creates a framework for making decisions when agents prefer different outcomes and disagree about how the decision should be made. This approach is then applied to Arrow's axioms and constitutional amendments to demonstrate that decisions can be made that satisfy everyone, even when agents do not all agree on a common set of axioms or a single most preferred outcome, and no single justification can satisfy everyone.

CHAPTER 10
FOUNDING AND AMENDING A CONSTITUTION

Consider a group of n agents that gather to form a self-governed community. To make collective decisions, they first need to agree on a rule by which to do so; but offering to vote on the voting rule leads to a problem of infinite regress.

Thus, in this paper we offer a different approach. In particular, we assume that the agents first agree on a common set of axioms, and aim to infer an initial decision rule from these axioms. Ideally, the axioms will logically imply an identifiable set of permissible decision rules, and a unique decision rule from this set to serve as the initial rule, so that an agent who agrees to the axioms cannot object to the initial rule they imply, nor to the set of permissible rules.

Furthermore, it is crucial that the decision rule be amendable. Once the agents have passed the first hurdle and founded their rudimentary constitution, agent preferences over the various permissible decision rules become material. Thus, we assume that the agents also agree on a common set of axioms that relate to the amendment procedure to change the decision rule itself. Again, ideally, the axioms will logically imply a unique amendment procedure such that, taking agent preferences over possible decision rules into account, agents who agree to the axioms would have no objections to the amendment procedure.

We view the initial decision rule as the foundation of the *constitution* of the community and the amendment procedure as the process by which the constitution changes itself. We look at our work as a step towards understanding how a group of agents can congeal into a self-governed, rule-based, egalitarian community with a decision making mechanism that is not dictated by an external agent.

10.1 Founding and Amending a Constitution, Informally

We begin by informally discussing the processes of founding and amending a constitution; later, we delve into a rigorous mathematical treatment.

This chapter is to appear as: B. Abramowitz, E. Shapiro, and N. Talmon, "In the beginning there were n agents: Founding and amending a constitution," in *7th Int. Conf. on Algorithmic Decision Theory.*, Nov. 2021.

10.1.1 Founding a Constitution, Informally

We assume that all n agents agree that they need a procedure for accepting or rejecting proposals. Proposals are considered individually, and for each proposal, each agent either approves it or not. The agents wish to establish a decision rule that determines whether the group as a whole accepts or rejects a given proposal, based on the votes of its members. To achieve that, they first agree on the following axioms that a decision rule must fulfill.

We first introduce the axioms without mathematical precision, to demonstrate how unassuming they are, and show that the claims that follow from them make intuitive sense. In the next Section 10.2 we rephrase the axioms mathematically, rewrite the claims as propositions, and prove the propositions formally.

Axioms for a Decision Rule

1. **Decisiveness**: For a given proposal, the decision is either to accept or to reject it, as determined by the votes.

2. **Monotonicity**: A decision to accept a proposal will remain so if some agents that do not approve the proposal change their vote to approve.

3. **Anonymity**: A decision to accept a proposal shall not depend on the identities of the agents approving it.

4. **Concordance**: If two proposals are incompatible, in the sense that no agent may approve both, then not both can be accepted.

5. **Minimality**: Among the decision rules consistent with the axioms above, choose an initial rule that requires a minimal number of approvals.

Axioms 1-3 imply that the decision rule is based on the size of the fraction of the agents approving the proposal.

Informal Claim 1 (Fractional Approval Voting)**.** Axioms 1-3 imply that the decision rule shall accept a proposal if approved by at least some fraction of the agents.

Adding Axiom 4 implies that the fraction shall be at least a half.

Informal Claim 2 (Supermajority Approval)**.** Axioms 1-4 imply that the fraction of approvals needed to accept a proposal is greater than $\frac{1}{2}$.

We refer to a fraction of approvals greater than $\delta \geq \frac{1}{2}$ as a δ-*supermajority*, with the case of $\delta = \frac{1}{2}$ being *simple majority* and $\delta = 1 - \epsilon$, where $0 < \epsilon \leq \frac{1}{n}$, being *unanimous consent*.

Note that, unlike Axioms 1-4, which effectively restrict the set of possible decision rules, Axiom 5 compares and prioritizes some decision rules over others.

Informal Claim 3 (Initially, Simple Majority Approval)**.** Axioms 1-5 imply that the initial decision rule is approval via a simple majority.

To summarize, we claim that, according to Axioms 1-5, the decision rule must always be a δ-supermajority approval rule, $\frac{1}{2} \leq \delta < 1$, with the initial supermajority being a simple majority $\delta = \frac{1}{2}$.

10.1.2 Amending a Constitution, Informally

The initial decision rule of the community is simple majority according to Axiom 5, but the axioms are consistent with amending it to any δ-supermajority, with $\frac{1}{2} \leq \delta < 1$. So, once the initial rule is established as simple majority, some agents may wish to amend it to a higher supermajority rule. Some agents may aspire for a decision by unanimous consent, while others may prefer to require more than simple majority but less than unanimous consent, for example to protect minorities. Agents may also wish to employ different rules for different decisions in the future. Furthermore, agents' preferences may change over time. Thus, below we focus on the fundamental question of how the decision rule can be amended.

Naturally, the axioms agreed upon by the community apply to the decision of whether to accept a newly-proposed decision rule just like they would for any other decision. Hence, the decision to change the present δ-supermajority decision rule must be approved by a δ-supermajority. However, proposals to amend the decision rule itself require additional considerations, reflected in the consistency requirements introduced below. But first we provide some background in order to express them.

We assume that agents have preferences over decision rules, and that an agent approves a proposal to amend the decision rule if and only if the agent prefers the newly-proposed rule over the decision rule in force. Concretely, we identify agents with their index $i \in [n]$, and make the simplifying assumption that agent preferences are single-peaked in the following sense: Every agent $i \in [n]$ has an *ideal point* δ_i, $\frac{1}{2} \leq \delta_i < 1$, such that: (i) agent i strictly

prefers δ_i over all other proposed δ's; (ii) among any two proposed values larger than δ_i, the agent prefers the smaller of the two; (iii) among any two proposed values smaller than δ_i, the agent prefers the larger of the two; and (iv) the agent i has no other pairwise preferences [231].

Secondly, we say that a proposal p is *preferred* over another proposal (aka *dominates*) p' if the set of agents that prefer p over p' are a majority. A proposal p is *most-preferred* if no other proposal p' is preferred over p. Note that, if p is the only most-preferred proposal, then it is a Condorcet winner [37]. Also, as preference is transitive and there are at most $\frac{n}{2}$ significant values of δ ($\frac{1}{2}, \frac{1}{2} + \frac{1}{n}, \ldots, \frac{n-1}{n}$), a knock-out tournament among these values for δ, choosing the preferred value in each match, may quickly result in a most-preferred value for δ.

Using these terms, we express two amendment axioms. The first axiom, named Posterior Consistency, says that, to change the current decision rule to a new decision rule, this amendment decision must be justifiable in retrospect according to the new rule, as opposed to being consistent only with current rule in force. This is in essence an "anti-hypocrisy" axiom. The second axiom requires that among all alternative decision rules consistent with the other axioms, we choose a most-preferred decision rule.

Axioms for a Decision Rule Applied to Amend Itself

6. **Posterior Consistency**: A proposal to amend the decision rule is accepted only if accepted according to the newly accepted decision rule as well.

7. **Condorcet Consistency**: Among the decision rules that satisfy all axioms above, choose an most-preferred one.

Given our assumptions on agent preferences and Axioms 1-5, we claim that there is a unique amendment process that is also consistent with Axioms 1-7

Informal Claim 4 (Condorcet Amendment Rule)**.** Axioms 1-7 together with our assumptions on agent preferences imply that the process to amend the current δ-supermajority approval decision rule must be:

1. Increase δ to $\delta' > \delta$, if and only if there is a δ'-supermajority that prefers δ' over δ and δ' is the **maximal** supermajority with this property;

2. Decrease δ to $\delta' < \delta$, if and only if there is a δ-supermajority that prefers δ' over δ and δ' is the **minimal** supermajority with this property; and

3. Retain the present δ-supermajority rule otherwise.

This completes the informal presentation.

10.2 Founding and Amending a Constitution, Formally

In this section we rephrase the assumptions, axioms, and claims formally.

10.2.1 Founding a Constitution, Formally

Let $B = \{0, 1\}$; these values represent approving or disapproving a given proposal by each agent, as well as accepting or rejecting the proposal by the group of agents as a whole. For the set of n agents, we refer to $V \in B^n$ as a voter *profile*. For a profile V, we define $[\![V]\!] := |\{i \mid V_i = 1\}|$ to be the number of approvals in V. For two profiles $V, V' \in B^n$, define $V \leq V'$ if $V_i \leq V'_i$ for all $1 \leq i \leq n$.

Indeed, we assume a set of n agents that wish to agree on a decision rule F that takes a proposal and a profile V, where each voter specifies either *approve* (1) or *not approve* (0); and produces (i.e., outputs) a decision of either to *accept* (1) or to *reject* (0). In the axioms below, $V, V' \in B^n$ and the given proposal is assumed and not specified as an explicit parameter.

Formal Axioms for a Decision Rule

9. **Decisiveness**: The decision rule F is a function $F : B^n \to B$.

10. **Monotonicity**: If $F(V) = 1$ and $V \leq V'$ then $F(V') = 1$.

11. **Anonymity**: If V' is a permutation of V then $F(V) = F(V')$.

12. **Concordance**: If $V_i + V'_i \leq 1$ for all $1 \leq i \leq n$, then $F(V) + F(V') \leq 1$.

13. **Minimality**: If F and F' are consistent with the axioms above, and $F'(V') \leq F(V)$ implies $[\![V']\!] \leq [\![V]\!]$ for all $V, V' \in B^n$, but not vice versa (exchanging F and F'), then prefer F' over F.

We now rephrase Informal Claims 1-3 into propositions, and prove them.

Proposition 6 (Fractional Approval Voting)**.** *Axioms 9-11 imply that there is some fraction* $0 \leq f < 1$ *such that* $F(V) = 1$ *if and only if* $[\![V]\!]/n > f$*.*

Proof. Let $V^* \in B^n$ be a voter profile $V^* = \underset{\substack{V \in B^n \\ F(V)=1}}{\arg\min} [\![V]\!]$, and let $f = \frac{[\![V^*]\!]-1}{n}$. Naturally, for any profile $[\![V']\!] < [\![V^*]\!]$, $F(V') = 0$ by the definition of V^*. Let V' be a profile for which $[\![V^*]\!] \leq [\![V']\!]$ and hence $f < [\![V^*]\!]/n \leq [\![V']\!]/n$. We wish to show that $F(V') = 1$. Consider a voter profile $V'' \in B^n$ such that $V'' \leq V'$ and $[\![V'']\!] = [\![V^*]\!]$. By anonymity, $F(V'') = F(V^*)$ and by monotonicity, $F(V'') \geq F(V') = 1$, assuming of course that F is a function $F : B^n \to B$ (Axiom 9). □

With n voters, a decision rule F using fraction $f = \frac{k}{n}$ for $k \in \mathbb{N}$ will return the same outcome as any decision rule F' using fraction $\frac{k}{n} + \epsilon$ for $\epsilon < \frac{1}{n}$, for all possible profiles. We therefore do not differentiate between these functionally equivalent rules. For example, the decision rule using fraction $f = \frac{1}{2}$ is majority rule whether the number of voters is even or odd.

Proposition 7 (Supermajority Approval)**.** *If* $F(V) = 1$ *then* $[\![V]\!] > \frac{n}{2}$*.*

Proof. By way of contradiction, assume $F(V) = 1$ and $[\![V]\!] \leq \frac{n}{2}$. Then V has a non-overlapping permutation V' such that $V_i + V'_i \leq 1$ for all $i \in [n]$. By Axiom 11, $F(V') = F(V)$, implying via the assumption that $F(V) + F(V') = 2$, contradicting Axiom 12. □

Proposition 8 (Initially, Simple Majority Approval)**.** *Axioms 1-5 imply that the initial decision rule is* $F(V) = 1$ *if* $[\![V]\!]/n > \frac{1}{2}$*.*

Proof. Let F be the initial decision rule. By Proposition 6 it is fractional, with some fraction $0 \leq f < 1$. By Proposition 7, $f \geq \frac{1}{2}$. We claim that $f \leq \frac{1}{2}$. By way of contradiction, assume $f = \frac{1}{2} + \frac{1}{n}$, and consider a decision rule F' with a fraction $f' = \frac{1}{2}$. Just like F, the rule F' is consistent with all the Axioms 9-13. Furthermore, it satisfies that $F'(V') \leq F(V)$ implies $[\![V']\!] \leq [\![V]\!]$, as $f < f'$. However, the converse is not true. Assume $[\![V]\!] = \frac{1}{2} + \frac{1}{n}$ and $[\![V']\!] = \frac{1}{2}$. Then $F(V) \leq F'(V')$ as both equal 0, but $[\![V]\!] > [\![V']\!]$, contradicting Axiom 13. Hence the initial decision rule is fractional with $f = \frac{1}{2}$, namely approval by a simple majority. □

Let $\Delta := [\frac{1}{2}, 1)$. To summarize, Propositions 6-8 prove that, according to Axioms 9-13, the decision rule must always be a δ-supermajority approval rule, $\delta \in \Delta$, with the initial supermajority being a simple majority $\delta = \frac{1}{2}$. Next, we rewrite Axioms 6-7(a, b) formally, recast Claims 4-6 as proper theorems, and prove them.

10.2.2 Amending a Constitution, Formally

We have established that all decision rules are δ-supermajority rules for some $\delta \in \Delta$. We therefore identify each decision rule F with its δ, and denote it by F^δ.

Agent preferences. Recall that each agent i has a preference $\preceq_i \subseteq \Delta \times \Delta$ over the decision rules, where $x \prec_i y$ if $x \preceq_i y$ but not vice versa. Furthermore, agent preferences are single-peaked in that every agent $i \in [n]$ has an *ideal point* $\delta_i \in \Delta$, such that: (i) $\delta \prec_i \delta_i$ for all $\delta \neq \delta_i \in \Delta$; (ii) if $\delta_i < \delta < \delta'$ then $\delta' \prec_i \delta$; (iii) if $\delta_i > \delta > \delta'$ then $\delta' \prec_i \delta$; and (iv) $\prec_i$ is the smallest relation satisfying (i)-(iii).

Hence, a voter approves a proposals if and only if it is in between the status quo and her ideal point, and the voter profile $V^{\delta,\delta'}$ on the amendment decision $F^\delta(\delta')$ for $\delta \neq \delta'$ is defined accordingly by $V_i^{\delta,\delta'} = 1$ iff $\delta_i \leq \delta' < \delta$ or $\delta_i \geq \delta' > \delta$.

We can now specify formally Posterior Consistency:

14. **Posterior Consistency**: Given an ideal points profile δ_i, $i \in [n]$, and $\delta, \delta' \in \Delta$, $F^\delta(V^{\delta,\delta'}) = 1$ only if $F^{\delta'}(V^{\delta,\delta'}) = 1$

According to Axiom 15, the amendment process is not a simple approval vote on a proposed amendment, but a selection of an most-preferred decision rule based on the reported ideal points of the agents. Hence, the decision on an amendment is of a different type, $\Delta \times \Delta^n \to \Delta$, which takes the δ in force and the ideal points profile δ_i, $i \in [n]$, as input, and produces the amended δ' as output.

Recall that a proposal p is *preferred* over (aka *dominates*) another proposal p' if $|\{i \in [n] : p' \prec_i p\}| > \frac{n}{2}$. A proposal p is *most-preferred* (aka *undominated*) if no other proposal p' is preferred over p. With this we can specify Condorcet Consistency:

15. **Condorcet Consistency**: Given an ideal points profile δ_i, $i \in [n]$, choose a most-preferred[2] $\delta' \in \Delta$ for which $F^\delta(V^{\delta,\delta'}) = 1$.

In other words, given the current δ-supermajority approval rule, Condorcet Consistency requires the choice of an amendment δ' that maximizes $[\![V^{\delta,\delta'}]\!]$.

We can now state our main theorem.

[2]The standard definition of Condorcet Consistency is the selection of the unique Condorcet winner (namely undominated alternative) when it exists. Our definition is a bit more general.

Theorem 18 (Condorcet Amendment)**.** *Given ideal points profile δ_i for $i \in [n]$, Axioms 9-15 and our assumptions on agent preferences imply that F^{δ} should be amended to $F^{\delta'}$ for $\delta' \neq \delta$ if either:*

1. $\delta' = \arg\max_{\delta < x < 1} |\{i \in [n] : \delta_i \geq x\}| > xn$ *exists, or*
2. $\delta' = \arg\min_{1/2 \leq x < \delta} |\{i \in [n] : \delta_i \leq x\}| > \delta n$ *exists*

Otherwise, no proposal to amend the decision rule is accepted.

Remark 1. Note that the two definitions of δ' could be made more similar textually by replacing the right-hand side of the equations with $> max(x, \delta) \cdot n$.

Proof. First, recall that, in all cases, the result is a δ-supermajority rule, which is consistent with Axioms 9-13. Next we argue that the δ' computed by the rules above is uniquely consistent with Axioms 9-15:

1. If such a $\delta' > \delta$ exists, then F^{δ} would accept a proposal to amend δ to δ' since δ' is preferable to δ by a δ'-supermajority, which is more than a δ-supermajority. Furthermore, it is Posterior Consistent as the proposal to amend δ to δ' will be approved by $F^{\delta'}$ for the same reason. Now we argue that δ' uniquely satisfies the axioms. A $\delta'' > \delta'$ cannot be chosen without violating Posterior Consistency, since there is no δ''-supermajority approval for that amendment by the maximality of δ'. A $\delta'' < \delta'$ cannot be chosen without violating Condorcet consistency, since δ' is preferred over δ'' by a δ'-supermajority and hence by a majority.

2. If such a $\delta' < \delta$ exists, note that accepting $F^{\delta'}$ as the new decision rule satisfies F^{δ} by the definition of δ', namely the use of δn on the right hand side. Furthermore, since δ' is preferable to δ by a δ-supermajority, which is greater than a δ'-supermajority, it is also Posterior Consistent. We now argue that δ uniquely satisfies the axioms. A $\delta'' < \delta'$ cannot be accepted by F^{δ}, since there is no δ-supermajority approval for amending δ to δ'', by the minimality of δ'. And a $\delta'' > \delta'$ cannot be chosen without violating Condorcet consistency, since δ' is preferred over δ'' by a δ'-supermajority and hence by a majority.

3. If there is no δ-supermajority to increase or decrease δ then of course δ should not change. No amendment proposal could be accepted according to δ itself, so remaining at δ in the face of any proposal is Posterior Consistent and trivially Condorcet-consistent.

This finishes the proof of Theorem 18. □

We have shown that a group of n agents that agree on a simple and self-evident axioms and a minimal set of assumption, can obtain from these a definite decision rule that also determines its own amendment process, via formal logical deduction.

10.3 Alternate Approaches: Conservative Amendment

Naturally, there are many possible alternatives to the axiomatic basis proposed herein, as well as to the assumptions made. Here we explore an alternative to Axiom 7.

7. **a. Conservatism**: Among the decision rules that satisfy all axioms above, choose the one closest to the current decision rule.

Given our assumptions on agent preferences and Axioms 1-5, we also claim that there is a unique amendment process that is consistent with Axioms 1-7(a).

In practice, Condorcet Amendment happens to select the furthest δ-supermajority rule from the current rule that satisfies Axioms 1-6. By contrast, Conservative Amendment requires that the rule selected be as close to the current rule as possible.

Informal Claim 5 (Conservative Amendment)**.** Axioms 1-7(a) together with our assumptions on agent preferences imply that the process to amend the current δ-supermajority approval decision rule must be:

1. Increase δ to $\delta' > \delta$, if and only if there is a δ'-supermajority that prefers δ' over δ and δ' is the **minimal** supermajority with this property;

2. Decrease δ to $\delta' < \delta$, if and only if there is a δ-supermajority that prefers δ' over δ and δ' is the **maximal** supermajority with this property; and

3. Retain the present δ-supermajority rule otherwise.

While the Condorcet and Conservative Amendment rules seem rather different, they are related in the sense that the second leads to the same result of the first if applied iteratively, in an exhaustive way – in case the rule in force is simple majority.

Informal Claim 6 (Iterate from Simple Majority)**.** For any given ideal points profile, if the rule in force is $\delta = \frac{1}{2}$ then the iterative application of the Conservative Amendment rule until no further amendments occur halts at the decision rule characterized by Informal Claim 4.

To recap the informal presentation, if agents found their constitution according to Axioms 1-5, their initial decision rule will be simple majority. If they amend their initial rule by applying iteratively either of two amendment rules that satisfy posterior consistency, with a given set of ideal points, the result would be the δ characterized by Informal Claim 4.

We now proceed with the formal analysis, considering Conservatism instead of Condorcet Consistency. First, we offer a formal statement of Conservatism:

15. **a. Conservatism**: Given an ideal points profile δ_i, $i \in [n]$, choose a $\delta' \neq \delta \in \Delta$ closest to δ for which $F^{\delta}(V^{\delta,\delta'}) = 1$.

The resulting Theorem 19 looks very similar to Theorem 18, but the two cases swap the minimization and maximization over values of δ. The proof is similar as well.

Theorem 19 (Conservative Amendment)**.** *Given ideal points δ_i for $i \in [n]$, Axioms 9-15(a) imply that F^{δ} should be amended to $F^{\delta'}$ for $\delta' \neq \delta$ if either:*

1. $\delta' = \underset{\delta<x<1}{\arg\min} \left|\{i \in [n] : \delta_i \geq x\}\right| > xn$ *exists, or*
2. $\delta' = \underset{1/2\leq x<\delta}{\arg\max} \left|\{i \in [n] : \delta_i \leq x\}\right| > \delta n$ *exists*

Otherwise, no proposal to amend the decision rule is accepted.

Proof. First, recall that in all cases the result is a δ-supermajority rule, which is consistent with Axioms 9-13. Next we argue that the δ' computed by the rules above is uniquely consistent with Axioms 9-15(a):

1. If such a $\delta' > \delta$ exists, then F^{δ} would accept a proposal to amend δ to δ' since δ' is preferable to δ by a δ'-supermajority, which is more than a δ-supermajority. Furthermore, it is Posterior Consistent as the proposal to amend δ to δ' will be approved

by $F^{\delta'}$ for the same reason. We now show that δ' uniquely satisfies the axioms. Any $\delta'' > \delta'$ cannot be chosen without violating Conservatism because δ' would be accepted. No $\delta'' < \delta'$ could be accepted, because this would imply that δ'' is preferred by a δ''-supermajority according to Posterior Consistency, which violates the definition of δ'.

2. If such a $\delta' < \delta$ exists, note that accepting $F^{\delta'}$ as the new decision rule satisfies F^{δ} from the definition of δ', namely the use of δn on the right hand side. Furthermore, since δ' is preferable to δ by a δ-supermajority, which is greater than a δ'-supermajority, it is also Posterior Consistent. We now argue that δ' uniquely satisfies the axioms. A $\delta'' > \delta'$ cannot be chosen because according to Posterior Consistency it must be preferred by a δ''-supermajority, which violates the definition of δ'. A $\delta'' < \delta'$ cannot be accepted without violating Conservatism because δ' would be accepted.

3. If there is no δ-supermajority to increase or decrease δ then of course δ should not change. No amendment proposal could be accepted according to δ itself, so remaining at δ is Posterior Consistent and trivially Conservative.

This finishes the proof of Theorem 19. □

Proposition 9 (Increasing Condorcet Amendment is Idempotent)**.** *Given the initial decision rule $\delta = \frac{1}{2}$ and any ideal points profile, the application Condorcet Consistent Amendment is idempotent.*

Proof. When the current rule is $\delta = \frac{1}{2}$, any change to the decision rule must increase δ. From Theorem 18, under Condorcet Amendment, δ' will be the largest such that the number of agents with ideal points at least δ' is strictly greater than $\delta' n$. Let us refer to this decision rule as $\hat{\delta} \in \Delta$. Since a δ'-supermajority prefer $\hat{\delta}$ over δ, we know the majority cannot prefer any $\delta'' < \hat{\delta}$ over $\hat{\delta}$, so once $\hat{\delta}$ takes over as the current decision rule, no proposal to decrease it will be accepted. For an amendment $\delta'' > \hat{\delta}$ to occur, there would have to be a δ''-supermajority who prefer δ'' over $\hat{\delta}$, but this violates the definition of $\hat{\delta}$. Thus, Condorcet Amendment is idempotent and chooses $\hat{\delta}$. □

Proposition 10 (Iterate from Simple Majority)**.** *Given the initial decision rule $\delta = \frac{1}{2}$ and any ideal points profile, the result of iterative application to completion of Conservative Amendment is the same as a single application of Condorcet Amendment.*

Proof. Suppose that the iterative application of Conservative Amendment were to select some $\delta' \neq \hat{\delta}$. First, suppose that $\delta' < \hat{\delta}$. A $\hat{\delta}$-supermajority prefers $\hat{\delta}$ to δ', which is larger than a δ-supermajority, so Conservative Amendment, given δ', would approve the proposal $\hat{\delta}$. This contradicts the fact that the iterative application of Conservative Amendment reached completion. Second, suppose $\delta' > \hat{\delta}$. From Theorem 19, this implies that more than $\delta' n$ agents have ideal points of at least δ', because it is accepted when proposed against a smaller δ. This contradicts the definition of $\hat{\delta}$. □

Here, we have offered an alternative axiomatic basis, replacing Condorcet Consistency (Axioms 7, 15) with Conservatism (Axioms 7(a) and 15(a)), discussed its consequences, and compared it to the initial axiomatic basis proposed.

CHAPTER 11
POLITICAL SOCIAL CHOICE

"To lay a broader foundation for a constructive social choice theory (broader than the framework Arrow chose), we have to make room for accommodating more information." - Amartya Sen [6]

11.1 Introduction

If scholars find it worthwhile to debate the normative ethics surrounding collective choice, then surely the agents being modeled should have similar concerns. What is needed is an extension of Social Choice Theory that accounts explicitly for both agents' preferences regarding the possible decision outcomes and their preferences about *how* the decision is made.

In utilitarian models, each agent receives a specific utility for every possible outcome. All relevant normative concerns of agents are assumed to be captured by their utility functions. However, there are significant issues with this approach. First, how can a utility function capture that an agent cares about whether the decision procedure satisfies axioms like Arrow's non-dictatoriality or independence of irrelevant alternatives? Their utility would need to be conditioned on the preference profile itself, which is supposed to be derived from, and consistent with, their utility. Resolving this paradox would require the introduction of an independent variable. For example, one may try to resolve it by allowing each agent to have two different types of utility, one of which is independent of the decision procedure and only dependent on the outcome, but this undermines the very premise of utility theory. Note that this criticism holds even in models with transferable utilities. Second, Amartya Sen's criticism of welfarism applies, as it may be unwise to assume that agents' expressed preferences capture all relevant aspects of welfare [6]. Third, human beings are not utilitarian. People behave in ways that are not strictly instrumental, and do not depend solely on consequentialist reasoning. Otherwise, people would not vote in large populations where the probability of casting a pivotal vote is negligible. It is important to note that such behavior is not irrational as is sometimes claimed [10]. A framework for social choice capable of producing models of rational collective decisions must be capable of capturing preferences over axioms or aggregation procedures. If we aggregate only preferences over outcomes,

without consideration for agent preferences for how the decision is made, we ignore relevant information.

In Chapter 10 the model consists of a set of agents who agree on a set of axioms that define a class of voting rules (δ-supermajority rules), and also on additional axioms that determine how the rule in force can amend itself based on agents' preferences over this class of rules. The present chapter does not assume all agents accept the same set of axioms.

The framework proposed below is founded on a simple observation: for a single problem instance with a single preference profile, many different collective choice procedures may select the same outcome. When we characterize collective choice procedures by the sets of axioms they do and do not satisfy, the multitude of possible axiomatic explanations for a given outcome can be seen as different, possibly unrelated, justifications for the same decision. When axioms depend on counterfactual reasoning (e.g. inter-profile comparisons), there is a possibility that agents may differ in their counterfactual reasoning. Agents may hold different views on *why* the outcome of a collective choice is what it is. Hence, different agents with the same information can be satisfied by the same decision for different reasons.

11.1.1 Concrete Problems

Section 11.2 will introduce the general abstract framework for Political Social Choice. Before then, it helps to first look at the two concrete problems to which we will apply this framework in the sections that follow.

Arrow's Axioms The first problem we consider is based on the axioms of Arrow's General Impossibility Theorem. Arrow's Theorem says that no social welfare function can satisfy three desirable axioms for all preference profiles. A social welfare function takes a profile of weak orders over a set of alternatives as input and returns a weak order over those alternatives called the social order as output. The phrase "all preference profiles" therefore refers to all profiles of weak orders, and excludes other kinds of preference information. The three axioms that no social welfare function can simultaneously satisfy for all profiles are Weak Pareto (also called Unanimity), Independence of Irrelevant Alternatives, and Non-dictatoriality.

If we want to select a social welfare function to use (before preferences over the alternatives are known), then we must give up one of Arrow's axioms. However, reasonable people can disagree about what axiom should be given up in any particular circumstance.

Arrow's set of axioms is minimal in the set that for any proper subset of the axioms there are social welfare functions that can satisfy all of its axioms at once. In other words, there are three pairs of two axioms that one could try to satisfy.

Now suppose we have three philosophers who disagree on what axiom should be given up, so they each demand a different pair of axioms be maintained. Once we observe a preference profile (not necessarily including preferences of the philosophers), can we find a set of three social welfare functions such each of the philosophers is satisfied by applying at least one of the functions to the given profile and all three functions select the same outcome?

Constitutional Amendments Recall the preference model in Chapter 10 for constitutional amendments. Each agent has an ideal supermajority rule corresponding to a threshold value of $\delta \in [\frac{1}{2}, 1)$ that they believe should be used for decisions, including the decision of whether to amend the current rule.

Suppose you are one of these agents and your ideal rule is $\delta_{you} = \frac{2}{3}$. The current rule is $\frac{1}{2}$, and a proposed conservative amendment to increment the rule by $\frac{1}{n}$ is accepted because it is approved by exactly a $\frac{3}{5}$-majority. How would you feel about this amendment being accepted? On the one hand, the rule has come closer to your ideal. If your ideal rule is $\frac{2}{3}$, assuming your preferences over δ values are single-peaked over the real line, then presumably you prefer the outcome to the status quo of $\frac{1}{2}$. However, had your ideal rule been used to make the same decision, the amendment would have failed and the current rule would remain $\frac{1}{2}$.

Also consider the symmetric case, where your ideal rule is $\delta_{you} = \frac{1}{2}$, the status quo is $\frac{2}{3}$ and a proposed conservative amendment to increment the rule is accepted. This time the same outcome would be selected if your ideal rule were used to make the decision, but you prefer the status quo to the proposal. How would you feel about this amendment being accepted?

11.2 Abstract Framework

Let N be a set of n agents and let X be a set of alternatives. The collection of agent preferences over the alternatives X constitutes a *profile* $\mathcal{P}$.

Preference Aggregation Procedures A preference aggregation procedure F is a function that maps profiles of preferences (within the relevant domain of possible profiles) to decision outcomes (within the relevant codomain Y). Sometimes the set of aggregation procedures considered will be restricted to a common class of procedures from related literature. When the aggregation procedure is a voting rule, the outcome will be a subset of alternatives. When the procedure is an Arrovian social welfare function, agent preferences and the outcome will be restricted to orderings of the alternatives. Generally, an aggregation procedure may yield any relation of the alternatives as an outcome for some profile.

Decisions The selection of an outcome $y^* \in Y$ based on a preference profile $\mathcal{P}$ is called a *decision*. Preference aggregation procedures can be used to make decisions. Given a fixed profile $\mathcal{P}$, every preference aggregation procedure F with $\mathcal{P}$ in its domain maps to a single outcome $F(\mathcal{P}) = y^*$. A decision is represented by a (profile, procedure, outcome) triple, $(\mathcal{P}, F, y^*)$ such that $F(\mathcal{P}) = y^*$. If $F(\mathcal{P}) \neq y^*$, then $(\mathcal{P}, F, y^*)$ is not a decision.

Remark 2. Notice that the notation for decisions contains redundancy. since $F(\mathcal{P}) = y^*$ for any decision, a decision can be identified entirely by $(F, \mathcal{P})$. This redundancy is for clarity and will make the application of axioms more intuitive.

Axioms We do not formally define what an axiom is, and talk rather about what an axiom does. Informally, an *axiom* λ restricts what decisions are permissible, and an axiom is said to be satisfied by any decision in its permissible set $S(\lambda)$. Taking terminology from [212], a *normative basis* $\Lambda = \{\lambda_1, \lambda_2, \ldots\}$ is a set of axioms that one wants to satisfy simultaneously. The permissible set of a normative basis is the intersection of the permissible sets of its axioms $S(\Lambda) = \bigcap_{\lambda \in \Lambda} S(\lambda)$, and the basis is said to be satisfied by any decision in its permissible set $S(\Lambda)$.

We generalize normative bases by introducing normative chassis. A *normative chassis* $\mathbf{\Lambda} = \{\Lambda_1, \Lambda_2, \ldots\}$ is a collection of normative bases such that a decision satisfies a chassis if it satisfies at least one of the normative bases in the chassis. In other words, the permissible set of a chassis is the union of the permissible sets of its bases $S(\mathbf{\Lambda}) = \bigcup_{\Lambda \in \mathbf{\Lambda}} S(\Lambda)$. Normative chassis reflect the idea that a person may be willing to accept multiple explanations for the same decision or different explanations for different decisions.

If the permissible set of a normative basis is the empty set, the basis is said to be

trivial; otherwise it is said to be non-trivial. A normative chassis is non-trivial if at least one of its normative bases is non-trivial.

Remark 3. Notice that since axioms, normative bases, and normative chassis all restrict the set of permissible decisions, one could try to represent any normative basis by a single highly complex axiom. However, this axiom would be hard to interpret. The introduction of new terminology is again for clarity and intuition. For example, we typically think of Arrow's Theorem as being built upon multiple incompatible axioms, not a single complicated axiom. This is part of why we stay away from trying to define what an axiom is, and focus instead on what an axiom does.

Implementation Invariance In reality, only a single procedure F is used to make a single decision. We say that this procedure is *actually implemented*. When a decision $(\mathcal{P}, F, y^*)$ is actually implemented, we say that all decisions $(\mathcal{P}, \tilde{F}, y^*)$ are *effectively implemented*, because had the procedure $\tilde{F}$ been used instead on the same profile, the same outcome would have resulted. If one observes only the input $(\mathcal{P})$ and output (y^*) of a decision $(\mathcal{P}, F, y^*)$, one cannot necessarily say what procedure was actually implemented to produce the outcome, and can say only what procedures were effectively implemented.

In some circumstances, agents may care strongly about what procedure is actually implemented. For example, it can be important to use an aggregation procedure whose outcome can be computed in polynomial time or one that preserves agent privacy. In other cases, agents may not care about what procedure was actually implemented, but care about more than just the decision outcome. Agents may seek a justification or explanation for the decision itself.

For the purposes of this chapter, all normative chassis (or the agents who impose them) are assumed to be implementation invariant, meaning that when a decision is made the chassis is satisfied if any of the decisions in its permissible set are effectively implemented. Our goal is to find decisions that can be actually implemented such that all agents are satisfied by at least one decision that was effectively implemented.

11.3 Arrow's Axioms

Let each agent $i \in N$ have preferences in the form of a weak order (reflexive, transitive, and complete binary relation) over the alternatives X. The profile $\mathcal{P}$ is the collection of the

agents' individual preference orders. A social welfare function F takes any profile as input and outputs a weak order over the alternatives called a social order. In other words, given any fixed profile $\mathcal{P}$, a social welfare function F maps the profile to a single social order $y^* \in Y$ where Y is the set of all possible weak orders over X.

Arrow's General Impossibility Theorem states that no social welfare function can satisfy three axioms at once - Non-dictatoriality (D), Weak Pareto (P), and Independence of Irrelevant Alternatives (I). This means that a normative basis Λ which only permits decisions made by social welfare functions and contains all three axioms $\{D, P, I\} \subset \Lambda$ must be trivial. There is no way to satisfy an agent who demands that a social welfare function be used that satisfies all three axioms, even if they only require it to be effectively implemented. Of course, if any single agent demands the impossible, then no decision can satisfy everyone. But as long as no single individual demands the impossible there remains hope, even if the union of their normative bases would be trivial.

Definition 11 (Weak Pareto (P))**.** A social welfare function F satisfies the Weak Pareto axiom if for all profiles $\mathcal{P}$ in which all agents $i \in N$ strictly prefer x over y ($x \succ_i y$) for $x, y \in X$, x is ranked above y ($x \succ y$) in the social order $F(\mathcal{P})$.

Definition 12 (Non-dictatoriality (D))**.** A social welfare function F satisfies the non-dictatoriality axiom if there is no agent $i \in N$ such that for all profiles $\mathcal{P}$ containing the preference order of agent i, $F(\mathcal{P})$ is exactly the preference order of i.

Definition 13 (Independence of Irrelevant Alternatives (I))**.** A social welfare function F satisfies independence of irrelevant alternatives if for any two preference profiles $\mathcal{P}$ and $\mathcal{P}'$ in which all agents $i \in N$ rank alternatives $x, y \in X$ in the same order in both profiles, x and y must appear in the same order in $F(\mathcal{P})$ and $F(\mathcal{P}')$.

Suppose we have three agents who each maintain that two of Arrow's axioms $\{P, I, D\}$ must be upheld, but disagree on which one can be dispensed with. Each agent will be satisfied by any effectively implemented decision that satisfies their basis. Even if the normative bases of the agents may consist of all proper subsets of Arrow's axioms, for any fixed profile we can compute a social order and set of social welfare functions that will satisfy everyone. In other words, if agents only use Arrow's axioms as their normative principles but accept effective implementation, as long as no single agent demands the impossible, we can satisfy everyone.

Theorem 20 (Political Possibility with Arrow's Axioms)**.** *Let X be a finite set of three or more alternatives and N be a set of $n \geq 2$ agents whose preferences are weak orders over X. For all profiles $\mathcal{P}$ of weak orders, there exists a set of social welfare functions $\{F_1, F_2, F_3\}$ such that $(\mathcal{P}, F_1, y^*) \in S(\{P, I\})$, $(\mathcal{P}, F_2, y^*) \in S(\{P, D\})$, $(\mathcal{P}, F_3, y^*) \in S(\{I, D\})$, and $F_1(\mathcal{P}) = F_2(\mathcal{P}) = F_3(\mathcal{P}) = y^*$.*

Proof. Let F_1 be the dictatorial social welfare function that always selects the preference order of agent $v_1 \in N$ as the social order. Notice first that $(\mathcal{P}, F_1, y^*) \in S(\{P, I\})$ since it is dictatorial. I will use $y^* = F_1(\mathcal{P})$ from here on.

Now let x and y be two alternatives in X such that v_1 prefers $x \succeq y$. I designate F_2 to be the social welfare function that selects the preference order of agent v_1 if and only if v_1 prefers $x \succeq y$, and selects the preference order of agent v_2 as the social order otherwise. Notice second that $(\mathcal{P}, F_2, y^*) \in S(\{P, D\})$ and $F_2(\mathcal{P}) = y^*$.

Lastly, let F_3 be the constant function that returns y^*, not just for this profile but for all possible profiles of weak orders, thereby ignoring the preferences of the agents entirely. Finally, we can see that $(\mathcal{P}, F_3, y^*) \in S(\{I, D\})$ and $F_3(\mathcal{P}) = y^*$. □

The example constructed in the proof above feels deeply dissatisfying for a number of reasons. First, while some people may advocate for dictatorship in some circumstances, it is unlikely to be on the basis that it preserves the Weak Pareto and Independence of Irrelevant Alternatives axioms. Second, it carves out an exception based on the single profile observed to provide a desperately *ad hoc* explanation to each agent. Third, the social welfare functions used are not very attractive, especially to agents other than v_1.

Despite the simplicity of the result and *ad hoc* nature of the construction, Theorem 20 illustrates an important possibility. There are situations where no single social welfare function can satisfy a set of axioms, but if agents only care about (or only now about) what functions are effectively implemented, and no normative basis individually demands the impossible, then some set of effectively implemented social welfare functions might be able to satisfy all normative bases at once. It is easy to see how this concept can be extended to other impossibility theorems that contain axioms relying upon inter-profile comparisons and other classes of preference aggregation procedures.

11.4 Revisiting Constitutional Amendments

In Chapter 10, a set of agents agreed upon a set of axioms that narrowed down the set of possible decisions, including amendments, to those made by approval vote using δ-supermajority rules, and a single initial rule to start. The agents could disagree on what δ should be used for decisions, and then use the current rule to update itself based on those preferences. Fortunately, even when the agents do not agree on the additional axioms for amendments - Posterior Consistency and Condorcet Consistency or Conservatism - amendments are still possible, as long as agents are willing to accept decisions based on effectively implemented rules.

Recall that in Chapter 10 each agent $i \in N$ has an ideal point δ_i representing their top choice of δ-supermajority rule, and their preferences over the set of possible δ values are single-peaked over the real line with their ideal point as the peak. Under the conservative amendment procedure, the agents start at majority rule and iteratively increment the δ, proposing increasing integer multiples of $\frac{1}{n}$ until the amendment procedure halts. This amendment procedure follows directly from a set of axioms agreed to by all agents. So for every potential amendment, the agents have different preferences over whether the amendment should occur (approve or do not approve), but given the profile they all agree on what rule should be used. We now show a stronger claim, that the founding and amending of a constitution with conservative amendments can satisfy all agents even when they disagree on what amendment rule should be used.

Suppose that each of our agents has a normative chassis $\mathbf{\Lambda} = \{\Lambda_1, \Lambda_2\}$ that an amendment decision can satisfy in two ways. The first is that an agent is satisfied by a decision if it corresponds to their vote. If an agent approves an amendment proposal ($V_i^{\delta,\delta'} = 1$), and the new proposed δ' is accepted, then the agent is satisfied. Likewise if the agent did not approve the amendment, then they are satisfied if the proposal is rejected. The second way is that an agent is satisfied by a decision $(V^{\delta,\delta'}, F, \delta^*)$ if the same decision would be made according to their ideal rule $F^{\delta_i}(V^{\delta,\delta'}) = \delta^*$, so the decision $(V^{\delta,\delta'}, F^{\delta_i}, \delta^*)$ is effectively implemented.

Theorem 21. *Every amendment under the Conservative Amendment procedure beginning at simple majority ($\delta = \frac{1}{2}$) is guaranteed to satisfy all agents with the normative chassis $\mathbf{\Lambda}$ defined above.*

Proof. The two cases we must consider are when a proposal is accepted and when it is

rejected. For each case, we must ensure that both voters who approved of the amendment and those who did not approve are satisfied with the decision. In both cases we assume the proposed amendment is from δ to $\delta' = \delta + \frac{1}{n}$.

1. Suppose $F(V^{\delta,\delta'}) = 1$. For all agents $\delta_i < \delta'$, they must be satisfied by the amendment because the amendment is accepted by a δ'-supermajority, which is greater than a δ_i-supermajority. All agents $\delta_i \geq \delta'$ must be satisfied by the amendment because they approved of the amendment that was accepted, so their preferred outcome prevailed.

2. Suppose $F(V^{\delta,\delta'}) = 0$. Since the amendment was rejected, it must not have been approved by a δ'-supermajority. Therefore all agents $\delta_i \geq \delta'$ must be satisfied because the proposal would also have been rejected according to their ideal rule. For all agents $\delta_i < \delta'$, they must be satisfied because they did not approve of the amendment and the amendment was rejected.

□

Recall from Chapter 10 that when the Conservative amendment procedure is applied iteratively, it ultimately selects the same δ-supermajority rule that the Condorcet Consistent amendment procedure selects in one step. Therefore, if the Condorcet Consistent amendment procedure is actually implemented, then the iterative Conservative amendment procedure is effectively implemented. The actual Condorcet Consistent amendment from $\delta = \frac{1}{2}$ directly to a new $\delta^* > \frac{1}{2} + \frac{1}{n}$ will also satisfy all agents according to the normative chassis we have described, because all agents $\delta_i \geq \delta^*$ will be satisfied because they approve, while all agents with $\delta_i < \delta^*$ will be satisfied because their ideal rule was effectively implemented.

Lastly, it is worth noting that our Condorcet Constitution is strategyproof over the domain of single-peaked preferences because it corresponds to a generalized median rule [232]–[234].

CHAPTER 12
FUTURE DIRECTIONS

12.1 Implicit Utilitarian Models

The study of implicit utilitarian models is the study of how the approximation that one can achieve depends on the strength of assumptions that can be made about preference structure and the amount and type of information agents provide. When stronger assumptions are made about agent preferences, better approximation algorithms can be devised. Similarly, the more information agents provide, the better approximation one should be able to construct.

In Chapter 5, extra information about preference strengths enabled better approximation algorithms. The same concept can be applied to the model of Chapter 4. It is an interesting question how much the approximation factors for ABC systems can be improved if agents provide preference strength information in the form of the ratio of the weights of adjacent edges. Crucially, additional information about agent preferences means that an algorithm can use more than just a partial order, so algorithms involving improvement cycles become viable again.

The model in Chapter 4 makes no assumptions about agent preferences other than that the utilities (edge weights) are non-negative and symmetric. One could consider either relaxing these assumptions or strengthening them. An interesting approach to strengthening the assumptions on preference structure is to assume a trinary relation on edge weights like triangle inequality. The metric assumption has been used for matching with non-negative utilities, though it is more common for the setting in which agents experience negative utilities (costs). Instead, a homophily assumption may be more appropriate to capture the idea that "a friend of a friend is a friend." This relation would look like $w(a, b) + w(b, c) \leq \beta w(a, c)$ parameterized by some factor β. Complementarily, a natural relaxation is not to assume that the weights are symmetric. The graph representation would be a directed graph with parallel edges can have different weights. It is an interesting question for what problems small constant-factor approximations can be preserved. This is an equivalent setup to assuming that the agents share the utility from an undirected edge unevenly.

For both chapters of Part I, one can consider variations in which the agents provide not a different quantity of information, but different types of information. For example, agents

can report approval preferences, with or without preference strengths, rather than ordinal preferences. It is an open question whether nice bounds can be achieved on ideal candidate distortion with categorical preferences.

Taking a broader view, there are a wide variety of objectives one might consider rather than maximizing Social Welfare or minimizing Social Cost, including measures of fairness. There are also many different problems for which implicit utilitarian models may be useful that have yet to be explored. One clear example of a problem conducive to the implicit utilitarian treatment is the problem of forking in Chapter 7.

12.2 Preferences Over Both Agents and Alternatives

Chapter 7 has agents give preference orders over communities based on their size and what alternative the community is centered around. While the focus of Chapter 7 is on stability, one can also consider a utilitarian approach. If agents' preference orders are assumed to be induced by underlying utility functions, we can ask many of the same questions posed in the previous two chapters. One particularly interesting idea is to define a *Price of Cohesion* which would compare the social welfare when agents are forced to all belong to the same community to the optimum social welfare among all possible assignments of agents to communities. This is based on the ideal of ideal candidate distortion, as it allows us to take the social welfare of a constructed solution among only a subset of the conceivable assignments and compare it to the best conceivable assignment.

There are many different variations of the models and problems in Chapter 7 and Chapter 8 worth consideration. The modeling of flexible representative voting in particular included many different parameters that had to be specified for analysis. One can easily see how the model can be extended to consider issues with more than two alternatives, different preference formats, different agent delegation behaviors, uncertainty, and many other considerations. Each of these modifications raises new challenges. For example, when there are multiple alternatives, if voters' preferences over candidates and representatives are derived from similarity in their preferences over the alternatives, then there must be a clear notion of distance between preferences. Another appealing direction is to consider that agents may have non-consequentialist preferences, such as preferences over the distribution of power.

The model of flexible representative voting presented in Chapter 8 abstracts away all

temporal considerations for simplicity, but this is an important direction for future work. I include here the prospect of building on Condorcet Jury Theorems for revealing ground truths and agents learning the quality of representatives over time as the representatives votes are revealed.

12.3 Preferences Over Axioms and Rules

When founding and amending a constitution, the agents of Chapter 10 are restricted to taking approval votes. An obvious next step is to determine how, after founding a constitution, they can select a rule for taking other kinds of votes when there are many alternatives to choose from, or how to order proposals when they must be considered sequentially and have dependencies.

The framework of Political Social Choice proposed in Chapter 11 opens a new space of preference aggregation methods, including models of deliberation, that rely on agents reasoning differently about the same set of facts. There is clearly a pressing need to identify collections of normative chassis that realistically reflect common preferences for which collective decisions exist that satisfy everyone, and to characterize the conditions under which this is possible. In addition, one might consider preference aggregation methods that only need to satisfy some subset of the agents, say to prevent a fork.

REFERENCES

[1] B. Abramowitz and E. Anshelevich, "Utilitarians without utilities: Maximizing social welfare for graph problems using only ordinal preferences," in *Proc. 32nd AAAI Conf. Artif. Intell.*, S. A. McIlraith and K. Q. Weinberger, Eds., Feb. 2018, pp. 894–901.

[2] B. Abramowitz, E. Anshelevich, and W. Zhu, "Awareness of voter passion greatly improves the distortion of metric social choice," in *Web and Internet Econ. - 15th Int. Conf. Proc.*, vol. 11920, I. Caragiannis, V. S. Mirrokni, and E. Nikolova, Eds., Dec. 2019, pp. 3–16, doi: 10.1007/978-3-030-35389-6_1.

[3] B. Abramowitz and N. Mattei, "Flexible representative democracy: An introduction with binary issues," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, S. Kraus, Ed., Macao, China, Aug. 2019, pp. 3–10, doi: 10.24963/ijcai.2019/1.

[4] B. Abramowitz, E. Shapiro, and N. Talmon, "How to amend a constitution? Model, axioms, and supermajority rules: (extended abstract)," in *Proc. 20th Int. Conf. Auton. Agents and Multi-agent Syst.*, May 2021, pp. 1443–1445.

[5] B. Abramowitz, E. Elkind, D. Grossi, E. Shapiro, and N. Talmon, "Democratic forking: Choosing sides with social choice," Mar. 2021, arXiv:abs/2103.03652 (Retrieved on 6/16/2021).

[6] A. Sen, *Collective Choice and Social Welfare - An Expanded Edition.* Cambridge, MA, USA: Harvard Univ. Press, 2018.

[7] H. A. Simon, *Administrative Behavior*, 4th ed. New York, NY, USA: Simon and Schuster, 2013.

[8] S. Wolfram, *A Project to Find the Fundamental Theory of Physics.* Champaign, IL, USA: Wolfram Media, 2020.

[9] ——, *A New Kind of Science.* Champaign, IL, USA: Wolfram Media, 2002.

[10] H. Gintis, *Individuality and Entanglement: The Moral and Material Bases of Social Life.* Princeton, NJ, USA: Princeton Univ. Press, 2016.

[11] B. Russell, *Proposed Roads to Freedom: Socialism, Anarchism and Syndicalism.* New York, NY, USA: Henry Holt, 1920.

[12] D. R. Hofstadter, *Gödel, Escher, Bach.* New York, NY, USA: Basic Books, 1979.

[13] C. Darwin, *On the Origin of Species.* New York, NY, USA: Dover, 2006.

[14] M. A. Bakunin, *God and the State.* New York, NY, USA: Dover, 1970.

[15] R. Rocker, *Anarcho-syndicalism: Theory and Practice*, 6th ed. Oakland, CA, USA: AK Press, 2004.

[16] P. A. Kropotkin, *The State: Its Historic Role*, 4th ed. London, UK: Freedom Press, 1911.

[17] S. Li, "Ethics and market design," *Oxford Rev. Econ. Policy*, vol. 33, no. 4, pp. 705–720, Nov. 2017, doi: 10.1093/oxrep/grx047.

[18] J. Rawls, *Lectures on the History of Political Philosophy*. Cambridge, MA, USA: Harvard Univ. Press, 2007.

[19] K. J. Arrow, *Social Choice and Individual Values*, 3rd ed. New Haven, CT, USA: Yale Univ. Press, 2012.

[20] J. C. Harsanyi, "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility," *J. Political Econ.*, vol. 63, no. 4, pp. 309–321, Aug. 1955, doi: 10.1086/257678.

[21] A. Sen, "On weights and measures: Informational constraints in social welfare analysis," *Econometrica*, vol. 45, no. 7, pp. 1539–1572, Oct. 1977, doi: 10.2307/1913949.

[22] A. D. Procaccia and J. S. Rosenschein, "The distortion of cardinal preferences in voting," in *Cooperative Inf. Agents, 10th Int. Workshop Proc.*, vol. 4149, M. Klusch, M. Rovatsos, and T. R. Payne, Eds., Sep. 2006, pp. 317–331, doi: 10.1007/11839354_23.

[23] E. Anshelevich, O. Bhardwaj, E. Elkind, J. Postl, and P. Skowron, "Approximating optimal social choice under metric preferences," *Artif. Intell.*, vol. 264, no. 2, pp. 27–51, Aug. 2018, doi: 10.1016/j.artint.2018.07.006.

[24] C. Boutilier, I. Caragiannis, S. Haber, T. Lu, A. D. Procaccia, and O. Sheffet, "Optimal social choice functions: A utilitarian view," *Artif. Intell.*, vol. 227, no. 7, pp. 190–213, Jun. 2015, doi: 10.1016/j.artint.2015.06.003.

[25] E. Anshelevich, A. Filos-Ratsikas, N. Shah, and A. A. Voudouris, "Distortion in social choice problems: The first 15 years and beyond," Mar. 2021, arXiv:abs/2103.00911. (Retrieved on 6/16/2021).

[26] E. Anshelevich, "Ordinal approximation in matching and social choice," *ACM SIGecom Exchanges*, vol. 15, no. 1, pp. 60–64, Jul. 2016.

[27] E. Anshelevich and S. Sekar, "Blind, greedy, and random: Algorithms for matching and clustering using only ordinal information," in *Proc. 13th AAAI Conf. Artif. Intell.*, D. Schuurmans and M. P. Wellman, Eds., Phoenix, AZ, USA, Feb. 2016, pp. 390–396.

[28] ——, "Truthful mechanisms for matching and clustering in an ordinal world," in *Web and Internet Econ. - 12th Int. Conf. Proc.*, vol. 10123, Y. Cai and A. Vetta, Eds., Montreal, Canada, Dec. 2016, pp. 265–278, doi: 10.1007/978-3-662-54110-4_19.

[29] E. Anshelevich and W. Zhu, "Tradeoffs between information and ordinal approximation for bipartite matching," *Theory of Comput. Syst.*, vol. 63, no. 7, pp. 1499–1530, Oct. 2019, doi: 10.1007/s00224-018-9886-x.

[30] G. Christodoulou, A. Filos-Ratsikas, S. K. S. Frederiksen, P. W. Goldberg, J. Zhang, and J. Zhang, "Social welfare in one-sided matching mechanisms," in *Proc. Int. Conf. Auton. Agents & Multiagent Syst.*, vol. 10002, N. Osman and C. Sierra, Eds., no. 3, Singapore, Sep. 2016, pp. 30–50, doi: 10.1007/978-3-319-46882-2_3.

[31] I. Caragiannis, A. Filos-Ratsikas, S. K. S. Frederiksen, K. A. Hansen, and Z. Tan, "Truthful facility assignment with resource augmentation: An exact analysis of serial dictatorship," in *Web and Internet Econ. - 12th Int. Conf. Proc.*, vol. 10123, Y. Cai and A. Vetta, Eds., Montreal, Canada, Dec. 2016, pp. 236–250, doi: 10.1007/978-3-662-54110-4_17.

[32] G. Amanatidis, G. Birmpas, A. Filos-Ratsikas, and A. A. Voudouris, "A few queries go a long way: Information-distortion tradeoffs in matching," in *35th AAAI Conf. Artif. Intell.*, May 2021, pp. 5078–5085.

[33] E. Anshelevich and W. Zhu, "Ordinal approximation for social choice, matching, and facility location problems given candidate positions," *ACM Trans. Econ. and Comput.*, vol. 9, no. 2, pp. 1–24, Jun. 2021, doi: 10.1145/3434417.

[34] T. Ma, V. Menon, and K. Larson, "Improving welfare in one-sided matching using simple threshold queries," Nov. 2020, arXiv:abs/2011.13977. (Retrieved on 6/16/2021).

[35] A. Filos-Ratsikas, S. K. S. Frederiksen, and J. Zhang, "Social welfare in one-sided matchings: Random priority and beyond," in *Int. Symp. Algorithmic Game Theory*, vol. 8768, Sep./Oct. 2014, pp. 1–12, doi: 10.1007/978-3-662-44803-8_1.

[36] E. Anshelevich and J. Postl, "Randomized social choice functions under metric preferences," *J. Artif. Intell. Res.*, vol. 58, no. 20, pp. 797–827, Jan. 2017, doi: 10.1613/jair.5340.

[37] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, Eds., *Handbook of Computational Social Choice*. Cambridge, UK: Cambridge Univ. Press, 2016, doi: 10.1017/CBO9781107446984.

[38] I. Caragiannis, S. Nath, A. D. Procaccia, and N. Shah, "Subset selection via implicit utilitarian voting," *J. Artif. Intell. Res.*, vol. 58, no. 3, pp. 123–152, Jan. 2017, doi: 10.1613/jair.5282.

[39] I. Caragiannis and A. D. Procaccia, "Voting almost maximizes social welfare despite limited communication," *Artif. Intell.*, vol. 175, no. 9-10, pp. 1655–1671, Apr. 2011, doi: 10.1016/j.artint.2011.03.005.

[40] A. Goel, A. K. Krishnaswamy, and K. Munagala, "Metric distortion of social choice rules: Lower bounds and fairness properties," in *Proc. 2017 ACM Conf. Econ. and Comput.*, C. Daskalakis, M. Babaioff, and H. Moulin, Eds., Cambridge, MA, USA, Jun. 2017, pp. 287–304, doi: 10.1145/3033274.3085138.

[41] A. Goel, R. Hulett, and A. K. Krishnaswamy, "Relating metric distortion and fairness of social choice rules," in *Proc. 13th Workshop on Econ. of Netw., Syst. and Comput.*, Irvine, CA, USA, Jun. 2018, p. 4:1, doi: 10.1145/3230654.3230658.

[42] S. Gross, E. Anshelevich, and L. Xia, "Vote until two of you agree: Mechanisms with small distortion and sample complexity," in *Proc. 31st AAAI Conf. Artif. Intell.*, S. P. Singh and S. Markovitch, Eds., San Francisco, CA, USA, Feb. 2017, pp. 544–550.

[43] P. K. Skowron and E. Elkind, "Social choice under metric preferences: Scoring rules and STV," in *Proc. 31st AAAI Conf. Artif. Intell.*, S. P. Singh and S. Markovitch, Eds., San Francisco, CA, USA, Feb. 2017, pp. 706–712.

[44] D. Mandal, N. Shah, and D. P. Woodruff, "Optimal communication-distortion tradeoff in voting," in *The 21st ACM Conf. Econ. and Comput.*, P. Biró, J. D. Hartline, M. Ostrovsky, and A. D. Procaccia, Eds., Jul. 2020, pp. 795–813, doi: 10.1145/3391403.3399510.

[45] X. Chen, M. Li, and C. Wang, "Favorite-candidate voting for eliminating the least popular candidate in a metric space," in *The 34th AAAI Conf. Artif. Intell.*, New York, NY, USA, Feb. 2020, pp. 1894–1901.

[46] A. Borodin, O. Lev, N. Shah, and T. Strangway, "Primarily about primaries," in *The 33rd AAAI Conf. Artif. Intell.*, Honolulu, HI, USA, Jan./Feb. 2019, pp. 1804–1811, doi: 10.1609/aaai.v33i01.33011804.

[47] Y. Cheng, S. Dughmi, and D. Kempe, "Of the people: Voting is more effective with representative candidates," in *Proc. 2017 ACM Conf. Econ. and Comput.*, C. Daskalakis, M. Babaioff, and H. Moulin, Eds., Cambridge, MA, USA, Jun. 2017, pp. 305–322, doi: 10.1145/3033274.3085155.

[48] ——, "On the distortion of voting with multiple representative candidates," in *Proc. 32nd AAAI Conf. Artif. Intell.*, S. A. McIlraith and K. Q. Weinberger, Eds., New Orleans, LA, USA, Feb. 2018, pp. 973–980.

[49] B. Fain, A. Goel, K. Munagala, and N. Prabhu, "Random dictators with a random referee: Constant sample complexity mechanisms for social choice," in *The 33rd AAAI Conf. Artif. Intell.*, Honolulu, HI, USA, Jan./Feb. 2019, pp. 1893–1900, doi: 10.1609/aaai.v33i01.33011893.

[50] M. Feldman, A. Fiat, and I. Golomb, "On voting and facility location," in *Proc. Conf. Econ. and Comput.*, V. Conitzer, D. Bergemann, and Y. Chen, Eds., Maastricht, The Netherlands, Jul. 2016, pp. 269–286, doi: 10.1145/2940716.2940725.

[51] M. Seddighin, M. Latifian, and M. Ghodsi, "On the distortion value of elections with abstention," *J. Artif. Intell. Res.*, vol. 70, no. 15, pp. 567–595, Jan. 2021, doi: 10.1613/jair.1.12306.

[52] G. Pierczynski and P. Skowron, "Approval-based elections and distortion of voting rules," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, S. Kraus, Ed., Macao, China, Aug. 2019, pp. 543–549, doi: 10.24963/ijcai.2019/77.

[53] V. Gkatzelis, D. Halpern, and N. Shah, "Resolving the optimal metric distortion conjecture," in *61st IEEE Annu. Symp. Found. Comput. Sci.*, Durham, NC, USA, Nov. 2020, pp. 1427–1438, doi: 10.1109/FOCS46700.2020.00134.

[54] D. Mandal, A. D. Procaccia, N. Shah, and D. P. Woodruff, "Efficient and thrifty voting by any means necessary," in *Proc. 33rd Int. Conf. Neural Inf. Process. Syst.*, no. 645, Dec. 2019, pp. 7180–7191, doi: 10.5555/3454287.3454932.

[55] G. Amanatidis, G. Birmpas, A. Filos-Ratsikas, and A. A. Voudouris, "Peeking behind the ordinal curtain: Improving distortion via cardinal queries," in *34th AAAI Conf. Artif. Intell.*, New York, NY, USA, Feb. 2020, pp. 1782–1789.

[56] B. L. Meek, "A transferable voting system including intensity of preference," *Math. et Sci. Humaines*, vol. 50, pp. 23–29, 1975.

[57] W. D. Cook and M. Kress, "Ordinal ranking with intensity of preference," *Manag. Sci.*, vol. 31, no. 1, pp. 26–33, Jan. 1985, doi: 10.1287/mnsc.31.1.26.

[58] J. S. Dyer and R. K. Sarin, "Group preference aggregation rules based on strength of preference," *Manag. Sci.*, vol. 25, no. 9, pp. 813–930, Sep. 1979, doi: 10.1287/mnsc.25.9.822.

[59] J. R. Hauser and S. M. Shugan, "Intensity measures of consumer preference," *Oper. Res.*, vol. 28, no. 2, Apr. 2008, doi: 10.1287/opre.28.2.278.

[60] P. Resnick, "On Consensus and Humming in the IETF." Internet Eng. Task Force, San Diego, CA, USA, Memo. RFC 7282, Jun. 2014, https://www.hjp.at/doc/rfc/rfc7282.html (Retrieved on 7/18/2021).

[61] R. Preis, "Linear time 1/2-approximation algorithm for maximum weighted matching in general graphs," in *Proc. 16th Annu. Symp. Theor. Aspects of Comput. Sci.*, vol. 1563, C. Meinel and S. Tison, Eds., Trier, Germany, Mar. 1999, pp. 259–269, doi: 10.1007/3-540-49116-3_24.

[62] J. R. Edmonds, "Matroids and the greedy algorithm," *Math. Program.*, vol. 1, no. 1, pp. 127–136, Dec. 1971, doi: 10.1007/BF01584082.

[63] R. Rado, "A theorem on independence relations," *The Quart. J. Math.*, no. 1, pp. 83–89, Jan. 1942.

[64] B. Korte and D. Hausmann, "An analysis of the greedy heuristic for independence systems," in *Ann. Discrete Math.* vol. 2, New York, NY, USA: Elsevier, 1978, pp.65–74.

[65] M. Dyer, L. R. Foulds, and A. Frieze, "Analysis of heuristics for finding a maximum weight planar subgraph," *Eur. J. Oper. Res.*, vol. 20, no. 1, pp. 102–114, Apr. 1985, doi: 10.1016/0377-2217(85)90288-7.

[66] J. Mestre, "Greedy in approximation algorithms," in *Proc. 14th Annu. Eur. Symp.*, vol. 4168, Y. Azar and T. Erlbach, Zurich, Switzerland, Sep. 2006, pp. 528–539, doi: 10.1007/11841036_48.

[67] O. Peters, "The ergodicity problem in economics," *Nature Phys.*, vol. 15, no. 12, pp. 1216–1221, Dec. 2019, doi: 10.1038/s41567-019-0732-0.

[68] U. Bhaskar, V. Dani, and A. Ghosh, "Truthful and near-optimal mechanisms for welfare maximization in multi-winner elections," in *Proc. 32nd AAAI Conf. Artif. Intell.*, S. A. McIlraith and K. Q. Weinberger, Eds., New Orleans, LA, USA, Feb. 2018, pp. 925–932.

[69] S. Merrill III, S. Merrill, and B. Grofman, *A Unified Theory of Voting: Directional and Proximity Spatial Models.* Cambridge, UK: Cambridge Univ. Press, 1999.

[70] K. Arrow, *Advances in the Spatial Theory of Voting.* Cambridge, UK: Cambridge Univ. Press, 1990.

[71] J. M. Enelow and M. J. Hinich, *The Spatial Theory of Voting: An Introduction.* Cambridge, UK: Cambridge Univ. Press, 1984.

[72] A. Filos-Ratsikas and A. A. Voudouris, "Approximate mechanism design for distributed facility location," Sep. 2020, arXiv:abs/2007.06304 (Retrieved on 6/16/21).

[73] A. Bogomolnaia and J.-F. Laslier, "Euclidean preferences," *J. Math. Econ.*, vol. 43, no. 2, pp. 87–98, Feb. 2007, doi: 10.1016/j.jmateco.2006.09.004.

[74] G. Benadè, S. Nath, A. D. Procaccia, and N. Shah, "Preference elicitation for participatory budgeting," *Manag. Sci.*, vol. 67, no. 5, pp. 2813–2827, May 2021, doi: 10.1287/mnsc.2020.3666.

[75] G. Benadè, A. D. Procaccia, and M. Qiao, "Low-distortion social welfare functions," in *The 33rd AAAI Conf. Artif. Intell.*, Honolulu, HI, USA, Jan./Feb. 2019, pp. 1788–1795, doi: 10.1609/aaai.v33i01.33011788.

[76] E. Anshelevich, Z. Fitzsimmons, R. Vaish, and L. Xia, "Representative proxy voting," in *35th AAAI Conf. Artif. Intell.*, 2021, pp. 5086–5093.

[77] A. Filos-Ratsikas, E. Micha, and A. A. Voudouris, "The distortion of distributed voting," *Artif. Intell.*, vol. 286, no. 103343, Sep. 2020, doi: 10.1016/j.artint.2020.103343.

[78] D. E. Campbell, "Social choice and intensity of preference," *J. Political Econ.*, vol. 81, no. 1, pp. 211–218, Jan./Feb. 1973, doi: 10.1086/260018.

[79] P. H. Farquhar and L. R. Keller, "Preference intensity measurement," *Ann. Oper. Res.*, vol. 19, no. 1, pp. 205–217, Dec. 1989, 10.1007/BF02283521.

[80] G. Gerasimou, "Simple preference intensity comparisons," *J. Econ. Theory*, vol. 192, no. 105199, Mar. 2021, doi: 10.1016/j.jet.2021.105199.

[81] K. Willmoore and G. W. Carey, "The "intensity" problem and democratic theory," *Amer. Political Sci. Rev.*, vol. 62, no. 1, pp. 5–24, Aug. 1968, doi: 10.1017/S0003055400115618.

[82] G. Călinescu, C. G. Fernandes, H. J. Karloff, and A. Zelikovsky, "A new approximation algorithm for finding heavy planar subgraphs," *Algorithmica*, vol. 36, no. 2, pp. 179–205, Jun. 2003, doi: 10.1007/s00453-002-1020-3.

[83] K. E. Paluch, M. Mucha, and A. Madry, "A 7/9 - approximation algorithm for the maximum traveling salesman problem," in *Approx., Randomization, and Combinatorial Optim. Algorithms and Techn., Proc. 12th Int. Workshop*, vol. 5687, I. Dinur, K. Jansen, J. Naor, and J. D. P. Rolim, Eds., Berkeley, CA, USA, Aug. 2009, pp. 298–311, doi: 10.1007/978-3-642-03685-9_23.

[84] A. Vince, "A framework for the greedy algorithm," *Discret. Appl. Math.*, vol. 121, no. 1-3, pp. 247–260, Sep. 2002, doi: 10.1016/S0166-218X(01)00362-6.

[85] W. Zhu, "Matching, social welfare and ordinal approximation," Ph.D. dissertation, Dept. Comp. Sci., Rensselaer Polytech. Inst., Troy, NY, USA, 2020.

[86] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2019, unpublished.

[87] A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder, *Bitcoin and Cryptocurrency Technologies*. Princeton, NJ, USA: Princeton Univ. Press, 2016.

[88] A. Phillip, J. Chan, and S. Peiris, "A new look at cryptocurrencies," *Econ. Lett.*, vol. 163, pp. 6–9, Feb. 2018, doi: 10.1016/j.econlet.2017.11.020.

[89] N. Webb, "A fork in the blockchain: Income tax and the bitcoin/bitcoin cash hard fork," *North Carolina J. Law & Technol.*, vol. 19, no. 4, p. 283, May 2018.

[90] V. Allombert, M. Bourgoin, and J. Tesson, "Introduction to the tezos blockchain," in *17th Int. Conf. High Perform. Comput. & Simul.*, Dublin, Ireland, Jul. 2019, pp. 1–10, doi: 10.1109/HPCS48598.2019.9188227.

[91] R. Beck, C. Müller-Bloch, and J. L. King, "Governance in the blockchain economy: A framework and research agenda," *J. Assoc. Inf. Syst.*, vol. 19, no. 10, p. 1, Jan. 2018.

[92] W. Reijers, I. Wuisman, M. Mannan, P. De Filippi, C. Wray, V. Rae-Looi, A. C. Vélez, and L. Orgad, "Now the code runs itself: On-chain and off-chain governance of blockchain technologies," *Topoi*, pp. 1–11, Dec. 2018, doi: 10.1007/s11245-018-9626-5.

[93] D. Grossi and G. Pigozzi, *Judgment Aggregation: A Primer*. San Rafael, CA, USA: Morgan & Claypool, 2014.

[94] P. Jain, N. Talmon, and L. Bulteau, "Partition aggregation for participatory budgeting," in *20th Int. Conf. Auton. Agents and Multiagent Syst.*, F. Dignum, A. Lomuscio, U. Endriss, and A. Nowé, Eds., May 2021, pp. 665–673.

[95] G. Chalkiadakis, E. Elkind, and M. Wooldridge, *Computational Aspects of Cooperative Game Theory*. San Rafael, CA, USA: Morgan & Claypool, 2011.

[96] E. Elkind, D. Grossi, E. Shapiro, and N. Talmon, "United for change: Deliberative coalition formation to change the status quo," in *35th AAAI Conf. Artif. Intell.*, Feb. 2021, pp. 5339–5346.

[97] A. Darmann, "Group activity selection from ordinal preferences," in *Algorithmic Decis. Theory - 4th Int. Conf.*, vol. 9346, T. Walsh, Ed., Lexington, KY, USA, Sep. 2015, pp. 35–51, doi: 10.1007/978-3-319-23114-3_3.

[98] A. Darmann, E. Elkind, S. Kurz, J. Lang, J. Schauer, and G. J. Woeginger, "Group activity selection problem with approval preferences," *Int. J. Game Theory*, vol. 47, no. 3, pp. 767–796, Oct. 2018, doi: 10.1007/s00182-017-0596-4.

[99] A. Darmann, "A social choice approach to ordinal group activity selection," *Math. Soc. Sci.*, vol. 93, no. 7, pp. 57–66, May 2018, doi: 10.1016/j.mathsocsci.2018.01.005.

[100] B. Swierczek, "Democratic file revision control with liquidfeedback," *The Liquid Democracy J.*, vol. 1, no. 4, pp. 8–41, Jul. 2015.

[101] S. Zhou, B. Vasilescu, and C. Kästner, "How has forking changed in the last 20 years?: a study of hard forks on github," in *42nd Int. Conf. Softw. Eng.*, G. Rothermel and D. Bae, Eds., Seoul, South Korea, Jun. 2020, pp. 445–456, doi: 10.1145/3377811.3380412.

[102] J. C. Miller, "A program for direct and proxy voting in the legislative process," *Public Choice*, vol. 7, no. 1, pp. 107–113, 1969.

[103] G. Tullock, *Toward a Mathematics of Politics*. Ann Arbor, MI, USA: Univ. Michigan Press, 1967.

[104] M. Shubik, "On homo politicus and the instant referendum," *Public Choice*, vol. 9, pp. 79–84, 1970.

[105] G. Tullock, "Computerizing politics," *Math. and Comput. Model.*, vol. 16, no. 8-9, pp. 59–65, Aug./Sep. 1992, doi: 10.1016/0895-7177(92)90087-2.

[106] D. Alger, "Voting by proxy," *Public Choice*, vol. 126, no. 1-2, pp. 1–26, Jan. 2006, doi: 10.1007/s11127-006-3059-1.

[107] G. Cohensius, S. Mannor, R. Meir, E. A. Meirom, and A. Orda, "Proxy voting for better outcomes," in *Proc. 16th Conf. Auton. Agents and Multi-agent Syst.*, K. Larson, M. Winikoff, S. Das, and E. H. Durfee, Eds., São Paulo, Brazil, May 2017, pp. 858–866.

[108] M. Brill and N. Talmon, "Pairwise liquid democracy," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, J. Lang, Ed., Stockholm, Sweden, Jul. 2018, pp. 137–143, doi: 10.24963/ijcai.2018/19.

[109] A. Kahng, S. Mackenzie, and A. D. Procaccia, "Liquid democracy: An algorithmic perspective," *J. Artif. Intell. Res.*, vol. 70, no. 30, pp. 1223–1252, Mar. 2021.

[110] D. Bloembergen, D. Grossi, and M. Lackner, "On rational delegations in liquid democracy," in *The 33rd AAAI Conf. Artif. Intell.*, Honolulu, HI, USA, Jan./Feb. 2019, pp. 1796–1803, doi: 10.1609/aaai.v33i01.33011796.

[111] Z. Christoff and D. Grossi, "Binary voting with delegable proxy: An analysis of liquid democracy," in *Proc. 16th Conf. Theor. Aspects of Rationality and Knowl.*, vol. 251, J. Lang, Ed., Liverpool, UK, Jul. 2017, pp. 134–150, doi: 10.4204/EPTCS.251.10.

[112] R. Colley, "Multi-agent ranked delegations in voting," in *20th Int. Conf. Auton. Agents and Multi-agent Syst.*, F. Dignum, A. Lomuscio, U. Endriss, and A. Nowé, Eds., May 2021, pp. 1802–1804.

[113] R. Colley, U. Grandi, and A. Novaro, "Smart voting," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, C. Bessiere, Ed., 2020, pp. 1734–1740, doi: 10.24963/ijcai.2020/240.

[114] J. Green-Armytage, "Direct voting and proxy voting," *Constitutional Political Econ.*, vol. 26, no. 2, pp. 190–220, Jun. 2015, doi: 10.1007/s10602-014-9176-9.

[115] C. Blum and C. I. Zuber, "Liquid democracy: Potentials, problems, and perspectives," *J. Political Philos.*, vol. 24, no. 2, pp. 162–182, Jun. 2016, doi: 10.1111/jopp.12065.

[116] M. Brill, "Interactive democracy," in *Proc. 17th Int. Conf. Auton. Agents and Multiagent Syst.*, E. André, S. Koenig, M. Dastani, and G. Sukthankar, Eds., Stockholm, Sweden, Jul. 2018, pp. 1183–1187.

[117] S. Hardt and L. C. Lopes, "Google votes: A liquid democracy experiment on a corporate social network," *Tech. Discl. Commons*, no. 79, Jun. 2015, https://www.tdcommons.org/dpubs_series/79/ (Retrieved on 7/18/2021).

[118] B. A. Ford, "Delegative democracy," 2002, unpublished. https://bford.info/deleg/deleg.pdf (Retrieved on 6/16/2021).

[119] B. Escoffier, H. Gilbert, and A. Pass-Lanneau, "The convergence of iterative delegations in liquid democracy in a social network," in *Algorithmic Game Theory - Proc. 12th Int. Symp.*, vol. 11801, D. Fotakis and E. Markakis, Eds., Athens, Greece, Sep./Oct. 2019, pp. 284–297, doi: 10.1007/978-3-030-30473-7_19.

[120] J. Harding, "Incorporating preference information into formal models of transitive proxy voting", M.S. thesis, Univ. Amsterdam, London, UK, 2019.

[121] J. Harding, "Proxy selection in transitive proxy voting," *Soc. Choice Welf.*, vol. 56, no. 5, Jun. 2021, doi: 10.1007/s00355-021-01345-8.

[122] Y. Zhang and D. Grossi, "Power in liquid democracy," in *35th AAAI Conf. Artif. Intell.*, Feb. 2021, pp. 5822–5830.

[123] P. Boldi, F. Bonchi, C. Castillo, and S. Vigna, "Voting in social networks," in *Proc. 18th ACM Conf. Inf. and Knowl. Manage.*, New York, NY, USA, 2009, pp. 777–786, doi: 10.1145/1645953.1646052.

[124] I. Caragiannis and E. Micha, "A contribution to the critique of liquid democracy," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, S. Kraus, Ed., Macao, China, Aug. 2019, pp. 116–122, doi: 10.24963/ijcai.2019/17.

[125] P. Boldi, F. Bonchi, C. Castillo, and S. Vigna, "Viscous democracy for social networks," *Commun. ACM*, vol. 54, no. 6, pp. 129–137, Jun. 2011, doi: 10.1145/1953122.1953154.

[126] P. Boldi, C. Monti, M. Santini, and S. Vigna, "Liquid FM: Recommending music through viscous democracy," in *Proc. 6th Italian Inf. Retrieval Workshop*, vol. 1404, P. Boldi, R. Perego, and F. Sebastiani, Eds., Cagliari, Italy, May 2015.

[127] B. Zhang and H. Zhou, "Brief announcement: Statement voting and liquid democracy," in *Proc. ACM Symp. Princ. Distrib. Comput.*, E. M. Schiller and A. A. Schwarzmann, Eds., Washington, DC, USA, Jul. 2017, pp. 359–361, doi: 10.1145/3087801.3087868.

[128] Y. Zhang and D. Grossi, "Tracking truth by weighting proxies in liquid democracy," Mar. 2021, arXiv:abs/2103.09081. (Retrieved on 6/16/2021).

[129] J. Behrens, A. Kistner, A. Nitsche, and B. Swierczek, *The Principles of LiquidFeedback*, 1st ed. Berlin, Germany: Interacktive Demokratie, 2014.

[130] A. Paulin, "An overview of ten years of liquid democracy research," in *The 21st Annu. Int. Conf. Digit. Government Res.*, S. Eom and J. Lee, Eds., Seoul, South Korea, Jun. 2020, pp. 116–121, doi: 10.1145/3396956.3396963.

[131] U. Endriss and U. Grandi, "Binary aggregation by selection of the most representative voters," in *Proc. 28th AAAI Conf. Artif. Intell.*, C. E. Brodley and P. Stone, Eds., Québec City, Québec, Canada, Jul. 2014, pp. 668–674.

[132] S. Nitzan and J. Paroush, "Collective decision making and jury theorems," *Oxford Handbook of Law and Econ.*, vol. 1, pp. 494–516, Apr. 2017, doi: 10.1093/oxfordhb/9780199684267.013.035.

[133] B. Liu, Y. Shen, Y. Chen, X. Chen, and Y. Wang, "A two-layer weight determination method for complex multi-attribute large-group decision-making experts in a linguistic environment," *Inf. Fusion*, vol. 23, pp. 156 – 165, May 2015, doi: 10.1016/j.inffus.2014.05.001.

[134] W. T. H. Koh, "Optimal sequential decision architectures and the robustness of hierarchies and polyarchies," *Soc. Choice Welf.*, vol. 24, no. 3, pp. 397–411, Jun. 2005, doi: 10.1007/s00355-003-0304-0.

[135] G. Valentini, H. Hamann, and M. Dorigo, "Self-organized collective decision making: The weighted voter model," in *Proc. 2014 Int. Conf. Auton. Agents and Multi-agent Syst.*, 2014, pp. 45–52.

[136] A. Reina, M. Dorigo, and V. Trianni, "Collective decision making in distributed systems inspired by honeybees behaviour," in *Proc. 2014 Int. Conf. Auton. Agents and Multi-agent Syst.*, 2014, pp. 1421–1422.

[137] B. Wu, X. Zhou, Q. Jin, F. Lin, and H. Leung, "Analyzing social roles based on a hierarchical model and data mining for collective decision-making support," *IEEE Syst. J.*, vol. 11, no. 1, pp. 356–365, Mar. 2017, doi: 10.1109/JSYST.2014.2386611.

[138] I. Caragiannis, A. D. Procaccia, and N. Shah, "When do noisy votes reveal the truth?" *ACM Trans. Econ. and Comput.*, vol. 4, no. 3, pp. 15:1–15:30, Jun. 2016, doi: 10.1145/2892565.

[139] R. C. Ben-Yashar and S. I. Nitzan, "The optimal decision rule for fixed-size committees in dichotomous choice situations: The general result," *Int. Econ. Rev.*, vol. 38, no. 1, pp. 175–186, Feb. 1997, doi: 10.2307/2527413.

[140] E. Baharad, J. Goldberger, M. Koppel, and S. Nitzan, "Distilling the wisdom of crowds: weighted aggregation of decisions on multiple issues," *Auton. Agents and Multi-agent Syst.*, vol. 22, no. 1, pp. 31–42, Dec. 2011, doi: 10.1007/s10458-009-9120-y.

[141] ——, "Beyond Condorcet: Optimal aggregation rules using voting records," *Theory and Decis.*, vol. 72, no. 1, pp. 113–130, Jan. 2012, doi: 10.1007/s11238-010-9240-5.

[142] B. Grofman and S. L. Feld, "Determining optimal weights for expert judgment," in *Inf. Pooling and Group Decis. Making*, 1983, pp. 167–172.

[143] S. Nitzan and J. Paroush, "Optimal decision rules in uncertain dichotomous choice situations," *Int. Econ. Rev.*, vol. 23, no. 2, pp. 289–297, Jun. 1982, doi: 10.2307/2526438.

[144] D. Berend and J. E. Harmse, "Expert rule versus majority rule under partial information," *Theory and Decis.*, vol. 35, no. 2, pp. 179–197, Sep. 1993.

[145] B. Grofman, G. Owen, and S. L. Feld, "Thirteen theorems in search of the truth," *Theory and Decis.*, vol. 15, no. 3, pp. 261–278, Sep. 1983.

[146] M. Fey, "A note on the condorcet jury theorem with supermajority voting rules," *Soc. Choice Welf.*, vol. 20, no. 1, pp. 27–32, Jul. 2003, doi: 10.1007/s003550200157.

[147] S. Nitzan and J. Paroush, "A general theorem and eight corollaries in search of correct decision," *Theory and Decis.*, vol. 17, no. 3, pp. 211–220, Nov. 1994, doi: 10.1007/BF00132608.

[148] L. Shapley and B. Grofman, "Optimizing group judgmental accuracy in the presence of interdependencies," *Public Choice*, vol. 43, no. 3, pp. 329–343, Jan. 1984, doi: 10.1007/BF00118940.

[149] S. L. Feld and B. Grofman, "The accuracy of group majority decisions in groups with added members," *Public Choice*, vol. 42, no. 3, pp. 273–285, Jan. 1984, doi: 10.1007/BF00124946.

[150] J. H. Smith, "Aggregation of preferences with variable electorate," *Econometrica: J. Econometric Soc.*, vol. 41, no. 6, pp. 1027–1041, Nov. 1973, doi: 10.2307/1914033.

[151] J. Paroush and D. Karotkin, "Robustness of optimal majority rules over teams with changing size," *Soc. Choice Welf.*, vol. 6, no. 2, pp. 127–138, May 1989.

[152] E. Auriol and R. J. Gary-Bobo, "On the optimal number of representatives," *Public Choice*, vol. 153, no. 3-4, pp. 419–445, Dec. 2012, doi: 10.1007/s11127-011-9801-3.

[153] D. Karotkin and J. Paroush, "Optimum committee size: Quality-versus-quantity dilemma," *Soc. Choice Welf.*, vol. 20, no. 3, pp. 429–441, 2003, doi: 10.1007/s003550200190.

[154] M. Magdon-Ismail and L. Xia, "A mathematical model for optimal decisions in A representative democracy," in *31st Annu. Conf. Neural Inf. Process. Syst.*, S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., Montréal, Canada, Dec. 2018, pp. 4707–4716.

[155] G. D'Angelo, E. Delfaraz, and H. Gilbert, "Computation and bribery of voting power in delegative simple games," Apr. 2021, arXiv:abs/2104.03692 (Retrieved on 6/16/2021).

[156] M. A. Rodriguez and D. J. Steinbock, "A social network for societal-scale decision-making systems," Dec. 2004, arXiv:abs/cs/0412047 (Retrieved on 6/16/2021).

[157] P. K. Skowron, "What do we elect committees for? A voting committee model for multi-winner rules," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, Q. Yang and M. J. Wooldridge, Eds., Buenos Aires, Argentina, Jul. 2015, pp. 1141–1148.

[158] R. Meir, F. Sandomirskiy, and M. Tennenholtz, "Representative committees of peers," Jun. 2020, arXiv:abs/2006.07837. (Retrieved on 6/16/2021).

[159] M. Pivato and A. Soh, "Weighted representative democracy," *J. Math. Econ.*, vol. 88, pp. 52–63, May 2020, doi: 10.1016/j.jmateco.2020.03.001.

[160] A. C. Soh Voutsa, "Approval voting & majority judgement in weighted representative democracy," Jan. 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3731754 (Retrieved on 7/20/2021).

[161] L. S. Penrose, "The elementary statistics of majority voting," *J. Roy. Statist. Soc.*, vol. 109, no. 1, pp. 53–57, 1946.

[162] K. O. May, "A set of independent necessary and sufficient conditions for simple majority decision," *Econometrica*, vol. 20, no. 4, pp. 680–684, Oct. 1952, doi: 10.2307/1907651.

[163] A. Paulin, "Through liquid democracy to sustainable non-bureaucratic government," in *Proc. Int. Conf. for E-democracy and Open Government*, vol. 6, no. 2, Nov. 2014, pp. 205–217, doi: 10.29379/jedem.v6i2.307.

[164] J. Ramos, "Liquid democracy and the futures of governance," in *The Future Internet*, ser. Public Admin. and Inf. Technol., vol. 17, J. Winter and R. Ono, Eds. New York, NY, USA: Springer, 2015, pp. 173–191.

[165] Z. Christoff and D. Grossi, "Liquid democracy: An analysis in binary aggregation and diffusion," Dec. 2016, arXiv:abs/1612.08048.

[166] S. Bastgen and K. Winkler, "Liquid democracy–participation model of the 21st century?" in *The Governance of Large-Scale Projects: Linking Citizens and the State*, vol. 5, A. Römmele and H. Schober, Eds.,. Baden-Baden Germany: Nomos Verlagsgesellschaft mbH & Co. KG, 2013, pp. 111–132, doi: 10.5771/9783845243566-111.

[167] P. Gölz, A. Kahng, S. Mackenzie, and A. D. Procaccia, "The fluid mechanics of liquid democracy," in *Web and Internet Econ. - Proc. 14th Int. Conf.*, vol. 11316, G. Christodoulou and T. Harks, Eds., Oxford, UK, Dec. 2018, pp. 188–202, doi: 10.1007/978-3-030-04612-5_13.

[168] M. R. Garey and D. S. Johnson, *Computers and Intractability, A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1979.

[169] U. Feige, "A threshold of ln n for approximating set cover," *J. ACM*, vol. 45, no. 4, pp. 634–652, Jul. 1998, doi: 10.1145/285055.285059.

[170] G. E. M. Anscombe, "On frustration of the majority by fulfillment of the majority's will," *Anal.*, vol. 36, no. 4, pp. 161–168, Jun. 1976, doi: 10.2307/3327179.

[171] L. S. Shapley and M. Shubik, "A method for evaluating the distribution of power in a committee system," *Amer. Political Sci. Rev.*, vol. 48, no. 3, pp. 787–792, Sep. 1954, doi: 10.2307/1951053.

[172] J. F. Banzhaf III, "Weighted voting doesn't work: A mathematical analysis," *Rutgers Law Rev.*, vol. 19, pp. 317–343, 1965.

[173] J.-J. Rousseau, *Rousseau: The Social Contract and Other Later Political Writings*. Cambridge, UK: Cambridge Univ. Press, 2018.

[174] E. Dekel, M. O. Jackson, and A. Wolinsky, "Vote buying: General elections," *J. Political Econ.*, vol. 116, no. 2, pp. 351–380, Apr. 2008, doi: 10.1086/587624.

[175] S. P. Lalley and E. G. Weyl, "Quadratic voting: How mechanism design can radicalize democracy," in *Papers and Proc. 130th Annu. Meeting Amer. Econ. Assoc.*, vol. 108, May 2018, pp. 33–37, doi: 10.2307/26452700.

[176] E. A. Posner and E. G. Weyl, "Voting squared: Quadratic voting in democratic politics," *Vanderbilt Law Rev.*, vol. 68, no. 2, pp. 441–500, Mar. 2015.

[177] ——, "Quadratic voting as efficient corporate governance," *Univ. Chicago Law Rev.*, vol. 81, no. 1, pp. 251–272, Jan. 2014.

[178] E. G. Weyl, "The robustness of quadratic voting," *Public Choice*, vol. 172, no. 1-2, pp. 75–107, Jul. 2017, doi: 10.1007/s11127-017-0405-4.

[179] E. A. Posner and E. G. Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton, NJ, USA: Princeton Univ. Press, 2018, doi: 10.1515/9780691196978.

[180] S. D. Kominers, "Good markets (really do) make good neighbors," *ACM SIGecom Exchanges*, vol. 16, no. 2, pp. 12–26, May 2019, doi: 10.1145/3331041.3331044.

[181] R. Meir, J. Lang, J. Lesca, N. Mattei, and N. Kaminsky, "A market-inspired bidding scheme for peer review paper assignment," in *35th AAAI Conf. Artif. Intell.*, vol. 35, no. 6, May 2021, pp. 4776–4784.

[182] M. J. Sandel, *What Money Can't Buy: The Moral Limits of Markets*. New York, NY, USA: Macmillan, 2012.

[183] A. D. Procaccia and M. Tennenholtz, "Approximate mechanism design without money," *ACM Trans. Econ. and Comput.*, vol. 1, no. 4, pp. 1–26, Dec. 2013.

[184] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge, UK: Cambridge Univ. Press, 2007.

[185] D. C. Mueller, *Public Choice III*. Cambridge, UK: Cambridge Univ. Press, 2003.

[186] Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet, "A short introduction to computational social choice," in *Int. Conf. Current Trends in Theory and Pract. Comput. Sci.*, vol. 4362, 2007, pp. 51–69, doi: 10.1007/978-3-540-69507-3_4.

[187] U. Endriss, *Trends in Computational Social Choice*. AI Access, 2017, https://dl.acm.org/doi/book/10.5555/3180776 (Retrieved on 7/20/2021).

[188] A. Downs, "An economic theory of political action in a democracy," *J. Political Econ.*, vol. 65, no. 2, pp. 135–150, Apr. 1957, doi: 10.1086/257897.

[189] J. A. Ferejohn and M. P. Fiorina, "The paradox of not voting: A decision theoretic analysis," *Amer. Political Sci. Rev.*, vol. 68, no. 2, pp. 525–536, Jun. 1974, doi: 10.1017/S0003055400117368.

[190] K. J. Arrow, A. Sen, and K. Suzumura, *Handbook of Social Choice Welfare*. vol. 2, New York, NY, USA: Elsevier, 2010.

[191] J. S. Kelly, *Social Choice Theory*. Berling, Germany: Springer, 1988, doi: 10.1007/978-3-662-09925-4.

[192] L. Hurwicz, "The design of mechanisms for resource allocation," *The Amer. Econ. Rev.*, vol. 63, no. 2, pp. 1–30, May 1973.

[193] E. S. Maskin, "Mechanism design: How to implement social goals," *Amer. Econ. Rev.*, vol. 98, no. 3, pp. 567–76, Jun. 2008, doi: 10.1257/aer.98.3.567.

[194] T. Suzuki and M. Horita, "How to order the alternatives, rules, and the rules to choose rules: When the endogenous procedural choice regresses," in *Int. Conf. Group Decis. and Negotiation*, Jun. 2015, pp. 47–59, doi: 10.1007/978-3-319-19515-5_4.

[195] F. Dietrich, "How to reach legitimate decisions when the procedure is controversial," *Soc. Choice Welf.*, vol. 24, no. 2, pp. 363–393, Apr. 2005, doi: 10.1007/s00355-003-0288-9.

[196] M. Messner and M. K. Polborn, "Voting on majority rules," *The Rev. of Econ. Stud.*, vol. 71, no. 1, pp. 115–132, Jan. 2004.

[197] P. C. Fishburn, "Axioms for approval voting: Direct proof," *J. Econ. Theory*, vol. 19, no. 1, pp. 180–185, Oct. 1978, doi: 10.1016/0022-0531(78)90062-5.

[198] ——, "Condorcet social choice functions," *SIAM J. Appl. Math.*, vol. 33, no. 3, pp. 469–489, Sep. 1977, doi: 10.1137/0133030.

[199] F. Brandt, M. Brill, and P. Harrenstein, "Extending tournament solutions," *Soc. Choice Welf.*, vol. 51, no. 2, pp. 193–222, Jan. 2018, doi: 10.1007/s00355-018-1112-x.

[200] G. Shahaf, E. Shapiro, and N. Talmon, "Sybil-resilient reality-aware social choice," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, S. Kraus, Ed., Macao, China, Aug. 2019, pp. 572–579, doi: 10.24963/ijcai.2019/81.

[201] H. Nurmi, "The choice of voting rules based on preferences over criteria," in *Int. Conf. Group Decis. and Negotiation*, vol. 218, 2015, pp. 241–252, doi: 10.1007/978-3-319-19515-5_19.

[202] K. Binmore, "An example in group preference," *J. Econ. Theory*, vol. 10, no. 3, pp. 377–385, Jun. 1975.

[203] S. Koray, "Self-selective social choice functions verify arrow and gibbard-satterthwaite theorems," *Econometrica*, vol. 68, no. 4, pp. 981–996, Jul. 2000, doi: 10.1111/1468-0262.00143.

[204] S. Koray and A. Slinko, "Self-selective social choice functions," *Soc. Choice Welf.*, vol. 31, no. 1, pp. 129–149, Sep. 2007.

[205] D. Acemoglu, G. Egorov, and K. Sonin, "Dynamics and stability of constitutions, coalitions, and clubs," *Amer. Econ. Rev.*, vol. 102, no. 4, pp. 1446–76, Jun. 2012.

[206] J. L. García-Lapresta and A. Piggins, "Voting on how to vote," *Soft Computing in Humanities and Social Science*, Berlin, Germany: Springer, 2012, ch. 17, pp. 301–321.

[207] N. Houy, "A note on the impossibility of a set of constitutions stable at different levels," Univ. Panthéon-Sorbonne (Paris 1), Paris, France, Tech. Rep. v04039, 2004.

[208] D. W. Rae, "Decision-rules and individual values in constitutional choice," *Amer. Political Sci. Rev.*, vol. 63, no. 1, pp. 40–56, Mar. 1969, doi: 10.1017/S0003055400261467.

[209] M. Bhattacharya, "Constitutionally consistent voting rules over single-peaked domains," *Soc. Choice Welf.*, vol. 52, no. 2, pp. 225–246, Feb. 2019, doi: 10.1007/s00355-018-1145-1.

[210] S. Barbera and M. O. Jackson, "Choosing how to choose: Self-stable majority rules and constitutions," *The Quart. J. Econ.*, vol. 119, no. 3, pp. 1011–1048, Aug. 2004, doi: 10.1162/0033553041502207.

[211] A. D. Procaccia, "Axioms should explain solutions," in *The Future of Economic Design*, J.-F. Laslier, H. Moulin, M. R. Sanver, and W. S. Zwicker, Eds. New York, NY, USA: Springer, 2019, pp. 195–199, doi: 10.1007/978-3-030-18050-8_27.

[212] A. Boixel and U. Endriss, "Automated justification of collective decisions via constraint solving," in *Proc. 19th Int. Conf. Auton. Agents and Multi-agent Syst.*, 2020, pp. 168–176.

[213] A. Boixel, "Computer-aided reasoning about collective decision making," in *Proc. 19th Int. Conf. Auton. Agents and Multi-agent Syst.*, 2020, pp. 2176–2178.

[214] D. Peters, A. D. Procaccia, A. Psomas, and Z. Zhou, "Explainable voting," in *33rd Annu. Conf. Neural Inf. Process. Syst.*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., Dec. 2020, pp. 1525–1534.

[215] A. Boixel and R. de Haan, "On the complexity of finding justifications for collective decisions," in *35th AAAI Conf. Artif. Intell.*, vol. 35, no. 6, May 2021, pp. 5194–5201.

[216] O. Cailloux and U. Endriss, "Arguing about voting rules," in *Proc. Int. Conf. Auton. Agents & Multiagent Syst.*, C. M. Jonker, S. Marsella, J. Thangarajah, and K. Tuyls, Eds., Singapore, May 2016, pp. 287–295.

[217] T. Dodson, N. Mattei, J. T. Guerin, and J. Goldsmith, "An English-language argumentation interface for explanation generation with Markov decision processes in the domain of academic advising," *ACM Trans. Interactive Intell. Syst.*, vol. 3, no. 3, pp. 1–30, Oct. 2013, doi: 10.1145/2513564.

[218] D. Baumeister, D. Neugebauer, and J. Rothe, "Collective acceptability in abstract argumentation," *J. Appl. Log.*, vol. 8, no. 6, pp. 1503–1542, Jul. 2021.

[219] R. Baumann, S. Doutre, J.-G. Mailly, and J. P. Wallner, "Enforcement in formal argumentation," *J. Appl. Log.*, vol. 8, no. 6, pp. 1623–1677, Jul. 2021.

[220] M. Caminada and G. Pigozzi, "On judgment aggregation in abstract argumentation," *Auton. Agents and Multi-agent Syst.*, vol. 22, no. 1, pp. 64–102, Nov. 2011, doi: 10.1007/s10458-009-9116-7.

[221] E. Awad, R. Booth, F. Tohmé, and I. Rahwan, "Judgement aggregation in multi-agent argumentation," *J. Log. and Comput.*, vol. 27, no. 1, pp. 227–259, Feb. 2017, doi: 10.1093/logcom/exv055.

[222] F. Dietrich and C. List, "Judgment aggregation by quota rules: Majority voting generalized," *J. Theor. Politics*, vol. 19, no. 4, pp. 391–424, Oct. 2007, doi: 10.1177/0951629807080775.

[223] C. List and P. Pettit, "Aggregating sets of judgments: An impossibility result," *Econ. & Philos.*, vol. 18, no. 1, pp. 89–110, Mar. 2002, doi: 10.1017/S0266267102001098.

[224] T. J. M. Bench-Capon, "Persuasion in practical argument using value-based argumentation frameworks," *J. Log. and Comput.*, vol. 13, no. 3, pp. 429–448, Jun. 2003, doi: 10.1093/logcom/13.3.429.

[225] J. F. Banzhaf III, "One man, ? votes: Mathematical analysis of voting power and effective representation," *George Washington Law Rev.*, vol. 36, no. 4, pp. 808–823, May 1967.

[226] Y. Zick, A. Skopalik, and E. Elkind, "The shapley value as a function of the quota in weighted voting games," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, T. Walsh, Ed., Barcelona, Spain, Jul. 2011, pp. 490–496.

[227] P. Faliszewski and L. A. Hemaspaandra, "The complexity of power-index comparison," *Theor. Comput. Sci.*, vol. 410, no. 1, pp. 101–107, Jan. 2009, doi: 10.1016/j.tcs.2008.09.034.

[228] J. Y. Halpern, *Actual Causality*. Cambridge, MA, USA: MIT Press, 2016.

[229] D. Lewis, "Causation as influence," *J. Philos.*, vol. 97, no. 4, pp. 182–197, Apr. 2004, doi: 10.2307/2678389.

[230] ——, *Counterfactuals*. Hoboken, NJ, USA: Wiley, 2013.

[231] D. Black, *The Theory of Committees and Elections*. New York, NY, USA: Springer, 1958.

[232] H. Moulin, "On strategy-proofness and single peakedness," *Public Choice*, vol. 35, no. 4, pp. 437–455, Jan. 1980.

[233] S. Barbera, "An introduction to strategy-proof social choice functions," *Soc. Choice Welf.*, vol. 18, no. 4, pp. 619–653, Oct. 2001.

[234] X. Sui, C. Boutilier, and T. Sandholm, "Analysis and optimization of multi-dimensional percentile mechanisms," in *23rd Int. Joint Conf. Artif. Intell.*, Aug. 2013, pp. 367–374.

APPENDIX A
Springer Nature License

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Jul 14, 2021

This Agreement between Ben Abramowitz ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5100310890433
License date	Jul 01, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	Awareness of Voter Passion Greatly Improves the Distortion of Metric Social Choice
Licensed Content Author	Ben Abramowitz, Elliot Anshelevich, Wennan Zhu
Licensed Content Date	Jan 1, 2019
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print
Portion	full article/chapter

Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	yes
Title	Deciding Who, What, Why, and How: Aggregating Preferences Over Agents, Alternatives, Axioms, and Rules
Institution name	Rensselaer Polytechnic Institute
Expected presentation date	Jul 2021
Requestor Location	Ben Abramowitz 110 8th St TROY, NY 12180 United States Attn: Ben Abramowitz
Total	0.00 USD

Terms and Conditions

Springer Nature Customer Service Centre GmbH
Terms and Conditions

This agreement sets out the terms and conditions of the licence (the **Licence**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

2. 3. Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

2. 4. Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

2. 5. An alternative scope of licence may apply to signatories of the STM Permissions Guidelines, as amended from time to time.

3. Duration of Licence

3. 1. A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

Scope of Licence	Duration of Licence
Post on a website	12 months
Presentations	12 months
Books and journals	Lifetime of the edition in the language purchased

4. Acknowledgement

4. 1. The Licensor's permission must be acknowledged next to the Licenced Material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

5. Restrictions on use

5. 1. Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

5. 2. You must not use any Licensed Material as part of any design or trademark.

5. 3. Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

6. Ownership of Rights

6. 1. Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

7. Warranty

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND
WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

8. Limitations

8. 1. ***BOOKS ONLY:***Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a

personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

8. 2. For content reuse requests that qualify for permission under the STM Permissions Guidelines, which may be updated from time to time, the STM Permissions Guidelines supersede the terms and conditions contained in this licence.

9. Termination and Cancellation

9. 1. Licences will expire after the period shown in Clause 3 (above).

9. 2. Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

Appendix 1 — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

Version 1.3

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.