Thesis Proposal
Social Choice and AI: Helping Humans Make Decisions

Anson Kahng

April 16, 2020

Abstract

This thesis studies how to help people make decisions by using tools from artificial intelligence and computational social choice. On a high level, this thesis can be broken down into two major thrusts. The first is the analysis and refinement of two new paradigms of democratic decision-making, virtual democracy and liquid democracy. The second is the design of new mechanisms that satisfy certain theoretical properties, namely strategyproofness and consensus.

Virtual democracy is a method for automating decision-making by learning models of people and letting these models vote in order to reduce cognitive and operational load on the electorate. However, in such a system, the question of how to aggregate the electorate's predicted votes remains. We prove that the Borda count is a provably robust rule for aggregating virtual votes, and show that pairwise-majority consistent rules are not robust in this setting. We then study a real-world implementation of virtual democracy in collaboration with a local nonprofit organization in Pittsburgh, 412 Food Rescue.

Liquid democracy, on the other hand, is a democratic paradigm where voters are allowed to transitively delegate their voting rights to other voters, who then may cast a vote of commensurately greater weight. We study liquid democracy from an algorithmic point of view and prove that, under certain assumptions about delegation behavior, the popular notion that liquid democracy will always be more accurate than direct democracy is false. We then focus to minimizing the maximum weight voter in liquid democracy systems, a problematic occurrence in both theory and practice.

Turning now to designing voting rules that satisfy certain desirable properties, we start with impartial peer assessment, where our goal is to design a mechanism by which a set of agents can provide input on relative rankings between themselves in order to produce an aggregate ranking. However, using many common voting rules would incentivize agents to misreport their true beliefs in order to manipulate their final ranking. We analyze impartial rules, or rules that satisfy the property that no agent's report affects the distribution of his final ranking, and prove accuracy guarantees on their performance. Additionally, we develop HirePeer, a novel alternative approach to hiring at scale that leverages peer assessment to evaluate worker qualities, and explore the accuracy and efficacy of impartial algorithms in online labor markets.

Finally, we lay out current and future work. First, we consider an ongoing project on designing voting rules in order to achieve consensus (i.e., compromise) between agents with heterogeneous beliefs. Second, we discuss an extension to the framework of virtual democracy. Finally, we discuss another project on designing and evaluating voting rules in order to ensure proportional representation in multiwinner elections.

Contents

1 Introduction

Humans must continually make difficult decisions throughout their lives, and, on the scale of states or nations, these decisions become incredibly complex due to the large number of heterogeneous opinions at play. Academics have studied multi-agent decision-making through the lens of social choice theory, which formalizes the problem of aggregating individual preferences in order to choose a collective outcome. However, as techniques and technologies develop, so too do the ways in which we study large-scale decisions. In particular, computational social choice is a recently developed field that applies tools from computer science to social choice theory [24]. Additionally, in recent years, the field of artificial intelligence has made notable strides toward systems capable of performing complex decision-making, and the combination of tools from computational social choice and artificial intelligence has the potential to drastically improve the way in which humans make decisions.

This thesis studies how to apply concepts from artificial intelligence and computational social choice to help people make principled decisions. While many researchers in the field of artificial intelligence use computational social choice as a tool for designing multi-agent systems, we apply tools from artificial intelligence to analyze and design social choice mechanisms in order to help people make decisions in a more principled and provably robust way.

On a high level, this thesis can be broken down into two major thrusts. The first is the analysis and refinement of two new paradigms of democratic decision-making, virtual democracy and liquid democracy. The second is the design of new mechanisms that satisfy certain theoretical properties, namely strategyproofness, representation, and consensus. Below, we present a short overview of our approach, results, and future directions for each project. Please note that, because this thesis proposal touches on projects in many different settings, we defer project-specific related work, models, and notation to the relevant section of the proposal.

General Background Computational social choice encompasses a large range of topics, including fair allocation (also known as fair division) [37, 73, 82] and matching problems [45, 75]. However, we focus on applying tools from voting theory in order to help large collectives of people make more general decisions, not necessarily about how to split resources among themselves or construct a matching from agents to resources. Throughout this thesis, we assume that voters provide ordinal (e.g., ranked ballots or binary approval votes) preferences over alternatives as input to voting rules, which then return either an output ranking or set of winners.

On a high level, there are three fundamental impossibility results in the field of voting theory: Condorcet's paradox, Arrow's Impossibility Theorem, and the Gibbard-Satterthwaite Theorem. Condorcet's paradox [35] states that, on a population level, majority judgments between pairs of alternatives may not be transitive; i.e., that there can exist settings where alternative a is preferred to alternative b by at least half of the voters, b is preferred to c by at least half of the voters, but over half of the voters prefer c to a, causing a cycle. Arrow's Impossibility Theorem [7] states that, under mild assumptions and in settings with at least three alternatives, every voting rule must either violate a property known as independence of irrelevant alternatives (IIA) or be a dictatorship. Finally, the Gibbard-Satterthwaite Theorem [46, 78] maintains that any non-dictatorship voting rule is not strategyproof; i.e., that agents can benefit by misreporting their preferences. These fundamental impossibility results mean that, in our work, we must restrict either the domain of (strategic) voter behavior or weaken the theoretical results we hope to prove.

1.1 Theoretical Foundations of Democratic Paradigms

One important aspect of this thesis is establishing the theoretical properties of new forms of democracy, which provide valuable insights about the further development of these paradigms. We apply this approach to two novel voting mechanisms, virtual democracy and liquid democracy.

Virtual Democracy (Section 2) *Virtual democracy* is an approach to automating decisions by learning models of the preferences of individual people, and, at runtime, aggregating the *predicted* preferences of those people on the dilemma at hand. One of the key questions is which aggregation method — or *voting rule* — to use; we offer a novel statistical viewpoint that provides guidance. Specifically, we seek voting rules that are *robust* to prediction errors, in that their output on people's true preferences is likely to coincide with their output on noisy estimates thereof. We prove that the classic Borda count rule is robust in this sense, whereas any voting rule belonging to the wide family of pairwise-majority consistent (PMC) rules is not [53].

In concert with our theoretical results, we have worked closely with a local nonprofit organization in Pittsburgh, 412 Food Rescue, to build a framework for virtual democracy that enables people to build algorithmic policy for their communities. Through this framework, we study how to design algorithmic policy in order to balance varying interests in a moral, legitimate way. Our findings suggest that the framework successfully enabled participants to build models that they felt confident represented their own beliefs. Participatory algorithm design also improved both procedural fairness and the distributive outcomes of the algorithm, raised participants' algorithmic awareness, and helped identify inconsistencies in human decision-making in the governing organization [59].

Liquid Democracy (Section 3) *Liquid democracy* is a collective decision making paradigm that allows voters to transitively delegate their votes. Our first work on this subject studied liquid democracy through an algorithmic lens [52]. In our model, there are two alternatives, one correct and one incorrect, and we are interested in the probability that the majority opinion is correct. Our main question is whether there exist delegation mechanisms that are *guaranteed* to outperform direct voting, in the sense of being always at least as likely, and sometimes more likely, to make a correct decision. Even though we assume that voters can only delegate their votes to better-informed voters, we show that *local* delegation mechanisms, which only take the local neighborhood of each voter as input (and, arguably, capture the spirit of liquid democracy), cannot provide the foregoing guarantee. By contrast, we design a non-local delegation mechanism that does provably outperform direct voting under mild assumptions about voters.

The above result corroborates a common critique of liquid democracy: often, a small subset of agents may gain massive influence. To address this, we propose to change the current practice by allowing agents to specify multiple delegation options instead of just one. Then, we seek to control the flow of votes in a way that balances influence as much as possible. Specifically, we analyze the problem of choosing delegations to approximately minimize the maximum number of votes entrusted to any agent, by drawing connections to the literature on confluent flow. We also introduce a random graph model for liquid democracy, and draw on prior work on the *power of choice*, which allows us to establish a doubly-exponential separation between the maximum weight of a voter in the case where each voter provides a single delegation and the maximum weight of a voter in the case where each voter provides even two possible delegations [47].

1.2 Designing Principled Voting Rules

While it is important to analyze and refine existing democratic paradigms in order to provide certain guarantees, it is also valuable to start with theoretical desiderata and design voting mechanisms that satisfy these axioms. Below, I outline two notable lines of inquiry along these lines, where we design and analyze strategyproof and consensus-based mechanisms.

Impartial Aggregation (Section 4) We study rank aggregation algorithms that take as input the opinions of players over their peers, represented as rankings, and output a social ordering of the players (which reflects, e.g., relative contribution to a project or fit for a job). To prevent strategic behavior, these algorithms must be *impartial*, i.e., players should not be able to influence their own position in the output ranking. We design several randomized algorithms that are impartial and closely emulate given (non-impartial) rank aggregation rules in a rigorous sense [51].

We complement these theoretical results with a study of impartial algorithms applied to online labor markets through *HirePeer*, a novel alternative approach to hiring at scale that leverages peer assessment to elicit honest assessments of fellow workers' job application materials, which it then aggregates using an impartial ranking algorithm [56]. Surprisingly, we find that applying peer assessment to online hiring—even without impartial ranking algorithms to remove conflicts of interest—was an accurate and pedagogically beneficial practice.

Consensus Mechanisms (Ongoing Work, Section 5.1) Incentivizing agents to reach consensus, or compromise by choosing an alternative that is acceptable in some formal sense to all agents, is an important problem in group decision-making. As evidenced by political deadlocks around the world, current voting systems do not adequately incentivize compromise. To this end, we study mechanisms that encourage agents to make concessions in order to naturally reach unanimous agreement in settings where agents have heterogeneous preferences.

Proportionality in Multiwinner Elections (Ongoing Work, Section 5.3) Finally, we analyze mechanisms for multiwinner elections, where a group of agents, or voters, selects a committee from a set of candidates based on the agents' preferences. In our setting, each agent expresses her preferences through an approval vote, where she designates a subset of candidates she approves for the committee, and all votes are then aggregated to select a winning committee from the pool of candidates.

The property we would like to satisfy is that of proportionality, which intuitively says that a $c \leq 1$ fraction of voters who agree on a c fraction of the alternatives should be able to guarantee themselves control over a c fraction of the committee. We propose a measure of proportionality for elections where the size of the committee is not fixed beforehand and hope to establish a clear separation between the theoretical guarantees of deterministic and randomized rules under this notion of proportionality.

2 Virtual Democracy

One of the most basic ideas underlying democracy is that complicated decisions can be made by asking a group of people to vote on the alternatives at hand. As a decision-making framework, this paradigm is versatile, because people can express a sensible opinion about a wide range of issues.

One of its seemingly inherent shortcomings, though, is that voters must take the time to cast a vote — hopefully an informed one — every time a new dilemma arises.

But what if we could *predict* the preferences of voters — instead of explicitly asking them each time — and then aggregate those predicted preferences to arrive at a decision? This is exactly the idea behind the work of Noothigattu et al. [69], who are motivated by the challenge of automating *ethical* decisions. Specifically, their approach consists of three steps: first, collect preferences from voters on example dilemmas; second, learn models of their preferences, which generalize to any (previously unseen) dilemma; and third, at runtime, use those models to predict the voters' preferences on the current dilemma, and aggregate the predicted preferences to reach a decision. The idea is that we would ideally like to consult the voters on each decision, but in order to automate those decisions we instead use the models that we have learned as a proxy for the flesh and blood voters. In other words, the models serve as virtual voters, which is why we refer to this paradigm as *virtual democracy*.

Since 2017, we have been building on this approach in a collaboration with a Pittsburgh-based non-profit, 412 Food Rescue, that provides on-demand food donation distribution services. The goal is to design and deploy an algorithm that would automatically make the decisions they most frequently face: given an incoming food donation, which recipient organization (such as a housing authority or food pantry) should receive it? The voters in our implementation are stakeholders: donors, recipients, volunteers (who pick up the food from the donor and deliver it to the recipient), and employees. We have collected roughly 100 pairwise comparisons from each voter, where in each comparison, the voter is provided information about the type of donation, as well as seven relevant features of the two alternatives that are being compared, e.g., the distance between donor and recipient, and when the recipient last received a donation. Using this data, we have learned a model of the preferences of each voter, which allows us to predict the voter's preference *ranking* over hundreds of recipients. And given a predicted ranking for each voter, we map them into a ranking over the alternatives by applying a *voting rule*.

While this implementation sounds simple enough, the choice of voting rule can have a major impact on the efficacy of the system. In fact, the question of which voting rule to employ is one of the central questions in computational social choice [24], and in social choice theory more broadly. A long tradition of impossibility results establishes that there are no perfect voting rules [6], so the answer, such as it is, is often context-dependent.

The central premise of this theoretical body of work is that, in the context of virtual democracy, certain statistical considerations should guide the choice of voting rule. Indeed, the voting rule inherently operates on noisy predictions of the voters' true preferences, yet one might hope that it would still output the same ranking as it would in the 'real' election based on the voters' true preferences (after all, this is the ideal that virtual democracy is trying to approximate). Our theoretical research question, therefore, is

> *... which voting rules have the property that their output on the true preferences is likely to coincide with their output on noisy estimates thereof?*

In addition to answering this theoretical question, we work with 412 Food Rescue in order to study the effects of virtual democracy as a tool for algorithmic governance. Our work makes three contributions. First, we offer a framework and tools that enable participatory algorithm design, contributing to emerging research on human-centered algorithms and participatory design for technology. Second, through a case study with stakeholders of a real-world nonprofit, we demonstrate the feasibility, potential, and challenges of community involvement in algorithm design. Finally, our work provides insights on the effects of procedurally-fair algorithms and tools that can further understanding of algorithmic fairness and moral expectations.

2.1 Related Work

A number of recent papers have explored the idea of automating ethical decisions via machine learning and social choice [33, 43, 69]. As mentioned above, our work builds on the framework proposed by Noothigattu et al. [69]. However, it is important to clarify why the questions we explore here do not arise in their work. Since they deal with 1.3 million voters, and split-second decisions (what should a self-driving car do in an emergency?), they cannot afford to consult the individual voter models at runtime. Hence, they have added an additional *summarization* step, whereby the individual voter models are summarized as a single, concise model of societal preferences (with possibly significant loss to accuracy). The structure of the summary model is such that, for any given set of alternatives, almost all reasonable voting rules agree on the outcome (this is their main theoretical result), hence the choice of voting rule is a nonissue under that particular implementation. By contrast, our work is motivated by the food bank application of the virtual democracy framework, where the number of voters is small and speed is not of the essence, hence we predict the preferences of individual voters at runtime.

It is worth mentioning that another prominent approach to the allocation of food donations is based on (online) fair division [3]. That said, it is important to emphasize that we study a general question about the foundations of the virtual democracy paradigm, that is, our work is not technically tied to any particular application.

Furthermore, the Mallows model underlies a large body of work in computational social choice [9, 10, 11, 27, 28, 31, 32, 38, 39, 50, 60, 64, 71, 87, 88, 89]. Our model is loosely related to that of Jiang et al. [50], where individual rankings are derived from a single ground truth ranking via a Mallows model, and then a second Mallows model is applied to obtain a noisy version of each voter's ranking. Our technical question is completely different from theirs.

Finally, there is a large body of work in social choice on finding aggregation rules that satisfy axiomatic properties that formally capture notions of fairness or efficiency [6, 84]. However, many common axiomatic properties in social choice do not apply to standard applications of virtual democracy, including the autonomous vehicle domain of Noothigattu et al. [69] and our setting of food rescue, although they may be relevant in other differently-constrained domains.

2.2 Theoretical Foundations of Virtual Democracy

Preliminaries

We deal with a set of alternatives A such that $|A| = m$. Preferences over A are represented via a ranking $\sigma \in \mathcal{L}$, where $\mathcal{L} = \mathcal{L}(A)$ is the set of rankings (or permutations) over A. We denote by $\sigma(j)$ the alternative ranked in position j in σ, where position 1 is the highest, and m the lowest. We denote by $\sigma^{-1}(x)$ the position in which $x \in A$ is ranked. We use $x \succ_\sigma y$ to denote that x is preferred to y according to σ, i.e., that $\sigma^{-1}(x) < \sigma^{-1}(y)$.

The setting also includes a set of voters $N = \{1, \ldots, n\}$. Each voter $i \in N$ is associated with a ranking $\sigma_i \in \mathcal{L}$. The preferences of N are represented as a *preference profile* $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n) \in \mathcal{L}^n$.

Given a preference profile $\boldsymbol{\sigma} \in \mathcal{L}^n$, we say that $x \in A$ *beats* $y \in A$ *in a pairwise comparison* if a majority of voters prefer x to y, that is,

$$|\{i \in N : \ x \succ_{\sigma_i} y\}| > n/2.$$

The profile $\boldsymbol{\sigma}$ induces a *weighted pairwise majority graph* $\Gamma(\boldsymbol{\sigma})$, where we have a vertex for each alternative in A. For each $x \in A$ and $y \in A \setminus \{x\}$, there is an edge from x to y if x beats y in a pairwise comparison; the weight on this edge is

$$w_{(x,y)}(\boldsymbol{\sigma}) \triangleq |\{i \in N : \ x \succ_{\sigma_i} y\}| - |\{i \in N : \ y \succ_{\sigma_i} x\}|.$$

Voting Rules A *voting rule* (formally known as a *social welfare function*) is a function $f : \mathcal{L}^n \to \mathcal{L}$, which receives a preference profile as input, and returns a 'consensus' ranking of the alternatives. We are especially interested in two families of voting rules.

- *Positional scoring rules.* Each such rule is defined by a score vector $(\alpha_1, \ldots, \alpha_m)$. Given a preference profile $\boldsymbol{\sigma}$, the score of alternative x is

  $$\sum_{i=1}^{n} \alpha_{\sigma_i^{-1}(x)}.$$

  In words, each voter who ranks x in position p gives α_p points to x. The positional scoring rule returns a ranking of the alternatives by non-increasing score, with ties broken arbitrarily.

  Our main positive result pertains to the classic *Borda count* voting rule, which is the positional scoring rule defined by the score vector $(m-1, m-2, \ldots, 0)$. Denote the Borda count score of $x \in A$ in $\boldsymbol{\sigma} \in \mathcal{L}^n$ by

  $$B(x, \boldsymbol{\sigma}) \triangleq \sum_{i=1}^{n} \left(m - \sigma_i^{-1}(x)\right).$$

- *Pairwise-majority consistent (PMC) rules* [28]: These rules satisfy a fairly weak requirement that extends the classic notion of Condorcet consistent social *choice* functions: Given a profile $\boldsymbol{\sigma}$, if the pairwise majority graph $\Gamma(\boldsymbol{\sigma}) = (A, E)$ is such that for all $x \in A$, $y \in A \setminus \{x\}$, either $(x, y) \in E$ or $(y, x) \in E$ (i.e., it is a *tournament*), and, moreover, Γ is acyclic, then $f(\boldsymbol{\sigma}) = \tau$ for the unique ranking τ induced by $\Gamma(\boldsymbol{\sigma})$. Caragiannis et al. [28] give many examples of prominent voting rules that are PMC, including the Kemeny rule, the Slater rule, the ranked pairs method, Copeland's method, and Schulze's method.

The Mallows Model Let the *Kendall tau* distance between two rankings $\sigma, \sigma' \in \mathcal{L}$ be

$$d_{\mathrm{KT}}(\sigma, \sigma') \triangleq |\{(x, y) \in A^2 : \ x \succ_{\sigma} y \ \wedge \ y \succ_{\sigma'} x\}|.$$

In words, it is the number of pairs of alternatives on which σ and σ' disagree. For example, if $\sigma = (a, b, c, d)$, and $\sigma' = (a, c, d, b)$, then $d_{\mathrm{KT}}(\sigma, \sigma') = 2$.

In the Mallows [62] model, there is a *ground truth ranking* $\sigma^\star$, which induces a probability distribution over perceived rankings. Specifically, the probability of a ranking σ, given the ground truth ranking $\sigma^\star$, is given by

$$\Pr[\sigma \mid \sigma^\star] \triangleq \frac{\phi^{d_{\mathrm{KT}}(\sigma, \sigma^\star)}}{Z},$$

where $\phi \in (0, 1]$ is a parameter, and

$$Z \triangleq \sum_{\sigma' \in \mathcal{L}} \phi^{d_{\mathrm{KT}}(\sigma', \sigma^\star)}$$

is a normalization constant. Note that for $\phi = 1$ this is a uniform distribution, whereas the probability of $\sigma^\star$ goes to 1 as ϕ goes to 0. For ease of exposition, we assume that $\phi < 1$.

From Predictions to Mallows

In the virtual democracy framework, we are faced at runtime with a dilemma that induces a set of alternatives A. For example, when a food bank receives a donation, the set of alternatives is the current set of recipient organizations, each associated with information specific to the current donation, such as the distance between the donor and the recipient. Each voter $i \in N$ has a ranking $\sigma_i^\star \in \mathcal{L}$ over the given set of alternatives; together these rankings comprise the true preference profile $\boldsymbol{\sigma}^\star$.

One of the novel components of this paper is the assumption that, for each voter $i \in N$, we obtain a *predicted* ranking σ_i drawn from a Mallows distribution with parameter ϕ and true ranking $\sigma_i^\star$. We emphasize that, in contrast to almost all work on the Mallows Model, in our setting each voter has her own true ranking.

Why is the Mallows Model a good choice here? Recall that we are building preference models using pairwise comparisons as training data. When validating a model, we therefore test its accuracy on pairwise comparisons. And the Mallows model itself, because it is defined via the Kendall tau distance, is essentially determined by pairwise comparisons. In fact, the Mallows model (with parameter ϕ and true ranking $\sigma^\star$) is equivalent to the following generative process: for each pair of alternatives x and y such that $x \succ_{\sigma^\star} y$, x is preferred to y with probability $1/(1+\phi)$, and y is preferred to x with probability $\phi/(1+\phi)$; if this preference relation corresponds to a ranking (i.e., it is transitive), return that ranking, otherwise restart.

In more detail, let β be the average probability that we predict a pairwise comparison correctly; in our food bank implementation, $\beta \approx 0.9$. Based on the preceding discussion, one might be tempted to set $\beta = 1/(1+\phi)$, i.e., set β to be the probability of getting the relative ordering of two *adjacent* alternatives correctly. While this is not unreasonable (and would have been very convenient for us), for $\beta \approx 0.9$ it would lead to extremely high probability of correctly ranking alternatives that are, say, 30 positions apart in the ground truth ranking. In order to moderate this effect, we define another parameter $\kappa \in \{2, \dots, m\}$, and assume that our observed pairwise comparisons are between $\sigma_i^\star(1)$ (the top-ranked alternative in the true ranking of i) and $\sigma_i^\star(\kappa)$ (the alternative ranked in position κ). Formally, the parameters β and κ are such that, for the ranking σ_i sampled from a Mallows Model with ϕ and $\sigma_i^\star$,

$$\Pr\left[\sigma_i^\star(1) \succ_{\sigma_i} \sigma_i^\star(\kappa)\right] = \beta. \tag{1}$$

It is worth noting that the implicit assumption that we are observing comparisons between $\sigma_i^\star(1)$ and $\sigma_i^\star(\kappa)$ specifically is not meant to be realistic. Rather, the idea is that there is *some* appropriate value of κ such that the observed accuracy β can be related to the underlying Mallows model through Equation (1), and, if we can establish results that are *general* with respect to the choice of κ, they would carry over to the real world.

Robustness of Borda Count

First, we rigorously establish the robustness of Borda count to prediction error. As we have already discussed, we do not have access to the Mallows parameter ϕ. Instead, we can measure β, the probability that we correctly predict a pairwise comparison of alternatives that are κ positions apart. On a very high level, the theorem bounds the probability that the noisy Borda ranking (based on the sampled profile) would disagree with the true Borda ranking (based on the true profile) on a given pair of alternatives.

Theorem 1. *For any $\beta > 1/2$ and $\epsilon > 0$ there exists a universal constant $T = T(\beta, \epsilon)$ such that for all $n, m, \kappa \in \mathbb{N}$ such that $n, m \geq 2$, for all $s \geq T\kappa \log \kappa$, for all $\boldsymbol{\sigma}^\star \in \mathcal{L}^n$, and for all $x, x' \in A$ such that $\frac{1}{n}B(x, \boldsymbol{\sigma}^\star) \geq \frac{1}{n}B(x', \boldsymbol{\sigma}^\star) + 2s$, it holds that*

$$\Pr\left[\frac{1}{n}B(x, \boldsymbol{\sigma}) > \frac{1}{n}B(x', \boldsymbol{\sigma})\right] \geq 1 - \epsilon^n,$$

where the probability is taken over the sampling of $\boldsymbol{\sigma}$.

Let us discuss the statement of the theorem. First, note that the probability of mistake, ϵ^n, converges to 0 exponentially fast as n grows, so the theorem immediately implies a "with high probability" statement. Moreover, one can easily derive such a statement with respect to all pairs of alternatives (whose Borda scores are sufficiently separated) *simultaneously*, using a direct application of the union bound. Second, it is intuitive that the separation in Borda scores has to depend on κ, but it is encouraging (and, to us, surprising) that this dependence is almost linear. In particular, *even* if κ is almost linear in m, i.e., $\kappa \in o(m/\log m)$, the theorem implies that our noisy Borda ranking is highly unlikely to make mistakes on pairs of alternatives whose average score difference is linear in m.

It is important to note that it should be possible to extend Theorem 1 to other positional scoring rules defined by a score vector $(\alpha_1, \ldots, \alpha_m)$ where $\alpha_j > \alpha_{j+1}$ for all $j = 1, \ldots, m-1$. However, Borda count is especially practical and easy to explain and implement, which is why we focus on it for our positive result.

Non-Robustness of PMC Rules

Theorem 1 shows that Borda count is robust against noisy perturbations of the preference profile. It is natural to ask whether 'many' voting rules satisfy a similar property. In this section we answer this question in the negative, by proving that any voting rule that belongs to the important family of PMC rules is *not* robust in a similar sense.

Specifically, recall that under a PMC rule, when the weighted pairwise majority graph is acyclic, the output ranking is the topological ordering of the pairwise majority graph. We show that there exist profiles in which the pairwise majority graph is acyclic and all edge weights are large, but, with high probability, the noisy profile also has an acyclic pairwise majority graph which induces a *different* ranking. This means that *any* PMC rule would return different rankings when applied to the true profile and the noisy profile.

Theorem 2. *For all $\delta > 0$, $\phi \in (0, 1)$, and $m \in \mathbb{N}$ such that $m \geq 3$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exists a profile $\boldsymbol{\sigma}^\star \in \mathcal{L}^n$ such that $\Gamma(\boldsymbol{\sigma}^\star)$ is acyclic and all edges have weight $\Omega(n)$, but with probability at least $1 - \delta$ $\Gamma(\boldsymbol{\sigma})$ is acyclic and there is a pair of alternatives on which* *the unique rankings induced by* $\Gamma(\boldsymbol{\sigma}^\star)$ *and* $\Gamma(\boldsymbol{\sigma})$ *disagree, where the probability is taken over the sampling of* $\boldsymbol{\sigma}$*.*

It is instructive to contrast our positive result, Theorem 1, with this negative result. On a very high level, the former result asserts that "if Borda count says that the gaps between alternatives are significant, then the alternatives will not flip under Borda count," whereas the latter says "even if a PMC rule says that the gaps between alternatives are very significant, some alternatives are likely to flip under that rule." On a technical level, a subtle difference is that Theorem 1 is stated for β and κ, whereas Theorem 2 is stated directly for ϕ. This actually *strengthens* the negative result, because a constant β and $\kappa \in \omega(1)$ lead to $\phi = 1 - o(1)$, i.e., very noisy distributions — and still the positive result of Theorem 1 holds. By contrast, the negative result of Theorem 2 is true *even* when ϕ is constant, i.e., for settings that are not nearly as noisy. That said, the two results are not directly comparable, as Borda count and PMC rules deal with very different notions of score or weight. Nevertheless, the take-home message is that the notion of score that defines Borda count is inherently more robust to random perturbations of the preference profile.

2.3 Virtual Democracy in Practice: 412 Food Rescue

In concurrent work to the above, we apply the theoretical findings above to the real-world scenario of *food rescue* [59] in order to see how people perceive, participate in, and react to algorithmic governance. Through a collaboration with 412 Food Rescue, a nonprofit that matches food donations with needy recipients, we build a platform based on virtual democracy that suggests possible destinations for donations based on the learnt preferences of various stakeholders (donors, recipients, volunteers who deliver donations, and 412 Food Rescue employees).

Throughout the process, we solicited stakeholder participation to adjudicate the tradeoffs involved in the algorithm's design, balancing equity and efficiency in donation distribution and managing the associated disparate impacts on different stakeholders. Over the course of a year, we had the stakeholders use the WeBuildAI framework to design the matching algorithm, and researched their experiences through a series of studies. The findings suggest that our framework successfully enabled participants to build models that they felt confident represented their own beliefs. In line with our original goals, participatory algorithm design also impacted both procedural fairness and distributive outcomes: participants trusted and perceived as fair the collectively-built algorithm, and developed an empathetic stance toward the organization. Compared to human dispatchers, the resulting algorithm improved equity in donation distribution without hurting efficiency when tested with historic data. Finally, we discovered that the individual model-building raised participants' algorithmic awareness and helped identify inconsistencies in human managers' decision-making in the organization, and that the design of the individual model-building method may influence the elicited beliefs.

In the interest of space, we do not discuss this further here; see Lee et al. [59] for the full paper.

3 Liquid Democracy

Liquid democracy is a modern approach to voting in which voters can either vote directly or delegate their vote to other voters. In contrast to the classic *proxy voting* paradigm [65], the key innovation underlying liquid democracy is that proxies — who were selected by voters to vote on their behalf — may delegate their own vote to a proxy, and, in doing so, further delegate all the votes entrusted to them. Put another way (to justify the liquid metaphor), votes may freely flow through the directed delegation graph until they reach a sink, that is, a vertex with outdegree 0. When the election takes place, each voter who did not delegate his vote is weighted by the total number of votes delegated to him, including his own. In recent years, this approach has been implemented and used on a large scale, notably by eclectic political parties such as the German Pirate Party (Piratenpartei) and Sweden's Demoex (short for Democracy Experiment).

One reason for the success of liquid democracy is that it is seen as a practical compromise between *direct democracy* (voters vote directly on every issue) and *representative democracy*, and, in a sense, is the best of both worlds. Under liquid democracy, voters who did not invest an effort to learn about the issue at hand (presumably, most voters) would ideally delegate their votes to well-informed voters. This should intuitively lead to collective decisions that are less random, and more likely to be correct, than those that would be made under direct democracy.

Our first goal is to rigorously investigate the intuition that liquid democracy "outperforms" direct democracy from an algorithmic viewpoint. Indeed, we are interested in *delegation mechanisms*, which decide how votes should be delegated based on how relatively informed voters are, and possibly even based on the structure of an underlying social network. Therefore, our first research question, explored in Section 3.2, is:

> *Are there delegation mechanisms that are* guaranteed *to yield more accurate decisions than direct voting?*

Our first result highlights the significance, and potential dangers, of delegating many votes to few voters. Importantly, there is evidence that this can happen in practice. For example, Der Spiegel reported[1] that one member of the German Pirate Party, a linguistics professor at the University of Bamberg, amassed so much weight that his "vote was like a decree." In other words, our algorithmic result indicates that, even if delegations go only to more competent agents, a high concentration of power might still be harmful for social welfare, by neutralizing benefits corresponding to the Condorcet Jury Theorem.

While all these concerns suggest that the weight of super-voters should be limited, the exact metric to optimize for varies between them and is often not even clearly defined. For our purposes, we choose to minimize the weight of the heaviest voter. As is evident in the Spiegel article, the weight of individual voters plays a direct role in the perception of super-voters. But even beyond that, we are confident that minimizing this measure will lead to substantial improvements across all presented concerns.

Just how can the maximum weight be reduced? One approach might be to restrict the power of delegation by imposing caps on the weight. However, as argued by Behrens et al. [17], delegation is always possible by coordinating outside of the system and copying the desired delegate's ballot. Pushing delegations outside of the system would not alleviate the problem of super-voters, just reduce transparency. Therefore, we instead adopt a voluntary approach: If agents are considering multiple potential delegates, all of whom they trust, they are encouraged to leave the decision for one of them to a centralized mechanism. With the goal of avoiding high-weight agents in mind, our second research challenge, addressed in Section 3.3, is twofold:

> *First, investigate the algorithmic problem of selecting delegations to minimize the maximum weight of any agent, and, second, show that allowing multiple delegation options*

[1] http://www.spiegel.de/international/germany/liquid-democracy-web-platform-makes-professor-most-powerful-pirate-a-818683.html

does indeed provide a significant reduction in the maximum weight compared to the status quo.

3.1 Related Work

There is a significant body of work on delegative democracy and proxy voting [2, 4, 25, 29, 65, 85]. In particular, Cohensius et al. [30] study a model where voters' positions on an issue are points in a metric space. In their version of direct democracy, a small subset of active voters report their positions, and an aggregation method (such as the median or mean when the metric space is the real line) outputs a single position. Under proxy voting, each inactive voter delegates his vote to the closest active voter. Cohensius et al. identify conditions under which proxy voting gives a more accurate outcome than direct voting, where the measure is proximity of the outcome to the aggregation method applied to all voters' positions.

Another paper by Green-Armytage [48] considers a setting where, similarly to Cohensius et al. [30], voters are identified with points on the real line, but in his model votes are noisy estimates of those positions. He then defines the *expressive loss* of a voter as the squared distance between his vote and his position and proves that delegation (even transitive delegation) can only decrease the expressive loss in his model.

Work by Christoff and Grossi [29] introduces a model of liquid democracy based on the theory of binary aggregation (i.e., their model has a mathematical logic flavor). Their results focus on two problems: the possibility of delegation cycles, and logical inconsistencies that can arise when opinions on interdependent propositions are expressed through proxies. Both are nonissues in our models (although the need to avoid cycles is certainly a concern in practice).

Bloembergen et al. [21] consider a game-theoretic version of liquid democracy in which voters must determine whether or not it is rational to delegate their votes to others. They introduce a *delegation game* in which each voter has a hidden true "type" that she knows imperfectly, and the goal of each voter is to communicate her true type to the mechanism either directly (by voting) or indirectly (by delegating). While this setting distills the problem of finding delegates that represent one's own opinion, it focuses on proving the existence of Nash equilibria under certain assumptions and provides only weak performance bounds in the setting we consider.

Further afield, there is a rich body of work in computational social choice [24] on the aggregation of objective opinions [9, 10, 27, 28, 31, 32, 39, 60, 64, 71, 72, 88, 89]. As in our work, the high-level goal is to pinpoint the correct outcome based on noisy votes. However, previous work in this area does not encompass any notion of vote delegation.

3.2 An Algorithmic Perspective on Liquid Democracy

The Model

We focus on a (common) setting where a decision is to be made on a binary issue, i.e., one of two alternatives must be selected. To model the idea of accuracy, we assume that one alternative is correct, and the other is incorrect. Each voter $i \in N$ is labeled by his *competence level* p_i. This is the probability that i has the correct opinion about the issue at hand, i.e., the probability that i will vote correctly.

We represent an instance of our problem using a directed, labeled graph $G = (N, E, \vec{p})$. $N = \{1, \ldots, n\}$ is a set of n *voters*, also referred to as *vertices* (we use the two terms interchangeably).

E represents a (directed) social network in which the existence of an edge (i, j) means that voter i knows (of) voter j.

Our setting is also parameterized by $\alpha \in [0, 1)$. Given this parameter and a labeled graph $G = (V, E, \vec{p})$, we define an approval relation between voters: $i \in N$ *approves* $j \in N$ if $(i, j) \in E$ and $p_j > p_i + \alpha$. In words, i approves his neighbor j if the difference in their competence levels is strictly greater than α. The strict inequality guarantees that the approval relation is acyclic. Denote

$$A_G(i) = \{j \in N : i \text{ approves } j\}.$$

Delegation Mechanisms The liquid democracy paradigm is implemented through a *delegation mechanism* M, which takes as input a labeled graph G, and outputs, for each voter i, a delegation probability distribution over $A_G(i) \cup \{i\}$ that represents the probability that i will delegate his vote to each of his approved neighbors, or to himself (which means he does not delegate his vote).

To determine whether a delegation mechanism M makes a correct decision on a labeled graph $G = (N, E, \vec{p})$, we use the following 4-step process (which is described in words to avoid introducing notation that will not be used again):

1. Apply M to G to output a delegation probability distribution for each voter i.
2. Sample the probability distribution for each vertex to obtain an acyclic *delegation graph*. Each sink i of the delegation graph (i.e., vertex with no outgoing edges) has weight equal to the number of vertices with directed paths to i, including i itself.
3. Each sink i votes for the correct alternative with probability p_i, and for the incorrect alternative with probability $1 - p_i$.
4. A decision is made based on the weighted majority vote.[2]

We denote the probability that the mechanism M makes a correct decision on graph G via this 4-step process by $P_M(G)$.

We are particularly interested in a special class of delegation mechanisms that we call *local* mechanisms. Intuitively, local mechanisms capture the natural setting where each voter makes an independent delegation decision without central coordination. Formally, a *local delegation mechanism* is a delegation mechanism such that the probability distribution of each vertex i depends only on the neighborhood of i in G, i.e., on $\{j \in N : (i, j) \in E\}$, and on $A_G(i)$, i.e., the subset of these neighbors that are approved.

Desiderata Recall that we are interested in comparing the likelihood of making correct decisions via delegative voting with that of direct voting. To this end, define the *gain* of delegation mechanism M on labeled graph G as

$$\text{gain}(M, G) = P_M(G) - P_D(G).$$

We would like to design delegation mechanisms that have positive gain (bounded away from zero) in some situations, and which never lose significantly to direct voting. Formally, we are interested in the following two desirable axioms:

[2] Ties can be broken arbitrarily.

- A mechanism M satisfies the *positive gain (PG)* property if there exist $\gamma > 0, n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ there exists a graph G_n on n vertices such that $\text{gain}(M, G_n) \geq \gamma$.
- A mechanism M satisfies the *do no harm (DNH)* property if for all $\varepsilon > 0$, there exists $n_1 \in \mathbb{N}$ such that for all graphs G_n on $n \geq n_1$ vertices, $\text{gain}(M, G_n) \geq -\varepsilon$.

The choice of quantifiers here is of great significance. PG asks for the *existence* of (large enough) instances where the gain is at least γ, for a constant γ. By contrast, DNH essentially requires that any loss would go to 0 as the size of the graph goes to infinity. That is, there may certainly be small instances where delegative voting loses out to direct voting, but that should not be the case in the large.

Impossibility for Local Mechanisms

First, we establish an impossibility result for local delegation mechanisms: such mechanisms cannot satisfy both PG and DNH. In a nutshell, the idea is that for any local delegation mechanism that satisfies PG we can construct an instance where few voters amass a large number of delegated votes, that is, delegation introduces significant *correlation* between the votes. The instance is such that, when the high-weight voters are incorrect, the weighted majority vote is incorrect; yet direct voting is very likely to lead to a correct decision.

Theorem 3. *For any $\alpha_0 \in [0, 1)$ such that $i \in N$* approves *$j \in N$ if $(i, j) \in E$ and $p_j > p_i + \alpha_0$, there is no local mechanism that satisfies the PG and DNH properties.*

Possibility for Non-Local Mechanisms

The main idea underlying Theorem 3 is that liquid democracy can correlate the votes to the point where the mistakes of a few popular voters tip the scales in the wrong direction. However, non-local delegation mechanisms can circumvent this issue. Indeed, consider the following delegation mechanism.

input: labeled graph G with n vertices, cap $C : \mathbb{N} \to \mathbb{N}$

$N' \leftarrow N$
while $N' \neq \emptyset$ **do**
  let $i \in \text{argmax}_{j \in N'} |A_G^{-1}(j) \cap N'|$
  $J \leftarrow A_G(i) \cap N'$
  if $|J| \leq C(n) - 1$ **then**
    $J' \leftarrow J$
  else
    let $J' \subseteq J$ such that $|J'| = C(n) - 1$
  end if
  vertices in J' delegate to i
  $N' \leftarrow N' \setminus (\{i\} \cup \{J'\})$
end while

Algorithm 1: GreedyCap

In words, the mechanism GreedyCap, given as Algorithm 1, receives as input a labeled graph G, and a *cap* C. It iteratively selects a voter with maximum approvals, and delegates votes to him, so that no more than $C(n) - 1$ votes are delegated to a single voter (that is, no voter can have weight more than $C(n)$). All voters involved in the current iteration are then eliminated from further consideration, which is why delegations under this mechanism are only 1-hop.

It is obvious that GreedyCap satisfies the PG property. Additionally, with the mild assumption that the natural assumption that competence levels are bounded away from 0 and 1, i.e., voters are never perfectly misinformed or informed, we can show that GreedyCap satisfies the DNH property.

Theorem 4. *Assume that there exists $\beta \in (0, 1/2)$ such that all competence levels are in $[\beta, 1-\beta]$. Then for any $\alpha \in (0, 1-2\beta)$,* GreedyCap *with cap $C : \mathbb{N} \to \mathbb{N}$ such that $C(n) \in \omega(1)$ and $C(n) \in o(\sqrt{\log n})$ satisfies the PG and DNH properties.*

3.3 Minimizing the Maximum Weight of Voters in Liquid Democracy

Our first study highlights the significance, and potential dangers, of delegating many votes to few voters. Importantly, there is evidence that this can happen in practice. For example, Der Spiegel reported[3] that one member of the German Pirate Party, a linguistics professor at the University of Bamberg, amassed so much weight that his "vote was like a decree." We now turn to minimizing the maximum weight of any voter under a centralized (non-local) mechanism.

Let us consider a delegative voting process where agents may specify multiple potential delegations. This gives rise to a directed graph, whose nodes represent agents and whose edges represent potential delegations. In the following, we will conflate nodes and the agents they represent. A distinguished subset of nodes corresponds to agents who have voted directly, the *voters*. Since voters forfeit the right to delegate, the voters are a subset of the sinks of the graph. We call all non-voter agents *delegators*.

Each agent has an inherent voting weight of 1. When the delegations will have been resolved, the weight of every agent will be the sum of weights of her delegators plus her inherent weight. We aim to choose a delegation for every delegator in such a way that the maximum weight of any voter is minimized.

Problem Statement

All graphs $G = (N, E)$ mentioned in this section will be finite and directed. Furthermore, they will be equipped with a set V of distinguished sinks in the graph. For the sake of brevity, these assumptions will be implicit in the notion "graph G with V".

Some of these graphs represent situations in which all delegations have already been resolved and in which each vote reaches a voter: We call a graph (N, E) with V a *delegation graph* if it is acyclic, its sinks are exactly the set V, and every other vertex has outdegree one. In such a graph, define the *weight* $w(n)$ of a node $n \in N$ as

$$w(n) \coloneqq 1 + \sum_{(m,n) \in E} w(m).$$

This is well-defined because E is a well-founded relation on N.

[3]http://www.spiegel.de/international/germany/liquid-democracy-web-platform-makes-professor-most-powerful-pirate-a-818683.html

Resolving the delegations of a graph G with V can now be described as the MinMaxWeight problem: Among all delegation subgraphs (N', E') of G with voting vertices V of maximum $|N'|$, find one that minimizes the maximum weight of the voting vertices.

Connections to Confluent Flow This task closely mirrors the problem of congestion minimization for confluent flow (with infinite edge capacity): There, a flow network is also a finite directed graph with a distinguished set of graph sinks, the *flow sinks*. Every node has a non-negative *demand*. If we assume unit demand, this demand is 1 for every node. Since the flow is confluent, for every non-sink node, the algorithm must pick exactly one outgoing edge, along which the flow is sent. Then, the *congestion* at a node n is the sum of congestions at all nodes who direct their flow to n plus the demand of n. The goal in congestion minimization is to minimize the maximum congestion at any flow sink.

Using this connection, we demonstrate the NP-hardness of approximating the MinMaxWeight problem to within a factor of $\frac{1}{2}\,\log_2 |V|$.

Theorem 5. *It is NP-hard to approximate the* MinMaxWeight *problem to a factor of* $\frac{1}{2}\log_2 |V|$*, even when each node has outdegree at most* 2.

Probabilistic Model and Results

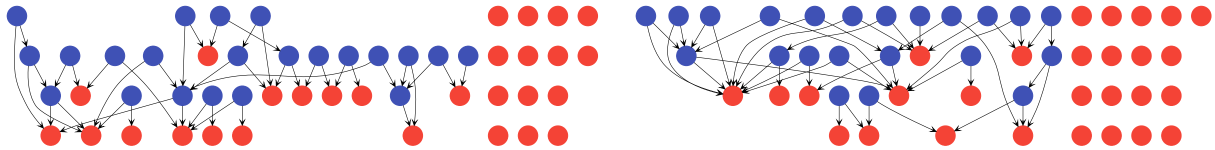
Our generalization of liquid democracy to multiple potential delegations aims to decrease the concentration of weight. Accordingly, the success of our approach should be measured by its effect on the maximum weight in real elections. Since, at this time, we do not know of any available datasets,[4] we instead propose a probabilistic model for delegation behavior, which can serve as a credible proxy. Our model builds on the well-known preferential attachment model, which generates graphs possessing typical properties of social networks.

In sections 3.3 and 3.3, for a certain choice of parameters in our model, we establish a striking separation between traditional liquid democracy and our system. In the former case, the maximum weight at time t is $\Omega(t^{\beta})$ for a constant β with high probability, whereas in the latter case, it is in $\mathcal{O}(\log\log t)$ with high probability, even if each delegator only suggests two options.

The Preferential Delegation Model Many real-world social networks have degree distributions that follow a power law [57, 68]. Additionally, in their empirical study, Kling et al. [55] observed that the weight of voters in the German Pirate Party was "power law-like" and that the graph had a very unequal indegree distribution. In order to meld the previous two observations in our liquid democracy delegation graphs, we adapt a standard preferential attachment model [16] for this specific setting. At a high level, our *preferential delegation* model is characterized by three parameters: $0 < d < 1$, the probability of delegation; $k \geq 1$, the number of delegation options from each delegator; and $\gamma \geq 0$, an exponent that governs the probability of delegating to nodes based on current weight.

At time $t = 1$, we have a single node representing a single voter. In each subsequent time step, we add a node for agent i and flip a biased coin to determine her delegation behavior. With probability d, she delegates to other agents. Else, she votes independently. If i does not delegate, her node has no outgoing edges. Otherwise, add edges to k many independently selected, previously

[4] There is one relevant dataset that we know of, which was analyzed by Kling et al. [55]. However, due to stringent privacy constraints, the data privacy officer of the German Pirate Party was unable to share this dataset with us.



(a) $\gamma = 0$ (b) $\gamma = 1$

Figure 1: Example graphs generated by the preferential delegation model for $k = 2$ and $d = 0.5$.

inserted nodes, where the probability of choosing node j is proportional to $(\mathit{indegree}(j) + 1)^\gamma$. Note that this model might generate multiple edges between the same pair of nodes, and that all sinks are voters. Figure 1 shows example graphs for different settings of γ.

In the case of $\gamma = 0$, which we term *uniform delegation*, a delegator is equally likely to attach to any previously inserted node. Already in this case, a "rich-get-richer" phenomenon can be observed, i.e., voters at the end of large networks of potential delegations will likely see their network grow even more. Indeed, a larger network of delegations is more likely to attract new delegators. In traditional liquid democracy, where $k = 1$ and all potential delegations will be realized, this explains the emergence of super-voters with excessive weight observed by Kling et al. [55]. We aim to show that for $k \geq 2$, the resolution of potential delegations can strongly outweigh these effects. In this, we profit from an effect known as the "power of two choices" in load balancing described by Azar et al. [8].

In our theoretical results, we focus on the cases of $k = 1$ and $k = 2$, and assume $\gamma = 0$ to make the analysis tractable. The parameter d can be chosen freely between 0 and 1. Note that our upper bound for $k = 2$ directly translates into an upper bound for larger k, since the resolution mechanism always has the option of ignoring all outgoing edges except for the first two. Therefore, to understand the effect of multiple delegation options, we can restrict our attention to $k = 2$. This crucially relies on $\gamma = 0$, where potential delegations do not influence the probabilities of choosing future potential delegations. Based on related results by Malyshkin and Paquette [63], it seems unlikely that increasing k beyond 2 will reduce the maximum weight by more than a constant factor.

Lower Bounds for Single Delegation ($k = 1$, $\gamma = 0$) As mentioned above, we first assume uniform delegation and a single delegation option per delegator, and derive a lower bound on the maximum weight.

Theorem 6. *In the preferential delegation model with $k = 1$, $\gamma = 0$, and $d \in (0, 1)$, with high probability, the maximum weight of any voter at time t is in $\Omega(t^\beta)$, where $\beta > 0$ is a constant that depends only on d.*

Upper Bound for Double Delegation ($k = 2$, $\gamma = 0$) Analyzing cases with $k > 1$ is considerably more challenging. One obstacle is that we do not expect to be able to incorporate optimal resolution of potential delegations into our analysis, because the computational problem is hard even when $k = 2$ (see Theorem 5). Therefore, we give a pessimistic estimate of optimal resolution via a greedy delegation mechanism, which we can reason about alongside the stochastic process. Clearly, if this stochastic process can guarantee an upper bound on the maximum weight with high probability, this bound must also hold if delegations are optimally resolved to minimize maximum weight.

In more detail, whenever a new delegator is inserted into the graph, the greedy mechanism immediately selects one of the delegation options. As a result, at any point during the construction of the graph, the algorithm can measure the weight of the voters. Suppose that a new delegator suggests two delegation options, to agents a and b. By following already resolved delegations, the mechanism obtains voters a^* and b^* such that a transitively delegates to a^* and b to b^*. The greedy mechanism then chooses the delegation whose voter currently has lower weight, resolving ties arbitrarily.

This situation is reminiscent of a phenomenon known as the "power of choice." In its most isolated form, it has been studied in the *balls-and-bins* model, for example by Azar et al. [8]. In this model, n balls are to be placed in n bins. In the classical setting, each ball is sequentially placed into a bin chosen uniformly at random. With high probability, the fullest bin will contain $\Theta(\log n/ \log\log n)$ balls at the end of the process. In the choice setting, two bins are independently and uniformly selected for every ball, and the ball is placed into the emptier one. Surprisingly, this leads to an exponential improvement, where the fullest bin will contain at most $\Theta\left(\log\log n\right)$ balls with high probability.

We show that, at least for $\gamma = 0$ in our setting, this effect outweighs the "rich-get-richer" dynamic described earlier:

Theorem 7. *In the preferential delegation model with $k = 2$, $\gamma = 0$, and $d \in (0, 1)$, the maximum weight of any voter at time t is $\log_2 \ln t + \Theta(1)$ with high probability.*

This establishes a doubly-exponential separation between the maximum weight of any voter in the single delegation and double delegation cases; this gap will only grow as the number of delegations increases.

4 Impartial Ranking

Our work is primarily motivated by online labor markets, such as Upwork or Freelancer. In the bigger markets, employers typically receive dozens of applications for a job, but there is an embarrassment of riches, because employers do not have the knowledge required to accurately evaluate applicants. For the past year we have been building a prototype of a new online labor market, where applicants for a job — who are well-suited to evaluate applications for that same job — rank each other. We would like to implement a mechanism that aggregates these rankings into a single ranking that is then shown to the employer.

However, the foregoing application has a clear problem, which gets in the way of applying standard rank aggregation rules: strategic behavior. Specifically, in these relatively high-stakes scenarios, it is likely that a player would try to improve his own position in the output ranking by manipulating his reported ranking. For example, he might weaken a strong contender for the top position by ranking him last. Our goal, therefore, is to design rank aggregation rules that are *impartial*, in the sense that the position of a player in the output ranking is completely independent of the report of that player.

On a high level, our approach is to design *randomized* rank aggregation rules that are impartial and closely emulate standard rank aggregation rules that are not impartial. Specifically, we focus on providing impartial approximations to the important class of *pairwise* rules, which, as input, only require information about the fraction of players ranking any one player above another. Our theoretical results crucially depend on the notion of approximation — or measure of error — in question.

4.1 Related Work

At this point there is a significant body of work on the design of impartial mechanisms [5, 18, 20, 22, 34, 41, 49, 61, 83], including several papers in major AI conferences [13, 58]. We only elaborate on the papers that are most closely related to ours.

The paper of de Clippel et al. [34] introduced the notion of impartiality, in the context of dividing *credit* for a joint project. Specifically, the output of their mechanism is the fraction of the total credit each player receives, and impartiality means that a player cannot affect his own share of the credit. This mechanism is deployed on the fair division website `Spliddit.org`, where one of the suggested applications is ordering authors on scientific papers. However, an impartial credit division mechanism does *not* induce an impartial ranking mechanism, because, when players are sorted by credit, a player can improve his own position by decreasing another player's share.

Berga and Gjorgjiev [18] study the impartial rank aggregation problem from an axiomatic viewpoint, but focus on deterministic rules and a stronger notion of impartiality. Their results suggest that deterministic impartial rank aggregation methods are severely limited, and support our focus on randomized algorithms.[5]

On a technical level, our k-PARTITE algorithm is reminiscent of an algorithm of Alon et al. [5], in that it randomly partitions the players into subsets, and the outcome of players in one subset is only determined by players in other subsets. But the details of the algorithm, and its analysis, are completely different.

4.2 Impartial Ranking: Theoretical Results

Preliminaries

Rankings and Aggregation For any $k \in \mathbb{N}$, let $[k] = \{1, \ldots, k\}$. Our setting involves a set of players $[n] = \{1, \ldots, n\}$. The opinions of players are represented as rankings over $[n]$, which we think of as *permutations*. Let $\mathcal{L}$ represent the set of all permutations of $[n]$, and let $\mathcal{L}^n$ represent the set of all *input profiles*. For any $\sigma \in \mathcal{L}$, let $\sigma(j)$ be the player at position j in σ and let $\sigma^{-1}(i)$ be the position of player i in the ranking σ (where position 1 is the highest and position n is the lowest).

A *deterministic rank aggregation rule* (also known as a *social welfare function*) is a function $f : \mathcal{L}^n \to \mathcal{L}$, which takes in an input profile and returns a ranking. A *randomized* rank aggregation rule returns a probability distribution over rankings. We sometimes find it convenient to slightly abuse notation and think of the domain of a rank aggregation rule as $\mathcal{L}^n \times 2^{[n]}$ — for $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ and $X \subseteq [n]$, $f(\vec{\sigma}, X)$ is the application of the rule to the input profile $(\sigma_i)_{i \in X}$.

[5] In addition, it is easy to prove that there are no deterministic rank aggregation rules that are both impartial (according to our definition) and Pareto efficient [66]. The latter property means that if everyone ranks one player above another, so does the output ranking.

Pairwise Rank Aggregation Rules An input profile $\vec{\sigma} = (\sigma_1, \dots, \sigma_n)$ induces a *pairwise comparison matrix* $A(\vec{\sigma})$, where

$$A(\vec{\sigma})_{ij} = \frac{|\{k \in [n] : \; \sigma_k^{-1}(i) < \sigma_k^{-1}(j)\}|}{n}.$$

In words, the (i, j) entry is the fraction of players who rank i above j. Let Ω be the set of pairwise comparison matrices. Therefore, we can think of $A : \mathcal{L}^n \to \Omega$ as a function that takes in an input profile and returns its associated pairwise comparison matrix. As before, we will also use the notation $A(\vec{\sigma}, X)$, for a subset of players $X \subseteq [n]$, to denote the pairwise comparison matrix associated with the rankings of the players in X.

Some rank aggregation rules only require the information encoded in the pairwise comparison matrix to compute their output. Formally, a deterministic *pairwise* rank aggregation rule is a function $f : \Omega \to \mathcal{L}$. We denote the class of all deterministic pairwise rules by $\mathcal{P}$.

Impartiality Recall that we are interested in designing rank aggregation rules that are impartial, that is, no player i can affect his probability of being ranked in position j, for all $i, j \in [n]$. Formally:

Definition 8. *A (possibly randomized) rank aggregation rule f is* impartial *if for all $i \in [n]$, all input profiles $(\sigma_1, \dots, \sigma_n) \in \mathcal{L}^n$, and all $\tilde{\sigma}_i \in \mathcal{L}$, it holds that $\vec{x} = \vec{y}$, where x_j is the probability i is ranked in position j in $f(\sigma_1, \dots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \dots, \sigma_n)$, and y_j is the probability i is ranked in position j in $f(\sigma_1, \dots, \sigma_{i-1}, \tilde{\sigma}_i, \sigma_{i+1}, \dots, \sigma_n)$.*

In an alternative model, we may assume that each player i has a value v_{ij} for being ranked in position j, and then impartiality would mean no player can affect his *expected* value for the outcome, regardless of his value function. Although this definition may seem weaker than Definition 8 at first glance, it is easy to verify that the two definitions are, in fact, equivalent.

Measures of Error Given that our goal is to approximate rank aggregation rules, the measure of error is critical to the statement of the formal problem. To define appropriate notions, we adapt concepts that are standard in scientific computing (e.g., in numerical stability analysis): forward error, backward error, and mixed error. We view these imported definitions as part of our conceptual contribution.

Definition 9. *Let f be a rank aggregation rule. A rank aggregation rule g is said to have (Δ_P, Δ_F)* forward error *with respect to f if for every input profile $\vec{\sigma} \in \mathcal{L}^n$, the probability that for all $i \in [n]$ it holds that*

$$\frac{|f(\vec{\sigma})^{-1}(i) - g(\vec{\sigma})^{-1}(i)|}{n} < \Delta_F$$

is at least $1 - \Delta_P$.

Intuitively, a low amount of forward error implies that every player i is placed near his correct rank (as determined by f) with high probability. Unfortunately, as the next theorem states, impartial rank aggregation rules cannot approximate the Borda rule. Since Borda is a pairwise rule, the theorem rules out the possibility of approximating all pairwise rules.

Theorem 10. *For all $n \geq 2$ and $\varepsilon > 0$, there exists no impartial rank aggregation rule g that gives a $(1/2 - \varepsilon, 1/3)$ forward error with respect to the Borda rule f.*

With this impossibility in hand, we set our sights on an alternate error measure, which is well defined only with respect to *pairwise* rank aggregation rules. For this definition and throughout the paper, we use the Frobenius norm and denote $\|A\|_\infty = \max_{i,j} |A_{i,j}|$.

Definition 11. *Let $f \in \mathcal{P}$. A rank aggregation rule g is said to have (Δ_P, Δ_B)* backward error *with respect to f if for every input profile $\vec{\sigma} \in \mathcal{L}^n$ the probability that for all $i \in [n]$ there exists a matrix $\tilde{A} \in \Omega$ such that*

1. $\left\|A(\vec{\sigma}) - \tilde{A}\right\|_\infty < \Delta_B$*, and*
2. $f(\tilde{A})^{-1}(i) = g(\vec{\sigma})^{-1}(i)$*,*

is at least $1 - \Delta_P$*.*

Intuitively, a low amount of backward error implies that every player i is placed in a rank that had the players altered their opinions slightly, i would be in the correct rank (according to f) with high probability.

Finally, we define a third measure of error, which, in a sense, is a union of the two previous notions.

Definition 12. *Let $f \in \mathcal{P}$. A rank aggregation rule g is said to have $(\Delta_P, \Delta_B, \Delta_F)$* mixed error *with respect to f if for every input profile $\vec{\sigma} \in \mathcal{L}^n$, the probability that for all $i \in [n]$ there exists a matrix $\tilde{A} \in \Omega$ such that*

1. $\left\|A(\vec{\sigma}) - \tilde{A}\right\|_\infty < \Delta_B$*, and*
2. $\frac{\left|f(\tilde{A})^{-1}(i) - g(\vec{\sigma})^{-1}(i)\right|}{n} < \Delta_F$*,*

is at least $1 - \Delta_P$*.*

The k-Partite Algorithm

We now introduce and analyze our first impartial rule, k-PARTITE, which is formally given as Algorithm 2. As it appears somewhat opaque, it is best to understand its ideas when we assume that all the X_i are the same size, i.e., k divides n, $|X_i| = n/k$, and $\gamma_i = k$ for all $i \in [k]$. Slight adjustments are made when this is not the case, which for purposes of intuition can be safely ignored.

First, players are randomly split into k groups of equal size $X_1, \ldots, X_k$, and then each such group separately ranks all n players producing rankings τ_i. The crux of the algorithm is the construction of the matrix Z, which, in turn, is the sum of $Z^{(i)}$ matrices. Intuitively, the $Z^{(i)}$ matrix represents X_i's contribution to Z, and its (a, b) entry indicates the probability that a should be placed in position b overall. Specifically, each player not in X_i is placed in his exact position dictated by τ_i with probability $1/k$, and in all positions that the players in X_i themselves were assigned to in τ_i with probability $1/(n(k-1))$. This information is encoded as the only non-zero entries in $Z^{(i)}$ — each column then sums to $1/k$, each row representing a player in X_i is zero, and all other rows sum to $1/(k-1)$. Because Z is doubly stochastic (its rows and columns sum to 1), we can apply the Birkhoff-von Neumann Theorem [19, 86] to sample from this distribution and remain faithful to the probabilities.

We now state the following theorem, which states the guarantees of k-PARTITE.

input: $f \in \mathcal{P}$ and $\vec{\sigma} \in \mathcal{L}^n$

Randomly split all n players into k groups $X_1, \ldots, X_k$ where $|X_i| \in \{\lfloor n/k \rfloor, \lceil n/k \rceil\}$
for $i = 1, \ldots, k$ **do**
  $\tau_i \leftarrow f(\vec{\sigma}, X_i)$
  $\gamma_i \leftarrow n / |X_i|$
  Let $Z^{(i)} \in \mathbb{R}^{n \times n}$ where

$$Z^{(i)}_{a,b} \leftarrow \begin{cases} \frac{1}{\gamma_i} & \text{if a} \notin X_i \text{ and } \tau_i(b) = a \\ \frac{1}{\gamma_i(\gamma_i - 1)|X_i|} & \text{if a} \notin X_i \text{ and } \tau_i(b) \in X_i \\ 0 & \text{otherwise} \end{cases}$$

end for
$Z \leftarrow \sum_{i \in [k]} \frac{\frac{n}{|X_i|} - 1}{k - 1} Z^{(i)}$
Sample a ranking σ such that a is ranked in position b with probability $Z_{a,b}$
return σ

Algorithm 2: k-PARTITE

Theorem 13. *k-PARTITE is impartial, and, for every $f \in \mathcal{P}$ and $\vec{\sigma} \in \mathcal{L}^n$, if $k = \lfloor (n/\ln n)^{1/3} \rfloor$, it gives at most*

$$(4/k, 4/k) \in \left(O\left(\left(\frac{\ln n}{n} \right)^{1/3} \right), O\left(\left(\frac{\ln n}{n} \right)^{1/3} \right) \right)$$

backward error with respect to f.

Note that, in particular, the error goes to 0 as n grows.

The Committee Algorithm

k-PARTITE demonstrates that there exist impartial rules that accurately imitate any $f \in \mathcal{P}$. Observe, however, that the algorithm is somewhat hamstrung by the fact that a player must be (with high probability) ranked in *exactly* the location that a small perturbation of the input rankings would give.

To allow more flexibility, we focus on mixed error, and consider COMMITTEE, given as Algorithm 3. Intuitively, this algorithm selects a random committee $X = \{x_1, \ldots, x_k\}$, which then determines the entire ranking. First, for each committee member x_i, we determine their rank using only the rankings given by the remaining $k - 1$ members. However, as directly placing each committee member in this fashion may cause collisions (i.e., multiple members may be assigned the same rank) we restrict placement of x_i to only the positions $i, i + k, i + 2k, \ldots$ Specifically, we assign x_i to the closest such position to the rank given to x_i by the other committee members. There are then k of the n positions assigned. Second, the committee ranks all of the n players, and the non-committee members are placed in the order ranked by the committee in the remaining $n - k$ slots.

The algorithm yields the following guarantees.

input: $f \in \mathcal{P}$ and $\vec{\sigma} \in \mathcal{L}^n$

Randomly select a subset $X = \{x_1, \ldots, x_k\} \subseteq [n]$
for $i = 1, \ldots, k$ **do**
  $c \leftarrow \arg\min_{j \in \{i, i+k, \ldots\}} \left| j - f\left(\vec{\sigma}, X \setminus \{x_i\}\right)^{-1}(x_i) \right|$
  $\sigma(c) \leftarrow x_i$
end for
$\tau \leftarrow f(\vec{\sigma}, X)$
$j \leftarrow 1$
for $i = 1, \ldots, n$ **do**
  if $\tau(i) \notin X$ **then**
    while $\sigma(j)$ is occupied **do**
      $j \leftarrow j + 1$
    end while
    $\sigma(j) \leftarrow \tau(i)$
  end if
end for
return σ

Algorithm 3: COMMITTEE

Theorem 14. COMMITTEE *is impartial, and, for every* $f \in \mathcal{P}$, $\vec{\sigma} \in \mathcal{L}^n$, *and* $\varepsilon > 0$, *if*

$$k = 1 + \frac{2}{\varepsilon^2} \ln\left(\frac{n^3}{\varepsilon}\right),$$

it gives at most $(\varepsilon, \varepsilon, (k+1)/n)$ *mixed error with respect to* f.

Importantly, this theorem allows for an incomparable error to Theorem 13. That is, we can reduce the backward error so long as we are willing to take on some forward error. For example, setting ε appropriately gives at most $\left(n^{-2/5}, n^{-2/5}, 2/n + (34/5)n^{-1/5} \ln n\right)$ mixed error.

4.3 HirePeer: Impartial Peer Assessment for Hiring in Online Labor Markets

Expert crowdsourcing (e.g., Upwork.com) provides promising benefits such as productivity improvements for employers, and flexible working arrangements for workers. Yet to realize these benefits, a key persistent challenge is effective hiring at scale. Current approaches, such as reputation systems and standardized competency tests, develop weaknesses such as score inflation over time, thus degrading market quality. In conjunction with the theoretical work described above, we develop *HirePeer*, a novel alternative approach to hiring at scale that leverages peer assessment to elicit honest assessments of fellow workers' job application materials, which it then aggregates using an impartial ranking algorithm [56]. We perform three studies that investigate both the costs and the benefits to workers and employers of impartial peer-assessed hiring. We find, to solicit honest assessments, algorithms must be communicated in terms of their impartial effects. Second, in practice, peer assessment is highly accurate, and impartial rank aggregation algorithms incur a small accuracy cost for their impartiality guarantee. Third, workers report finding peer-assessed hiring useful for receiving targeted feedback on their job materials.

In the interest of space, see Kotturi et al. [56] for the complete paper.

5 Future Directions

We now discuss ongoing and future work. In particular, there are three main avenues of active research. The first and third deal with developing rules that satisfy certain theoretical properties—namely, consensus and representation. The second is an extension of virtual democracy to a new domain: school choice.

5.1 Consensus Mechanisms

In group decision-making, one other desirable property of a voting mechanism is one that pushes agents to reach consensus by agreeing on a compromise that is tolerable to all agents. For instance, one intuitive shortcoming of modern political systems is deadlock among parties that refuse to negotiate with each other. One current project tackles this problem by designing voting rules that incentivize agents to compromise with each other.

Related Work Economists and computer scientists have studied the bargaining problem for decades [67, 74, 76]. Perhaps the best-known work in the bargaining literature is that of Nash bargaining [67]. In the Nash bargaining setting, agents must unanimously agree on an outcome from a known and common set of outcomes A, and if no agreement is reached, some external alternative will be chosen such that each agent's utility for this external alternative is less than their minimum utility for any alternative in A. While Nash characterizes a solution for this setting (corresponding to the notion of maximum Nash welfare), this relies on the existence of a universally disliked external alternative, or threat.

Additionally, Rubinstein famously studied the equilibrium of a class of bargaining games with alternating offers in a setting with an infinite time horizon and (additively or multiplicatively) decreasing utility functions [76]. In contrast to his work, our setting does not consider an infinite time horizon or decreasing utilities over time; we aim to achieve consensus purely through the voting mechanism we use.

Preliminaries Given a set of n agents N and a set of m alternatives A, we denote the utility of agent $i \in N$ for alternative $a \in A$ as $u_i(a)$. We assume that utilities are bounded; i.e., $u_i(a) \in [0, 1]$ for all $i \in N$ and $a \in A$. We now define our notion of consensus.

Definition 15. *Alternative $a \in A$ is a* consensus alternative *if, for all agents $i \in N$, $u_i(a) \geq \frac{1}{m}\sum_{x\in A} u_i(x)$; i.e., if all agents value a as much as their mean utility alternative.*

As a first step, we also assume that agents know each other's utilities as an analysis tool; we plan to extend our analysis and results to the setting of incomplete information.

Our first goal is to design a voting rule that incentivizes agents to choose a consensus alternative when one exists. To this end, we propose Algorithm 4, an algorithm that proposes alternatives in a random order, one at a time, and asks agents to cast approval votes over each alternative. It continues until either all agents approve the current proposed alternative, at which point it stops and returns the unanimously approved alternative, or all alternatives have been proposed, at which point it randomly chooses an alternative uniformly at random and returns that random choice.

In the full information setting, it is possible to show that Algorithm 4 satisfies our first goal.

Theorem 16. *If the set of consensus alternatives is non-empty and agents have full knowledge of others' utilities, Algorithm 4 is guaranteed to return a consensus alternative.*

input: agents N, alternatives A, utility functions $u_1, \ldots, u_n$

```
randomly generate an order Π over A
for a ∈ Π do
    propose a as a possible consensus alternative
    each agent i submits a binary approval vote on a
    if all agents approve of a then
        return a
    end if
end for
return uniformly randomly drawn alternative from A
```

Algorithm 4: Iterative consensus mechanism

We are currently working to establish a separation between this type of algorithm and a large class of algorithms.

5.2 Virtual Democracy in School Choice

We are also interested in applying virtual democracy to high school matching in New York City. School matching is an instance of the stable marriage problem, and many cities, including New York and Boston, use variants of the Gale-Shapley algorithm [45] (also known as deferred acceptance) to match students to high schools [1]. The Gale-Shapley algorithm relies on each school and student having an ordered preference list over their counterparts, and the mechanism proceeds in rounds, with schools (resp. students) “proposing” to students (resp. schools) and students conditionally accepting offers. This proposal and conditional acceptance process continues until every student has accepted an offer. The mechanism is strategyproof for the proposing side, and the outcome is guaranteed to be stable.

However, it has recently come to our attention that the process by which schools generate rankings over students is remarkably ad hoc and sensitive to specific feature values. This problem can be addressed via virtual democracy: instead of asking a committee from each school to generate rules by hand for ranking students (in practice, by putting weight on different features), learn a model for how each member of each school’s committee evaluates students and let committees “vote” over the aggregate ranking of all students. The first question, then, is whether we can prove that the Gale-Shapley algorithm is robust to virtual democracy in the sense that the output of the matching algorithm given a noisy estimate of schools’ preferences will align with the output of the matching algorithm given schools’ true preferences. The second question, which is conditional on the first being false, is whether we can design another matching algorithm that will be robust to such noise.

5.3 Representation in Multiwinner Elections

We study *multiwinner approval-based elections*, where a group of agents, or voters, selects a committee from a set of candidates based on the agents’ preferences. Each agent expresses her preferences through an approval vote, where she designates a subset of candidates she approves for the committee, and all votes are then aggregated to select a winning committee from the pool of candidates.

Some multiwinner elections include a fixed committee size: voters must fill exactly k seats on a committee. This is known as the fixed number of winners (FNW) setting, and there is a large body of work on the complexity and proportionality of various voting rules in the FNW setting [12, 14, 15, 26, 70, 77, 81]. In contrast, we are interested in the setting in which there is no a priori fixed committee size, also known as the variable number of winners (VNW) setting. In this case, both the size of the committee and the candidates chosen to sit on the committee are informed by agents' votes.

In order to study proportionality in FNW elections, researchers have proposed the axioms of justified representation (JR), proportional justified representation (PJR), extended justified representation (EJR), and average satisfaction (AS) [14, 77], which capture the intuition that all sufficiently large groups that agree on sufficiently many candidates should achieve some measure of satisfaction. However, to our knowledge, we are the first to study representation in VNW elections.

Related Work There is a significant body of work studying proportionality in FNW elections. As mentioned above, Aziz et al. [14] put forward the compelling axiom of justified representation (JR), as well as a stronger version of this axiom, extended justified representation (EJR) to capture the notion that any sufficiently large and cohesive group of voters deserves some measure of representation in the elected committee. This idea was built upon by Sánchez-Fernández et al. [77], who introduced the intermediate axiom of proportional justified representation (PJR), a relaxation of EJR that is more stringent than JR.

Average satisfaction (AS) was first defined by Sánchez-Fernández et al. [77] In this paper, they studied the average satisfaction guaranteed by extended justified representation (EJR). Further work by Aziz et al. [15] showed that Proportional Approval Voting (PAV) guarantees a level of average satisfaction that implies EJR. Additionally, Skowron et al. [80] extend the notion of average satisfaction to the context of complete rankings as opposed to committee selection. Further work by Skowron [79] studies the proportionality degree of various multiwinner rules by considering the average satisfaction of all groups of a certain size.

There is also a significant body of work studying VNW elections; however, to the best of our knowledge, none of these works consider proportionality. Kilgour [54] proposes a multitude of rules for VNW elections; however, none of the proposed VNW rules are proportional. Fishburn and Pekeč [42] study threshold approaches to committee selection, which are VNW rules in the sense that the size of the selected committee depends on the approval votes. However, threshold approaches are also not proportional. Additionally, the Borda Mean Rule, which was characterized by Brandl and Peters [23] and also studied by Duddy et al. [36], can be seen as a VNW rule with approval votes, but is also not proportional. Finally, Faliszewski et al. [40] study the computational complexity of various VNW rules, but do not consider proportionality in their analysis.

Our Contributions

In the interest of space, we sketch our results so far below. For a more complete discussion, see Freeman et al. [44].

Our main research goal is to study proportionality in multiwinner elections with a variable number of winners. In particular, we study the proportionality measure of average satisfaction (AS) and show that there is a separation between the performance of deterministic and randomized voting rules.

As our first contribution, we develop a framework for thinking about proportionality in VNW elections. Previous work on proportionality in FNW elections is largely based on the concept of justified representation (and extensions thereof). However, JR-based notions of proportionality are less compelling in VNW elections than in FNW elections. Therefore, we instead base our approach on the concept of average satisfaction, which is arguably a more robust version of justified representation.

Second, we consider the proportionality guarantees of deterministic rules in the VNW setting. We extend three existing deterministic rules for the FNW setting to the VNW setting, and show that these rules do not guarantee good approximations to average satisfaction. We also prove upper bounds on the level of average satisfaction that any deterministic rule can provide.

Finally, motivated by the shortcomings of deterministic rules, we turn our attention to randomized rules and show that a natural randomized rule provides a good approximation to average satisfaction (and is strategyproof).

Future Work

However, many interesting open problems remain. In particular, we do not have matching upper and lower bounds for the average satisfaction guarantees that can be provided by deterministic and randomized rules, and we do not know whether there exists a randomized rule that can satisfy AS.

Additionally, determining the existence of rules that satisfy EJR is also an interesting question; while we have argued that natural extensions of JR and PJR make less sense for VNW elections than for FNW, EJR remains a compelling property.

More broadly, we have assumed voters gain utility whenever they agree with the placement of a candidate. A natural extension of this model would be one in which voters derive different levels of utility for an approved candidate being selected and a disapproved candidate being excluded. Extending our results to this setting appears nontrivial.

References

[1] A. Abdulkadiroğlu, P. A. Pathak, and A. E. Roth. The New York City high school match. *American Economic Review*, 95(2):364–367, 2005.

[2] B. Abramowitz and N. Mattei. Flexible representative democracy: An introduction with binary issues. *arXiv:1811.02921*, 2018.

[3] M. Aleksandrov, H. Aziz, S. Gaspers, and T. Walsh. Online fair division: Analysing a food bank problem. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2540–2546, 2015.

[4] D. Alger. Voting by proxy. *Public Choice*, 126(1–2):1–26, 2006.

[5] N. Alon, F. Fischer, A. D. Procaccia, and M. Tennenholtz. Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 101–110, 2011.

[6] K. Arrow. *Social Choice and Individual Values*. Wiley, 1951.

[7] K. J. Arrow. A difficulty in the concept of social welfare. *Journal of political economy*, 58(4): 328–346, 1950.

[8] Y. Azar, A. Z. Broder, A. R. Karlin, and E. Upfal. Balanced allocations. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing (STOC)*, pages 593–602, 1994.

[9] H. Azari Soufiani, D. C. Parkes, and L. Xia. Random utility theory for social choice. In *Proceedings of the 26th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 126–134, 2012.

[10] H. Azari Soufiani, D. C. Parkes, and L. Xia. Preference elicitation for general random utility models. In *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 596–605, 2013.

[11] H. Azari Soufiani, D. C. Parkes, and L. Xia. Computing parametric ranking models via rank-breaking. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, pages 360–368, 2014.

[12] H. Aziz and S. Huang. A polynomial-time algorithm to achieve extended justified representation. arXiv preprint arXiv:1703.10415, 2017.

[13] H. Aziz, O. Lev, N. Mattei, J. S. Rosenschein, and T. Walsh. Strategyproof peer selection: Mechanisms, analyses, and experiments. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, pages 397–403, 2016.

[14] H. Aziz, M. Brill, V. Conitzer, E. Elkind, R. Freeman, and T. Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017.

[15] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. On the complexity of extended and proportional justified representation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 902–909, 2018.

[16] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[17] J. Behrens, A. Kistner, A. Nitsche, and B. Swierczek. *The Principles of LiquidFeedback*. Interaktive Demokratie, 2014.

[18] D. Berga and R. Gjorgjiev. Impartial social rankings. Manuscript, 2014.

[19] G. Birkhoff. Three observations on linear algebra. *Universidad Nacional de Tucumán, Revista A*, 5:147–151, 1946.

[20] A. Bjelde, F. Fischer, and M. Klimm. Impartial selection and the power of up to two choices. In *Proceedings of the 11th Conference on Web and Internet Economics (WINE)*, pages 146–158, 2015.

[21] D. Bloembergen, D. Grossi, and M. Lackner. On rational delegations in liquid democracy. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1796–1803, 2019.

[22] N. Bousquet, S. Norin, and A. Vetta. A near-optimal mechanism for impartial selection. In *Proceedings of the 10th Conference on Web and Internet Economics (WINE)*, pages 133–146, 2014.

[23] F. Brandl and D. Peters. An axiomatic characterization of the Borda mean rule. *Social Choice and Welfare*, 52(4):685–707, 2019.

[24] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[25] M. Brill and N. Talmon. Pairwise liquid democracy. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 137–143, 2018.

[26] M. Brill, R. Freeman, S. Janson, and M. Lackner. Phragmén's voting methods and justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 406–413, 2017.

[27] I. Caragiannis, A. D. Procaccia, and N. Shah. Modal ranking: A uniquely robust voting rule. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*, pages 616–622, 2014.

[28] I. Caragiannis, A. D. Procaccia, and N. Shah. When do noisy votes reveal the truth? *ACM Transactions on Economics and Computation*, 4(3): article 15, 2016.

[29] Z. Christoff and D. Grosssi. Binary voting with delegable proxy: An analysis of liquid democracy. In *Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 134–150, 2017.

[30] G. Cohensius, S. Mannor, R. Meir, E. Meirom, and A. Orda. Proxy voting for better outcomes. In *Proceedings of the 16th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 858–866, 2017.

[31] V. Conitzer and T. Sandholm. Common voting rules as maximum likelihood estimators. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 145–152, 2005.

[32] V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 109–115, 2009.

[33] V. Conitzer, W. Sinnott-Armstrong, J. Schaich Borg, Y. Deng, and M. Kramer. Moral decision making frameworks for artificial intelligence. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 4831–4835, 2017.

[34] G. de Clippel, H. Moulin, and N. Tideman. Impartial division of a dollar. *Journal of Economic Theory*, 139:176–191, 2008.

[35] M. de Condorcet. Essai sur l'application de l'analyse à la probabilité de décisions rendues à la pluralité de voix. Imprimerie Royal, 1785. Facsimile published in 1972 by Chelsea Publishing Company, New York.

[36] C. Duddy, A. Piggins, and W. S. Zwicker. Aggregation of binary evaluations: a Borda-like approach. *Social Choice and Welfare*, 46(2):301–333, 2016.

[37] J. Edmonds and K. Pruhs. Cake cutting really is not a piece of cake. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 271–278, 2006.

[38] E. Elkind and N. Shah. Electing the most probable without eliminating the irrational: Voting over intransitive domains. In *Proceedings of the 30th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 182–191, 2014.

[39] E. Elkind, P. Faliszewski, and A. Slinko. Good rationalizations of voting rules. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, pages 774–779, 2010.

[40] P. Faliszewski, A. Slinko, and N. Talmon. The complexity of multiwinner voting rules with variable number of winners. arXiv preprint arXiv:1711.06641, 2017.

[41] F. Fischer and M. Klimm. Optimal impartial selection. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, pages 803–820, 2014.

[42] P. C. Fishburn and A. Pekeč. Approval voting for committees: Threshold approaches. Technical Report, 2004.

[43] R. Freedman, J. Schaich Borg, W. Sinnott-Armstrong, J. P. Dickerson, and V. Conitzer. Adapting a kidney exchange algorithm to align with human values. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1636–1643, 2018.

[44] R. Freeman, A. Kahng, and D. M. Pennock. Proportionality in approval-based elections with a variable number of winners, 2020. In submission.

[45] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *Americal Mathematical Monthly*, 69(1):9–15, 1962.

[46] A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587–602, 1973.

[47] P. Gölz, A. Kahng, S. Mackenzie, and A. Procaccia. The fluid dynamics of liquid democracy. In *Proceedings of the 13th Conference on Web and Internet Economics (WINE)*, pages 188–202, 2018.

[48] J. Green-Armytage. Direct voting and proxy voting. *Constitutional Political Economy*, 26(2): 190–220, 2015.

[49] R. Holzman and H. Moulin. Impartial nominations for a prize. *Econometrica*, 81(1):173–196, 2013.

[50] A. X. Jiang, L. S. Marcolino, A. D. Procaccia, T. Sandholm, N. Shah, and M. Tambe. Diverse randomized agents vote to win. In *Proceedings of the 28th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 2573–2581, 2014.

[51] A. Kahng, Y. Kotturi, C. Kulkarni, D. Kurokawa, and A. D. Procaccia. Ranking wily people who rank each other. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1087–1094, 2018.

[52] A. Kahng, S. Mackenzie, and A. D. Procaccia. Liquid democracy: An algorithmic perspective. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, 2018. Forthcoming.

[53] A. Kahng, M. K. Lee, R. Noothigattu, A. D. Procaccia, and C. Psomas. Statistical foundations of virtual democracy. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 3173–3182, 2019.

[54] D. M. Kilgour. Approval elections with a variable number of winners. *Theory and Decision*, 81(2):199–211, 2016.

[55] C. C. Kling, J. Kunegis, H. Hartmann, M. Strohmaier, and S. Staab. Voting behaviour and power in online democracy. In *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM)*, pages 208–217, 2015.

[56] Y. Kotturi, A. Kahng, A. D. Procaccia, and C. Kulkarni. Hirepeer: Impartial peer-assessed hiring at scale in expert crowdsourcing markets. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

[57] R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *Proceedings of the 12th International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 611–617, 2006.

[58] D. Kurokawa, O. Lev, J. Morgenstern, and A. D. Procaccia. Impartial peer review. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 582–588, 2015.

[59] M. K. Lee, D. Kusbit, A. Kahng, J. T. Kim, X. Yuan, A. Chan, R. Noothigattu, D. See, S. Lee, C. Psomas, and A. D. Procaccia. WeBuildAI: Participatory framework for fair and efficient algorithmic governance. In *Proceedings of the 22nd ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, pages 1–35, 2019.

[60] T. Lu and C. Boutilier. Robust approximation and incremental elicitation in voting protocols. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 287–293, 2011.

[61] A. Mackenzie. Symmetry and impartial lotteries. *Games and Economic Behavior*, 94:15–28, 2015.

[62] C. L. Mallows. Non-null ranking models. *Biometrika*, 44:114–130, 1957.

[63] Y. Malyshkin and E. Paquette. The power of choice over preferential attachment. *Latin American Journal of Probability and Mathematical Statistics*, 12(2):903–915, 2015.

[64] A. Mao, A. D. Procaccia, and Y. Chen. Better human computation through principled voting. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1142–1148, 2013.

[65] J. C. Miller. A program for direct and proxy voting in the legislative process. *Public Choice*, 7(1):107–113, 1969.

[66] H. Moulin. Personal communication, 2014.

[67] J. F. Nash Jr. The bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 155–162, 1950.

[68] M. E. J. Newman. Clustering and preferential attachment in growing networks. *Physical review E*, 64(2):1–13, 2001.

[69] R. Noothigattu, S. S. Gaikwad, E. Awad, S. Dsouza, I. Rahwan, P. Ravikumar, and A. D. Procaccia. A voting-based system for ethical decision making. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 1587–1594, 2018.

[70] D. Peters and P. Skowron. Proportionality and the limits of welfarism. arXiv preprint arXiv:1911.11747, 2019.

[71] A. D. Procaccia, S. J. Reddi, and N. Shah. A maximum likelihood approach for selecting sets of alternatives. In *Proceedings of the 28th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 695–704, 2012.

[72] A. D. Procaccia, N. Shah, and Y. Zick. Voting rules as error-correcting codes. *Artificial Intelligence*, 231:1–16, 2016.

[73] J. M. Robertson and W. A. Webb. *Cake Cutting Algorithms: Be Fair If You Can.* A. K. Peters, 1998.

[74] A. E. Roth and M. W. Malouf. Game-theoretic models and the role of information in bargaining. *Psychological review*, 86(6):574, 1979.

[75] A. E. Roth and M. Sotomayor. *Two-Sided Mathing: A Study in Game-Theoretic Modelling and Analysis.* Cambridge University Press, 1990.

[76] A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109, 1982.

[77] L. Sánchez-Fernández, E. Elkind, M. Lackner, N. Fernández, J. A. Fisteus, P. B. Val, and P. Skowron. Proportional justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 670–676, 2017.

[78] M. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10: 187–217, 1975.

[79] P. Skowron. Proportionality degree of multiwinner rules. arXiv preprint arXiv:1810.08799, 2018.

[80] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. Proportional rankings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 409–415, 2017.

[81] P. Skowron, M. Lackner, E. Elkind, and L. Sánchez-Fernández. Optimal average satisfaction and extended justified representation in polynomial time. arXiv preprint arXiv:1704.00293, 2017.

[82] H. Steinhaus. The problem of fair division. *Econometrica*, 16:101–104, 1948.

[83] S. Tamura and S. Ohseto. Impartial nomination correspondences. *Social Choice and Welfare*, 43:47–54, 2014.

[84] M. Tennenholtz and A. Zohar. The axiomatic approach and the Internet. In F. Brandt, V. Conitzer, U. Endress, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 18. Cambridge University Press, 2016.

[85] G. Tullock. Computerizing politics. *Mathematical and Computer Modelling*, 16(8–9):59–65, 1992.

[86] J. von Neumann. A certain zero-sum two-person game equivalent to the optimal assignment problem. In W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games*, volume 2, pages 5–12. Princeton University Press, 1953.

[87] L. Xia. Bayesian estimators as voting rules. In *Proceedings of the 32nd Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, 2016.

[88] L. Xia and V. Conitzer. A maximum likelihood approach towards aggregating partial orders. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 446–451, 2011.

[89] L. Xia, V. Conitzer, and J. Lang. Aggregating preferences in multi-issue domains by using maximum likelihood estimators. In *Proceedings of the 9th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 399–408, 2010.