TU WIEN Informatics

Perpetual Individualized Fair Rank Aggregation

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Logic and Computation

eingereicht von

Kaan Unalan, BSc
Matrikelnummer 11907099

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Dipl.-Ing. Dr.techn. Stefan Woltran
Mitwirkung: Oliviero Nardi, MSc

Wien, 29. Juli 2025

Kaan Unalan | Stefan Woltran

TU WIEN Informatics

Perpetual Individualized Fair Rank Aggregation

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Logic and Computation

by

Kaan Unalan, BSc
Registration Number 11907099

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Dipl.-Ing. Dr.techn. Stefan Woltran
Assistance: Oliviero Nardi, MSc

Vienna, July 29, 2025

Kaan Unalan   Stefan Woltran

Erklärung zur Verfassung der Arbeit

Kaan Unalan, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.
Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel" habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, habe ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT-Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 29. Juli 2025

Kaan Unalan

Acknowledgements

I would like to thank everyone, every institution, and every organization who contributed to this thesis, perhaps without even knowing it. As naming everyone is beyond the scope of the acknowledgements section of this thesis, I will focus on the people directly involved.

First of all, I thank to Stefan Woltran for his guidance and years of excellent cooperation, not just as part of this thesis but in general during my time at TU Wien starting from the Bachelor's thesis. I would also like to thank Oliviero Nardi for his valuable support and time. He always provided high-quality feedback and was always available.

Of course, I am grateful to my family, particularly my mother Nesrin and my father Gökhan. Although they were not directly involved in the writing of this thesis, they always ensured that I could stay focused and motivated throughout the entire process. A special thanks goes to my sister Ceyda, who also reviewed this section and English and German abstracts of this thesis.

Lastly, I thank my friends for their understanding during my recent unavailability.

Kurzfassung

Kollektive Entscheidungsprozesse sind oft kein isoliertes einmaliges Ereignis, sondern wiederholen sich häufig, wobei sich die Wählerschaft, die Präferenzen und die Alternativen ständig ändern. In dieser Diplomarbeit untersuchen wir eine bestimmte Variante des Problems der *Rank Aggregation*. Wir stellen ein Modell vor, das die zeitlichen Aspekte eines Präferenzaggregationsszenarios berücksichtigt. Um Fairness zu gewährleisten, enthält unser Modell Mechanismen zur Aufrechterhaltung einer proportionalen Vertretung im Laufe der Zeit.

Wir erweitern diese Konstellation, indem wir einen besonderen Wähler einführen, der eine besondere Bedeutung und daher einen größeren Einfluss auf das Ergebnis hat. Wir passen unser Modell an die Situation an, in der der besondere Wähler andere Anforderungen als die anderen hat, sodass die Proportionalität unter Wählerinnen und Wählern weiterhin gewahrt bleibt.

Zu diesem Zweck erweitern wir den Formalismus des *Perpetual Votings* auf die *Rank Aggregation*. Wir führen neue Aktualisierungsregeln (AR) ein und analysieren deren Verhalten in Kombination mit drei bekannten *Social-Preference-Funktionen* (SPFs): Borda, Kemeny und Squared Kemeny. Um die Fairness zu bewerten, entwickeln wir axiomatische Eigenschaften. Unsere theoretische Analyse zeigt, dass verschiedene Konfigurationen unterschiedliche Stärken und Einschränkungen haben.

Darüber hinaus untersucht unser Modell mehrere Strategien zur Initialisierung der Gewichte, einschließlich Szenarien, in denen der besondere Wähler priorisiert wird. Durch Simulationen bewerten wir, wie sich verschiedene Kombinationen von AR und SPFs in Bezug auf langfristige Proportionalität auswirken. Unsere Ergebnisse zeigen, dass die Wahl der SPF in Bezug auf Fairness wichtiger ist als die der AR, aber einige AR sind dennoch vielversprechender als andere. In einigen Fällen beobachten wir, dass axiomatische Ergebnisse das reale Durchschnittsverhalten nicht vollständig widerspiegeln.

Abschließend demonstrieren wir die praktische Relevanz unseres Modells, indem wir einen Entwurf für fairnessorientierte Empfehlungssysteme mit mehreren Interessengruppen vorstellen. Diese Arbeit liefert einen theoretischen Beitrag zum Verständnis von Fairness bei sich ständig ändernden kollektiven Entscheidungsprozessen und unterstützt die Entwicklung gerechter algorithmischer Systeme in der Praxis.

Abstract

In collective decision making, the process is often not an isolated one-time event. It is often repeated, with evolving voters, voter preferences, and alternatives. We study a particular variant of the rank aggregation problem, focusing on decision-making over time. We introduce a model that considers the temporal aspects of a preference aggregation scenario that produces a ranking in each round while taking the evolving voter preferences into account. In order to ensure fairness, our model incorporates mechanisms to maintain proportional representation over time.

We further extend this setting by introducing a special voter who has special importance and, consequently, greater influence over the outcome. We adapt our model to the setting where the special voter has different representation requirements compared to others, while still maintaining the proportionality among voters.

To this end, we extend the formalism of *perpetual voting* to rank aggregation. Our formalism adapts update rules from perpetual voting, which is based on approval ballots, and also introduces novel update rules for ranking-based outcomes. We analyze the behavior of the update rules in combination with three prominent social preference mechanisms: Borda, Kemeny, and Squared Kemeny. To evaluate fairness, we develop axiomatic properties that capture proportionality and equitable influence across time. Our theoretical analysis shows that different settings have different strengths and limitations.

Furthermore, our framework explores multiple weight initialization strategies, including scenarios where the special voter is prioritized. Through simulations, we assess how various combinations of update rules and social preference mechanisms perform with respect to long-term fairness and proportionality. Our results indicate that the choice of the social preference mechanism is more important than the update rule in terms of fairness but some update rules are still more promising than others. In some cases, we observe that axiomatic results do not always capture the real average behavior in its full complexity.

Finally, we demonstrate the practical relevance of our model by proposing a design for *fairness-aware multi-stakeholder recommender systems*. Our approach serves as a design concept for recommendation algorithms that aim to balance fairness and different stakeholders' objectives with user-centric accuracy. This thesis thus contributes both to the theoretical understanding of fairness in evolving collective decisions over time and to the practical development of equitable algorithmic systems.

Contents

CHAPTER 1

Introduction

In numerous real-world scenarios, collective decisions must be made based on the preferences of a group. Examples include elections, committee selections, university rankings, or identifying the most livable cities. These decision-making processes are rarely isolated or one-time events; instead, they are often repeated over time, with evolving preferences, voters, and alternatives.

In some cases, it is advantageous to take past decisions into account. This often helps to increase the diversity and proportionality [34, 42, 23, 39] in order to satisfy as many voters as possible instead of letting a majority group to dominate the outcome in every round. This ensures a more balanced representation of voter preferences in the final rankings observed over time.

Consider, for example, a music streaming service that publishes a monthly ranking of songs based on users' ranked preferences. Each week, listeners submit a ranking of new songs, and the platform aggregates these into a single chart. However, the platform also wants to ensure that different listener groups — defined, for example, by region, musical taste (e.g., rock, hip-hop, classical), or engagement level — are represented fairly over time.

In many cases, the goal is not merely to identify a single winner but to produce a ranking of alternatives based on aggregated preferences. This is particularly relevant in applications such as city rankings, hotel recommendations, or university evaluations, where not only one winner or a set of winners is to be determined but the ranking of the alternatives is important. The problem of how to combine individual preferences into a collective ranking is known as *rank aggregation*, a fundamental topic in computational social choice [8, 9, 40], which aims to develop accurate and efficient algorithms for collective decision-making.

In many settings, it is preferable to model each voter as having equal weight. This simplifies the model and satisfies further desirable properties such as anonymity, where replacing one voter by another does not cause any change in the outcome. Moreover, equal representation of each voter is considered to be a core principle in democratic elections. Nevertheless, there are use cases where a special voter with a greater influence exists. In this case, it may be desirable to assign a higher weight to this special voter.

Consider a scenario where one voter has a higher influence than the others but the opinions of all voters are still considered. One might ask why this voter has to be treated differently in a formal model, given that it is only a normal voter with higher weight. However, over time, voter weights can be adjusted by a weight update mechanism from round to round in order to maximize the degree of proportionality. By explicitly distinguishing the special voter, the update mechanism can also treat it differently. A concrete example is a city that allows residents to vote on some local projects by ranking them, where the mayor of the city — acting as a special agent — has a higher influence on the outcome. This scenario is related to *participatory budgeting* [12] but differs in that it does not necessarily require a fixed budget constraint.

This thesis investigates a specific variant of rank aggregation that operates over time. We consider a setting where the votes, which are linearly ordered ranked alternatives, can be updated in each round, and one special voter is treated differently by receiving a stronger influence in the aggregation process. We assign an enhanced initial weight to the special voter, update agent weights dynamically based on a satisfaction measure, and aggregate rankings using one of the following methods: (i) Borda, (ii) Kemeny, (iii) Squared Kemeny [46]. We refer to this problem as *perpetual rank aggregation*, inspired by perpetual voting introduced by Lackner [42].

Although we primarily study scenarios where voters and alternatives remain constant across rounds, most of the results are also applicable to settings with changing voters and alternatives, as the model can easily be generalized to cases, where these parameters change over time.

The study of such a voting setting not only provides theoretical insights into weighted perpetual rank aggregation problems but also models many real-world use cases. In this chapter, we begin with motivating examples that introduce the problem of rank aggregation and demonstrate its practical relevance (in Section 1.1). We then present the key research questions and contributions of the thesis in Section 1.2, followed by an overview of the thesis structure in Section 1.3. The chapter concludes with a discussion of related work in Section 1.4.

1.1 Motivation: Potential Real-World Applications

Consider again the music streaming service that publishes a monthly ranking of songs based on users' ranked preferences, introduced at the beginning of the chapter.

This example can be modeled in a way that users are voters and songs are alternatives. Voting is repeated weekly so there is also a temporal aspect because the former decisions are decisive in determining the next rankings and listener preferences evolve: new songs arrive, tastes shift, and communities grow or shrink. As rankings, and not a set of winners, are selected, we are dealing here with the problem of *rank aggregation*. Moreover, since we model the scenario as a rank aggregation problem, we do not require predefined groups but can infer them from (evolving) preferences and respond to changes accordingly.

Furthermore, listeners have different interests based on factors such as age, region, and musical taste. These differences must be considered in terms of proportionality over time. For example, the rankings should not always reflect only the preferences of the largest listener group such that they always have the most songs in the top 10. Instead, less preferred rankings (e.g., classical music listeners) should also be represented more than larger groups in some rounds.

Therefore, ensuring proportionality within a single round is sometimes insufficient. If all users (e.g., both pop listeners and classical music listeners) must share the charts every week, neither group may be fully satisfied. On the one hand, pop listeners might feel underserved in the long run if non-pop music consistently appears in the charts, despite their majority status. Classical music listeners, on the other hand, may never see a chart dominated by their preferences, leading to dissatisfaction. This may drive those users away. Instead, the service seeks an over-time proportionality guarantee: across any sequence of, say, 15 weeks, each listener group should dominate roughly the same number of weeks in the top 10 as their share of the total vote.

In addition, we can also have premium listeners, who pay more for the platform and they can be represented by a special voter, with a higher initial weight and special weight update mechanisms.

This scenario illustrates why proportionality over time — not just fairness per round — is essential whenever repeated rankings determine users' experiences. Without such a mechanism, minority or niche listener communities would lose visibility, undermining both user satisfaction and the platform's long-term engagement goals.

We conclude this section with the sketch of a second motivating example, this time from the domain of recommender systems.

A *recommender system* is a tool that suggests items to a user. In the age of big data and information overload, recommender systems support users in making decisions across many areas such as which news to read, which products to buy, and which films to watch [53].

Fairness and achievement of objectives in recommender systems are complex and challenging problems [27], as there are diverse goals and fairness concerns, each relevant to different stakeholders, and these interests sometimes cannot be aligned [2, 1]. Moreover, fairness should be addressed over time in dynamic settings, where the preferences of stakeholders or their interactions with the system and items may evolve over time [3, 36]. In addition, recommender systems emphasize individualized recommendations for each user. Therefore, the problem of multi-stakeholder fair recommendation can be modeled by our framework of *Perpetual Individualized Fair Rank Aggregation* that we introduce in this thesis. Details of the concept are provided in Chapter 5.

1.2 Contributions

The goal of this thesis is to answer the following questions:

1. How should the rank aggregation process be configured to achieve an outcome that maximizes and even guarantees a satisfactory representation for every voter over time?

2. What are the axiomatic properties that define a fair outcome and how can fairness be measured experimentally?

3. How does the special voter interact with the overall rank aggregation process?

4. How do the diffrent rank aggregation configurations behave on average and in comparison to each other in empirical evaluations?

To answer these questions, this thesis makes the following three main contributions:

First, it advances the state of the art in perpetual voting by extending the existing formalism, which is based on approval ballots, to the more complex setting of rank aggregation. An axiomatic analysis gives deeper insights into the properties of perpetual rank aggregation mechanisms. In particular, it investigates notions of proportionality and a specific form of manipulability, laying a solid theoretical foundation for future work in this area. Additionally, the model is extended by a special voter, having distinct representation requirements. To meet this requirements but still be able to maintain the proportional representation for other voters, new update rules that satisfy the corresponding properties are introduced.

Second, the thesis conducts simulations to investigate the empirical behavior of the framework under various parameter configurations. A range of evaluation metrics is used to assess aspects such as equality of voter influence and proportional representation. The experimental results reveal important interactions between update mechanisms and aggregation rules, providing practical evidence to complement the theoretical analysis.

Finally, the thesis introduces a novel system design for potential applications of multi-stakeholder recommender systems. We propose a design concept based on our social choice model in which a special agent ensures individualized outcomes. We also provide an overview of current approaches to make use of computational social choice as a tool to improve recommender systems and compare them to our proposed system.

1.3 Structure of the Thesis

This thesis is structured as follows:

This chapter provides a general introduction, including the problem motivation (Section 1.1), main contributions (Section 1.2), and related work (Section 1.4).

Chapter 2 offers a brief introduction to computational social choice in Section 2.1, including its historical development and a brief introduction to voting theory. Readers familiar with the basics of computational social choice may skip directly to Section 2.2, where the relevant background information and notation used throughout the thesis are introduced, followed by perpetual rank aggregation (2.3) and a brief discussion of complexity theory (Section 2.4).

Chapter 3 presents the perpetual individualized fair rank aggregation framework and studies several social preference rules and weight update mechanisms, including the ones that prioritze the representation of the special voter. Section 3.1 introduces the temporal social preference functions and Section 3.2 provides the axiomatic analysis.

Chapter 4 presents the simulations and experimental evaluation of different configurations. Section 4.2 introduces the evaluation metrics, Section 4.3 formulates hypotheses, Section 4.4 presents the results and Section 4.5 interprets the results in consideration of the hypotheses.

Chapter 5 showcases a potential application for our model in recommender systems.

Finally, Chapter 6 concludes with a discussion of results and outlines directions for future research.

1.4 Related Work

Rank aggregation is one of the classical problems of computational social choice and has been widely studied for over half a century [8, 9]. Recently, proportional rank aggregation has attracted considerable attention in the literature [55, 46], where each voter is represented proportionally in the final ranking. Skowron et al. [55] study proportionality for rankings, where rank aggregation rules are based on approval ballots. Lederer et al. [46] demonstrate the proportionality perspective of the Squared Kemeny rule, which has been introduced by John Kemeny in [40] together with the Kemeny rule, which is less suitable for proportional rank aggregation.

Proportional rank aggregation has also been studied in [54] and [11]. Although it is not the main focus of [54], Schulze discusses proportionality for linearly ordered preference ballots, particularly in context of single transferable vote (STV) but does not compare different rules and definitions of proportionality measurement. Aziz et al. [11] show some proportionality results for multi-winner voting and rank aggregation in context of *committee monotonicity*.

In addition, proportionality is also studied in multi-winner voting, where a subset of candidates has to be selected. In [31], many techniques and challenges including proportionality requirements are summarized (surveyed). Axiomatization of proportional representation requirements for multi-winner voting with approval ballots are formulated in [10, 32].

Participatory budgeting is an extension of multi-winner voting where each alternative has a cost and the total cost of the selected alternatives should not exceed the given budget. Therefore, it is possible to study proportionality in participatory budgeting, as it has been done by Peters et al. [51].

On the other hand, many real-world voting scenarios involve several different rounds of elections, making it increasingly interesting to study social choice in dynamic settings, where voters' opinions may change and proportional representation should be maintained over time.

In [49], the voter preferences evolve over time based on past decisions and the dynamic social choice setting is modeled by a Markov decision process.

Tennenholtz [56] introduces dynamic voting and studies three axioms but ignores proportionality. Moreover, voter preferences do not change but new voters are added in each iteration. Tennenholtz defines continuity and monotonicity in terms of newly joining voters.

Freeman et al. [34] allow evolving preferences. However, they model the problem without any ballots but with valuations for every alternative (allocation problem). Then, the goal is to maximize Nash welfare.

Lackner [42] introduces perpetual voting, where voters submit approval ballots iteratively and preferences evolve over time. The axioms and some weight update mechanisms studied in our work are also adapted from [42] to our rank aggregation settings. In addition, we extend our framework by a special voter. In [43, 44], the framework of perpetual voting is studied in more detail, particularly in the context of proportionality.

Chandak et al. [23] use also approval ballots. However, they also study offline rules — rules where all rounds and preference rankings are known beforehand — and consider axioms that consider proportionality of voter groups instead of individual voters.

Isreal and Brill [39] produce in contrast to the model of perpetual voting a ranking out of approval ballots. In their model, alternatives are removed once they have been selected. They also introduce a new axiom called satisfaction monotonicity.

Although recommender systems have been seen as a potential application domain for Computational Social Choice for years [28], the work in the intersection of both fields is limited and focuses mainly on group recommendation [59, 47, 14]. However, there are also some papers proposing the application of social choice theory to evaluate or achieve fairness in recommender systems [50, 3, 5, 4].

Pennock et al. [50] provide a formal study of collaborative filtering in terms of social choice axioms. The system introduced by Aird et al. in [3, 5, 4] is similar to our idea. They also view the different fairness concerns from a multi-stakeholder perspective and propose a system that uses first an allocation problem and applies a social choice mechanism afterwards.

CHAPTER 2

Background

In this chapter, we introduce the foundational concepts, terminology, and notation required to understand the material presented in the rest of the thesis. In Section 2.1, we begin with a general overview of the field of *computational social choice*, including its historical development and a brief introduction to voting theory. Readers already familiar with these topics may skip directly to Section 2.2, where we define the problem of *rank aggregation* and present the notation used throughout the thesis. We conclude the chapter with a brief overview of concepts from *computational complexity theory* in Section 2.4. While this thesis relies only on basic notions from that field, references are provided for readers interested in exploring the topic further.

2.1 Computational Social Choice

Preferences are a fundamental concept studied in many disciplines including economics, philosophy, psychology, operational research, and computer science. In computer science, fields such as artificial intelligence, databases, or recommender systems frequently deal with preference-related questions, as they often involve analyzing decision-making and choice behavior [52].

Computational social choice is an interdisciplinary field at the intersection of computer science and economics. It studies the algorithmic and computational aspects of aggregating of individual preferences into a collective decision. Typically, it involves formalizing preference aggregation problems, designing efficient solution procedures, and analyzing whether these methods satisfy desirable properties specified axiomatically. The field encompasses several branches such as *voting theory*, *fair allocation*, *matching under preferences*, or *judgement aggregation* [21].

One of the core areas of computational social choice is *voting theory*, which is concerned with selecting one or more options — called alternatives — based on the preferences of a group of voters. A wide variety of voting rules have been proposed and studied in the literature, each with different axiomatic properties [20].

Fair allocation addresses the problem of distributing goods among agents who have preferences over those goods. The goal is to achieve allocations that satisfy criteria such as *Pareto optimality* — no allocation possible such that an agent ends up in a better situation without worsening another agent's situation — and *envy-freeness* — no agent wants to get the allocation of another agent instead of their own — [24]. Numerous algorithms and mechanisms have been developed for different variants of the fair allocation problem, such as allocating indivisible goods or cake cutting algorithms [57, 58].

Matching under preferences deals with problems where sets of agents or entities on two sides (e.g., students and universities, workers and companies) are paired in a stable and efficient way based on the preferences of agents [41]. The field originates from the seminal paper by Gale and Shapley [35], in which the stable marriage algorithm was introduced. This algorithm efficiently solves two-sided matching problems (matching problems that take both sides' preferences into account).

Judgment aggregation, on the other hand, focuses on aggregating logically interrelated individual judgments (e.g., yes/no decisions of propositional nature) from multiple agents into a consistent collective judgment [25]. It is relevant in domains such as legal reasoning, political philosophy, and also in artificial intelligence, mainly in systems involving autonomous software agents [30]. It is closely related to *abstract argumentation* [26], another tool from the field of artificial intelligence concerned with modeling of arguments and attack relations between them. There is already ongoing research exploring the potential combination of both concepts, particularly in aggregating conflicting arguments from multiple agents [22, 17].

2.1.1 History of Computational Social Choice

Computational social choice is a relatively new field. However, the study of voting rules and preference aggregation has a long history, dating back several centuries. Early voting rules have been proposed centuries ago, with notable contributions from thinkers like Borda and Condorcet from the end of 18th century [33].

The modern theory of social choice starts in the middle of the 20th century, when Arrow [7, 8, 9], Black [16], and Guilbaud [37] published foundational work on social-choice-related problems. Unlike earlier approaches, they introduced a normative mathematical framework for social choice theory [33]. However, these early formal models focused mainly on axiomatic properties and desirable criteria, often ignoring the computational resources required to compute the results [21].

With the rise of complexity theory and increasing computational power, social choice theory has also become interesting for computer scientists. This led to the emergence of *computational social choice*, a field that applies algorithmic and complexity-theoretic methods to social choice problems. In the age of the internet — starting from the late

1990s and early 2000s — many new application scenarios emerged, both within computer science and directly through novel computational social choice problems and techniques [21].

Nowadays, computational social has two main goals: to use methods from algorithmics, complexity theory, and computer science in general in order to develop efficient algorithms for typical social choice problems, and to understand the computational complexity of these problems. At the same time, methods developed for social choice theory are increasingly aimed to be applied to solve problems in computer science and artificial intelligence [24, 21]. Some examples include problems in multi-agent systems [29], group recommendation [47], real-world matching mechanisms such as college admissions [15], or several preference aggregation scenarios on the internet [19].

2.1.2 Voting Theory

In *voting theory*, different types of elections are studied depending on the application scenario. The most basic setting is when only one candidate is to be selected from a set of candidates. These elections are referred to as *single-winner voting*. A presidential election is a common example. In contrast, in *multi-winner voting*, also known as *committee election*, the goal is to select a subset of candidates rather than a single winner. Parliamentary elections typically follow this model, where a committee or a parliament consisting of multiple representatives is elected [31].

Rank aggregation, however, involves selecting an entire ranking of the alternatives, not just one or several winners. This setting is particularly relevant in applications where the overall ordering of candidates matters — for example, ranking sports teams or athletes, where we are not only interested in who comes first but also in the relative positions of all participants.

There are also various ways for voters to express their preferences. A common method involves *preference rankings*, where each voter provides an ordering of the candidates [20]. Another method is *approval voting*, where voters submit a ballot indicating which candidates they approve of, without providing a ranking [45]. In *majority judgment*, voters assign a grade to each candidate (e.g., "excellent", "good", "acceptable", etc.), and the final outcome is the highest median grade [13].

Each election requires voting rules that fit the type of the concrete election. For example, in parliamentary elections, it may be desirable to achieve *proportional representation*, ensuring that groups of voters with different preferences are represented in the outcome proportionally to their size. In contrast, in *excellence-based selection* [31] — such as when choosing a set of job candidates for interviews — the focus is on selecting the most promising individuals, possibly without regard for proportionality. Both of these examples are instances of multi-winner voting, but they illustrate how different contexts can require divergent features properties for the voting rule.

In this thesis, we focus on voting theory, and in particular on the specific problem of *rank aggregation*.

2.2 Rank Aggregation

Rank aggregation is one of the oldest problems in computational social choice [8, 9]. It combines the preference rankings of the voters to one single ranking and is used in scenarios, where not only the winner(s) but also the entire ranking is important.

We are interested in rank aggregation in this thesis so we define the problem of rank aggregation in the following: Let $N = \{v_1, \ldots, v_n\}$ be a finite set of *voters (agents)*, $v_n \in N$ is the *special voter*. The set $A = \{a_1, \ldots, a_m\}$ is the set of *alternatives (candidates)*. Each voter v has a *ranking* $\succ_v$ of alternatives from A, which is a total order. A *preference profile* $P = (\succ_1, \ldots, \succ_n)$ contains preference rankings of voters in N. $\mathcal{L}(A)^n$ is the set of all such profiles.

A *social preference function (SPF)* is a map $f : \mathcal{L}(A)^n \to \mathcal{L}(A)$ that returns a preference ranking $\succ_{out}$ for each given profile $P \in \mathcal{L}(A)^n$. The returned preference ranking $\succ_{out}$ is called the *output ranking*.

An SPF f satisfies *continuity* if for every pair of preference profiles $P \in \mathcal{L}(A)^\ell$, $P' \in \mathcal{L}(A)^m$ and for every output ranking $\succ_{out} \in \mathcal{L}(A)$ with $f(P') = \succ_{out}$, there exists a $k \in \mathbb{N}$ such that $f(P \cup P'^{\cup j}) = \succ_{out}$ for all $j \geq k$, where $P'^{\cup j}$ denotes the profile P' repeated j times. This repetition can be interpreted as either a proportionally larger group of voters or a group with higher total weight. Intuitively, continuity means that a group of voters with sufficiently large total weight or size can dominate the outcome.

The *Kendall tau metric* d_K *(KT distance)* measures the number of disagreements between two preference lists $\succ$ and $\succ'$:

$d_K(\succ, \succ') = |\{(a, b) \in A^2 \mid a \succ b \wedge b \succ' a\}|.$

We use this metric to compare the similarity between any two preference rankings $\succ, \succ'$.

The *rank* of an alternative $a \in A$ in the preference ranking $\succ_v$ of voter $v \in N$ is denoted by $r_v(a)$.

For the experiments, we also use another metric, the *Spearman footrule distance*, to compare it with the Kendall-tau metric. Spearman footrule distance measures the difference between two preference lists $\succ_v$ of voter v and $\succ_u$ of voter u as the sum of displacements of each alternative in one list compared to the other:

$d_S(\succ_v, \succ_u) = \sum_{a \in A} |r_v(a) - r_u(a)|.$

2.3 Perpetual Rank Aggregation

The perpetual rank aggregation problem extends conventional rank aggregation by adding time as a parameter.

Let $T = \{1, \ldots, o\}$ be a set of *time steps (rounds)*. $\succ_v^t$ denotes the preference ranking of voter v in round t, $\succ_{out}^t$ denotes the *output ranking* in round t.

In temporal setting, we introduce weights to make sure that the satisfaction is propotionally distributed among different rounds. Let a *weight function* take as input a sequence of preference profiles, the history of past output rankings, a voter $v \in N$, and the current round $t \in T$, and return a real number. This can be interpreted as the weight of voter v in round t. Throughout the thesis, $\omega(t, v)$ denotes the weight of voter v at time step t.

A *temporal social preference function* (TSPF) is an extension of an SPF that takes as input a sequence of preference profiles, the history of past output rankings, a weight function, and the current round $t \in T$, and returns a pereference ranking. This is the output ranking $\succ_{out}^{t}$ at time step t.

Accordingly, a TSPF satisfies *continuity* if its corresponding SPF satisfies continuity.

2.3.1 Social Preference Functions

We define the SPFs used throughout this thesis. We introduce them in this section and not in the previous one, as we extend the definitions by incorporating a temporal component.

The rules that we introduce — with the exception of Random Serial Dictatorship — are not *resolute*, i.e., the ranking does not have to be a total order. In order to obtain resolute rules, the following tie-breaking mechanism is applied throughout the thesis, consistent with the notion that the special voter has stronger influence: The special voter's preference ranking decides the tie-break. Specifically, if there are two tied alternatives $a_1, a_2 \in A$, then $a_1 \succ_{out} a_2$ if $a_1 \succ_n a_2$ and $a_2 \succ_{out} a_1$ otherwise, where $\succ_{out}$ denotes the output ranking and $\succ_n$ the preference ranking of the special voter. Even if there is no special voter in a particular setting, we still assume that voter n's preference determines the tie-break — at least in our experiments in Chapter 4 — as we want to observe whether this tie-breaking mechanism affects the average results. However, other mechanisms such as random tie-breaking are also possible.

Borda Rule

For a set of voters N, a corresponding preference profile P, the *Borda score* for an alternative $a \in A$ is defined as follows:

$$Borda(a) = \sum_{v \in N} m - r_v(a)$$

The definition can be extended with a temporal component as follows. For a set of voters N, a corresponding preference profile P, and a set of rounds T, the *Borda score* for an alternative $a \in A$ and time step $t \in T$ is defined as:

$$Borda(a, t) = \sum_{v \in N} m - r_v^t(a)$$

The *Borda* rule is an SPF that ranks the alternatives according to their Borda score decreasingly.

Kemeny Rule

Let $\succ^t_{out}$ be an arbitrary output ranking in round t.

The *Kemeny* rule returns the set of output rankings with the minimal $\sum\limits_{v \in N} d_K(\succ^t_v, \succ^t_{out})$:

$$\arg \min_{\succ^t_{out} \in \mathcal{L}(A)} \sum_{v \in N} d_K(\succ^t_v, \succ^t_{out})$$

Squared Kemeny Rule

Let $\succ^t_{out}$ be an arbitrary output ranking in round t.

The *Squared Kemeny* rule [40, 46] returns the set of output rankings with the minimal $\sum\limits_{v \in N} d_K(\succ^{t_k}_v, \succ^t_{out})^2$:

$$\arg \min_{\succ^t_{out} \in \mathcal{L}(A)} \sum_{v \in N} d_K(\succ^t_v, \succ^t_{out})^2$$

Random Serial Dictatorship Rule

The *Random Serial Dictatorship* rule chooses the ranking of a randomly selected voter $v \in N$ as output ranking s. t. $\succ^t_v = \succ^t_{out}$ for an arbitrary round $t \in T$.

2.4 Complexity Theory

Computational complexity theory classifies problems according to required computational resources to solve them [6].

Throughout this thesis, we rely on standard complexity classes to classify the computational difficulty our model. Complexity-theoretic considerations are discussed only in Section 3.3, where we evalute the complexity of our model. However, this is not the main focus of this thesis, as the complexity of the components that build up our model is already known, and combining these results to derive the overall complexity is straightforward. Therefore, the discussion of complexity is kept at an informal level.

The only complexity classes that we need in this thesis are P and Θ^P_2. We expect that the reader is familiar with the basics of complexity theory. Otherwise, the main takeaway is that problems in class P — problems solvable in polynomial time — are considered tractable, whereas problems in Θ^P_2 are generally regarded as intractable.

For a detailed introduction to these concepts, we refer the reader to comprehensive textbooks on complexity theory, such as [48, 6].

CHAPTER 3

The Perpetual Individualized Fair Rank Aggregation Model

This is the chapter where we introduce our model. First, we define the temporal social preference functions (TSPFs) in Section 3.1. We then present four axiomatic properties (perpetual lower quota, simple proportionality, independence of uncontroversial decisions, and bounded dry spells) and related theorems in Section 3.2. The axiomatic properties are adapted from Lackner [42].

3.1 Temporal Social Preference Functions

We have already defined what a temporal social preference function (TSPF) is (see Section 2.3). Now, we delve into the details and introduce some concrete TSPFs. When selecting a TSPF, three parameters are important:

- Weight initialization strategy
- Social preference function
- Weight update mechanism

3.1.1 Weight Initialization Strategy

The weight of each voter should be initialized with a specific value in round $1 \in T$. In other words, we should assign a weight to each voter $v \in N$ in round 1 such that there exists a $w \in \mathbb{Q}$ with $\omega(1, v) = w$ for each v. Of course, each voter may be assigned the same weight but differences between voter weights are also possible. This initial weight assignment is referred to as *weight initialization strategy*, which can significantly influence output rankings in each round, particularly in the early rounds and if it is used in combination with a suitable social preference function and a weight update mechanism.

For the initialization of weights, we study the following configurations:

- Equal weights: $\omega(1, v) = 1$ for each $v \in N$
- Special voter n's weight about 25% of the total weight: $\omega(1, n) = \left\lceil \frac{\sum_{v=0}^{n-1} \omega(1,v)}{3} \right\rceil$

If not specified otherwise, we assume equal weights, as this represents a simpler and more natural setting, particularly when there is no special voter. Note that other configurations are also possible and could be interesting for further study. However, our focus is on examining the difference between scenarios in which the special voter receives significantly greater weight and those in which all voters are weighted equally.

3.1.2 Social Preference Function

The selection of an SPF is crucial in determining the properties and behavior of a TSPF. Since a TSPF operates in a temporal setting, SPFs are applied in their weighted versions in order to ensure adaptability through weight updates after each round.

In this thesis, we study four social preference functions: *weighted Borda*, *weighted Kemeny*, *weighted Squared Kemeny*, and *weighted Random Serial Dictatorship (weighted RSD)*.

Weighted RSD serves as a baseline for experimental comparison with the other rules, but is not analyzed theoretically in this thesis. Weighted Borda is a positional scoring rule and is easy to compute, distinguishing it from weighted Kemeny and weighted Squared Kemeny. However, Kemeny and Squared Kemeny are good at minimizing the Kendall-tau distance, which is used to measure similarity between rankings. While Kemeny primarily selects alternatives preferred by the majority, it can adapt to proportionality over time. Squared Kemeny is more sensitive to underrepresented opinions [46] and generates rankings that better account for them.

The *weighted Borda score* for an alternative $a \in A$ and a time step $t \in T$ is defined as follows:

$$BordaW(a, t) = \sum_{v \in N} \omega(t, v) \cdot (m - r_v^t(a)) \tag{3.1}$$

The *weighted Borda rule* is an SPF that ranks the alternatives according to their weighted Borda score decreasingly.

The *weighted Kemeny rule* returns the set of output rankings with the minimal $\sum_{v \in N} \omega(t, v) \cdot d_K(\succ_v^t, \succ_{out}^t)$:

$$\arg \min_{\succ_{out}^t \in \mathcal{L}(A)} \sum_{v \in N} \omega(t,v) \cdot d_K(\succ_v^t, \succ_{out}^t) \tag{3.2}$$

The *weighted Squared Kemeny rule* returns the set of output rankings with the minimal $\sum_{v \in N} \omega(t,v) \cdot d_K(\succ_v^t \succ_{out}^t)^2$:

$$\arg \min_{\succ_{out}^t \in \mathcal{L}(A)} \sum_{v \in N} \omega(t,v) \cdot d_K(\succ_v^t, \succ_{out}^t)^2 \tag{3.3}$$

The *weighted random serial dictatorship rule* chooses the ranking of a randomly selected voter $v \in N$ as output ranking s. t. $\succ_v^t = \succ_{out}^t$ for an arbitrary round $t \in T$, where the probability of selecting $\succ_v^t$ in round t is v's relative weight in round t:

$$\frac{\omega(t,v)}{\sum\limits_{u \in N} \omega(t,u)}$$

3.1.3 Weight Update Mechanism

To maintain fairness and proportionality over time, a TSPF must not only make good decisions in each round but also adapt to past outcomes. This is achieved through a weight update mechanism, which adjusts the weight of each voter after every round based on their satisfaction. The central idea is that voters who have been less satisfied in earlier rounds should gain more influence in future ones, ensuring a balanced, long-term representation.

The voter weights are updated after each round. It is possible to define several weight update functions but we introduce here some of them that we study in this thesis and demonstrate diverse approaches:

Constant Weight Update Function

The most straightforward approach is to keep the weights fixed over time. In this setting, past decisions have no influence on future rounds, and every voter $v \in N$ maintains the same weight throughout the entire process:

$$\omega(t+1, v) = \omega(t, v) \tag{3.4}$$

This update rule effectively reduces the TSPF to a repetition of the original SPF, as weights do not change from round to round. Since it does not seem to be designed to achieve long-term proportionality, this method primarily serves as a baseline to evaluate the performance and fairness of more dynamic approaches.

Myopic KT-Distance-Based Weight Update Functions

Another intuitive method updates all weights based on the Kendall tau distance between the corresponding voter's ranking and the output ranking, as this distance measures the similarity between two rankings:

$$\omega(t+1, v) = d_K(\succ_v^t, \succ_{out}^t) \tag{3.5}$$

The larger the distance, the more disagreement between a voter's preferences and the outcome, leading to a higher weight in the next round. This reflects the principle that unsatisfied voters should gain more power in the next rounds.

Of course, we can also use the squared Kendall tau distance like in *Squared Kemeny*, which results in a faster increase in the weights of less satisfied voters:

$$\omega(t+1, v) = d_K(\succ_v^t, \succ_{out}^t)^2 \tag{3.6}$$

Both of these methods ignore the past decisions from round 1 to $t-1$ and do not accumulate dissatisfaction over time. Therefore, they are called *myopic weight update functions.*

KT-Distance-Based Weight Update Functions

If the update function is not myopic, past weights can still play a role with the idea of capturing long-term fairness better than myopic weight update functions. In this case, we consider cumulative update rules that take into account not only the current round's outcome but also past satisfaction. The new weight is calculated as the sum of the previous weight and the distance between the voter's ranking and the current outcome:

$$\omega(t+1, v) = \omega(t, v) + d_K(\succ_v^t, \succ_{out}^t) \tag{3.7}$$

Similarly, a squared version emphasizes recurring dissatisfaction more aggressively, and large disagreements are penalized heavier:

$$\omega(t+1, v) = \omega(t, v) + d_K(\succ_v^t, \succ_{out}^t)^2 \tag{3.8}$$

These functions preserve a memory of past disagreements, allowing voters who have consistently been underrepresented to accumulate more weight and thus gain greater influence in future rounds. This makes the KT-distance-based update functions particularly suitable for enforcing long-term proportionality across the entire timeline of decisions.

Reset-Based Weight Update Functions

Reset-based update functions introduce a mechanism to reset the weight of satisfied voters while increasing it for others. Specifically, if a voter's ranking is sufficiently different from the outcome (as measured by the Kendall tau distance exceeding a fixed threshold α), they are considered to be unsatisfied and their weight increases. Otherwise, their weight is reset to a baseline value, often 1:

$$\omega(t+1, v) = \begin{cases} \omega(t, v) + d_K(\succ_v^t, \succ_{out}^t), & \text{if } d_K(\succ_v^t, \succ_{out}^t) > \alpha, \\ 1, & \text{if } d_K(\succ_v^t, \succ_{out}^t) \leq \alpha. \end{cases} \tag{3.9}$$

This version is sensitive to the number of disagreements, using the Kendall tau distance as an update increment.

A more coarse-grained version updates weights uniformly for unsatisfied voters, regardless of the distance, which simplifies influence adjustments:

$$\omega(t+1, v) = \begin{cases} \omega(t, v) + 1, & \text{if } d_K(\succ_v^t, \succ_{out}^t) > \alpha, \\ 1, & \text{if } d_K(\succ_v^t, \succ_{out}^t) \leq \alpha. \end{cases} \tag{3.10}$$

Both variants share the idea of resetting the weight of satisfied voters, thereby giving them equal opportunity in future rounds and reducing their influence temporarily. This approach promotes fairness over time, preventing any voter from dominating the outcome if their preferences are regularly satisfied.

AV-Inspired Weight Update Functions

These weight update functions are inspired by analogous mechanisms from perpetual approval voting and were introduced by Lackner [42]. We adapted them to our ranking-based settings. The central idea is again to adjust weights based on whether a voter is satisfied according to a predefined satisfaction threshold α.

The unit cost weight update function inspired by perpetual unit-cost in [42] increases the weight of unsatisfied voters by a fixed amount (1 in our case), while leaving satisfied voters' weights unchanged:

$$\omega(t+1, v) = \begin{cases} \omega(t, v) + 1, & \text{if } d_K(\succ_v^t, \succ_{out}^t) > \alpha, \\ \omega(t, v), & \text{if } d_K(\succ_v^t, \succ_{out}^t) \leq \alpha. \end{cases} \tag{3.11}$$

Unlike the reset-based variant, this approach does not penalize satisfied voters but rather freezes their influence, allowing underrepresented voters to gradually catch up without reducing others such that a comeback is easier than in reset-based approaches.

The perpetual Kendall-tau weight update function (inspired by Perpetual PAV [42]) reduces the weights of satisfied voters:

$$\omega(t+1,v) = \begin{cases} \omega(t,v), & \text{if } d_K(\succ_v^t, \succ_{out}^t) > \alpha, \\ \frac{\omega(t,v)}{\omega(t,v)+1}, & \text{if } d_K(\succ_v^t, \succ_{out}^t) \leq \alpha. \end{cases} \tag{3.12}$$

Special-Voter-Sensitive Weight Update Functions

There are some use cases where the behavior of a powerful actor, central planner, etc., can be simulated via a special voter. Special-voter-sensitive weight update functions distinguishes the special voter from other voters, and update the special voter's weight according to a separate rule, which mostly favors the special voter.

For instance, the myopic KT maximum value weight update function uses the maximum KT distance across all voters to update the special voter's weight:

$$\omega(t+1,v) = \begin{cases} d_K(\succ_v^t, \succ_{out}^t), & \text{if } v \in N \setminus \{n\}, \\ max(d_K(\succ_1^t, \succ_{out}^t), d_K(\succ_2^t, \succ_{out}^t), \ldots, d_K(\succ_n^t, \succ_{out}^t)), & \text{if } v = n. \end{cases} \tag{3.13}$$

This update function corresponds to the myopic KT-distance-based weight update function 3.5. The only difference is the treatment of the special voter.

The KT sum weight update function emphasizes the representation of the special voter even more, and adds up all the weight gains of the voters in one round and assigns them to the special voter:

$$\omega(t+1,v) = \begin{cases} \omega(t,v) + d_K(\succ_v^t, \succ_{out}^t), & \text{if } v \in N \setminus \{n\}, \\ \sum\limits_{u \in N} \omega(t,u), & \text{if } v = n. \end{cases} \tag{3.14}$$

This update function is an extension of the KT-based function 3.7 with a special treatment of the special voter.

3.2 Axiomatic Analysis

In this section, we introduce axiomatic properties to understand whether the proposed TSPFs are suitable for perpetual proportional voting scenarios and how they differ from each other.

The properties are mainly adapted by Lackner's work on perpetual voting [42] to our settings. As we can easily observe, the concept of winning an approval voting election facilitates the definition of concepts such as satisfaction of a voter, support of a voter by other voters, and all related properties. In our context, there is no one clearly distinguished winner and many losers. This oversimplifies our model and disregards the rank aggregation setting if we consider only the first k alternatives as winners.

On the other hand, if we evaluate the entire lists and require them to be exactly the same in order to "win" or satisfy a voter's preference, this assumption would be too restrictive. Therefore, although other metrics could also be applied, we propose minimizing $d_K(\succ_v, \succ_{out})$ (Kendall tau distance) to "win" or at least approximate a win. However, we also use other distance measures $d_K(\succ_v, \succ_{out})^2$ (squared Kendall tau distance) and $d_S(\succ_v, \succ_{out})$ (Spearman footrule distance) in the empirical part and show that our TSPFs do not only optimize for KT distance.

We define the *satisfaction* of a voter v as follows:

Let satisfaction of a voter v in round t be the number of disagreements under a treshold for previous rankings:

$$sat_t(v) = |\{\ell \leq t : d_K(\succ_v^\ell, \succ_{out}^\ell) \leq \alpha\}| \tag{3.15}$$

The support of a voter $v \in N$ in round t is defined as:

$$supp_t(v) = \frac{1}{|N|} \cdot |\{u \in N \mid d_K(\succ_u^t, \succ_v^t) \leq \alpha\}| \tag{3.16}$$

The support indicates the ratio of voters that have similar preferences to voter v in round t.

This definition forces us to determine a threshold value α, which represents the boundary between satisfaction and dissatisfaction (analogous to the perpetual voting case, where there is a clear distinction between winning and losing). Of course, the challenge lies in the fact that there is no universally correct value for α. This suggests to have a more gradual definition of satisfaction of a voter v in round t. For this purpose, we define the maximal possible d_K in one round:

$$d_{max} = \frac{m \cdot (m-1)}{2} \tag{3.17}$$

Now, we can provide an alternative definition of satisfaction as the sum of differences between the worst case distance and the real distance between voter V's preference list and output list until the current round:

$$sat_o(v) = \sum_{t=1}^{o} d_{max} - d_K(\succ_v^t, \succ_{out}^t) = o \cdot d_{max} - \sum_{t=1}^{o} d_K(\succ_v^t, \succ_{out}^t) \tag{3.18}$$

The support of a voter $v \in N$ in round t is defined as:

$$supp_t(v) = \frac{1}{|N|} \cdot \sum_{u \in N} d_{max} - d_K(\succ_u^t, \succ_v^t) = d_{max} - \frac{1}{|N|} \cdot \sum_{u \in N} d_K(\succ_u^t, \succ_v^t) \tag{3.19}$$

These alternative definitions enable a more fine-grained measurement of similarity between preference rankings. However, the first definition with a threshold value α is still interesting, as it is simpler, closer to the approval voting setting, and α might give some interesting bounds for theoretical analysis.

In the following, we formulate axioms for both approaches, where we call the former definition the *threshold-based definition (TBD)* and the latter definition the *continuous definition (CD)*. However, for the theoretical and empirical analysis, we focus on the TBD for the rest of this thesis. Therefore, our theorems are applicable only to the TBD. Moreover, when we introduce the theoretical results, we assume the satisfaction threshold $\alpha < d_{max}$, as the problem becomes trivial when all voters are satisfied in every round such that $\alpha = d_{max}$.

Quota of voter $v \in N$ in round t is defined as the sum of the supports of v until now:

$$qu_t(v) = \sum_{t=1}^{o} supp_t(i) \tag{3.20}$$

For the theoretical part, we assume equal initial weights to keep the model simple. Different weight initialization strategies are compared only in the empirical part.

3.2.1 Perpetual Lower Quota

A TSPF satisfies *perpetual lower quota* if, for each voter $v \in N$, it holds

$$sat_o(v) \geq \lfloor qu_o(v) \rfloor$$

in the final round $o \in T$, where $o = |T|$. The property requires that a voter whose preferences are supported by $x\%$ of the voters should be satisfied in at least $x\%$ of the rounds. Clearly, this is a very strong proportionality requirement and is generally too difficult to be satisfied. We do not further analyze this axiom directly but instead introduce later in Section 4.2 an empirical relaxation called *perpetual lower quota compliance*, which serves as a metric for the experimental evaluation. Nevertheless, in some cases, the results are strong enough that even the strict perpetual lower quota condition is met, despite not being explicitly intended.

3.2.2 Simple Proportionality

Assume that, for each round $t \in T$, the ranking $\succ_v^t$ is always the same for any voter $v \in N$, where $|T| = |N| = n$. A TSPF satisfies *simple proportionality* if $sat_n(v) = qu_n(v)$ for every voter $v \in N$.

Simple proportionality requires that, for every voter, the overall satisfaction after a certain number of rounds —equal to the number of voters— matches the number of voters with similar preferences up to that round.

This property relies on the following simplifying assumptions: The number of rounds equals the number of voters, the set of alternatives and voters remain unchanged, and voter preferences do not vary over time. However, just as many perpetual voting rules in [42] do not satisfy simple proportionality, numerous TSPFs also fail to satisfy it. In fact, TSPFs struggle more than approval-voting-based perpetual voting, as it is actually a relatively strict property, particularly for the complexity of preference rankings. We introduce therefore additional restrictions such that simple proportionality holds.

Theorem 1. *If an SPF satisfies continuity, the corresponding TSPF with the constant weight update function does not satisfy simple proportionality.*

Proof. Consider the following counterexample: Voter $v \in N$ submits the preference ranking $\succ_v$ and all the other voters $u \in N \setminus \{v\}$ with $|N \setminus \{v\}| = n-1$ submit $\succ_u$, the completely reversed ranking of $\succ_v$. Let f be an SPF that satisfies continuity. By the definition of continuity, there is a k such that if $n-1 \geq k$, f outputs $f(P_1^t \cup P_2^t) = \succ_{out}^t = \succ_u$ for any $t \in T$, where $P_1^t = (\succ_v)$ is a preference profile in round t and $P_2^t = (\succ_{u_1}, \succ_{u_2}, \dots, \succ_{u_{n-1}})$ is another preference profile in round t with $f(P_2^t) = \succ_{out}^t = \succ_u$ and $\succ_{u_1} = \succ_{u_2} = \dots = \succ_{u_{n-1}}$. As $\succ_{out}^t = \succ_u \neq \succ_v$ for any round t, $sat_t(v) = 0$ and $qu_t(v) = 1$ follows (if $\alpha < d_{max}$). Therefore, $sat_t(v) \neq qu_t(v)$, i.e., simple proportionality is not satisfied.

In other words, by the definition of continuity, if there are sufficiently many voters, then for any round t, the output ranking is $\succ_v \neq \succ_u = \succ_{out}^t$. This implies that $sat_t(v) = 0$ and $qu_t(v) = 1$ (for $\alpha < d_{max}$), as in simple proportionality the ranking $\succ_x^t$ is always the same for any voter $x \in N$ for each round $t \in T$. Therefore, $sat_n(v) \neq qu_n(v)$, i.e., simple proportionality is not satisfied.

□

Corollary 1.1. *Borda, Kemeny and Squared Kemeny with the constant weight update function do not satisfy simple proportionality (if $\alpha < d_{max}$).*

Proof. This result follows directly by Theorem 1, as Borda, Kemeny and Squared Kemeny satisfy continuity [46]. □

Theorem 2. *If an SPF satisfies continuity, the corresponding TSPF with a myopic weight update function does not satisfy simple proportionality (if $|N| > 2$).*

Proof. Consider the following counterexample: Voter $v \in N$ submits the preference ranking $\succ_v$ and all the other voters $u \in N \setminus \{v\}$ with $|N \setminus \{v\}| = n-1$ submit $\succ_u$, the completely reversed ranking of $\succ_v$ in every round. Let f be an SPF that satisfies continuity. By the definition of continuity, there is a k such that if $n-1 \geq k$, f outputs $f(P_1^1 \cup P_2^1) = \succ_{out}^1 = \succ_u$ for round 1, where $P_1^1 = (\succ_v)$ is a preference profile in round 1 and $P_2^1 = (\succ_{u_1}, \succ_{u_2}, \dots, \succ_{u_{n-1}})$ is another preference profile in round 1. Then, the weights are updated such that $\omega(2, v) = d_{max}$ and $\omega(2, u) = 0$ for all $u \in N \setminus \{v\}$, as $\succ_v$ is the completely reversed and maximally different ranking from the outcome of the current round. Now consider how a myopic weight update function works: it updates each voter's weight based only on the disagreement between their ranking and the outcome of the current round, without considering the past rounds.

Consequently, in the next round, only $v's$ weight is non-zero so $\succ_v^2 = \succ_{out}^2$. However, the weights are updated again such that $\omega(3, v) = 0$ and $\omega(3, u) = d_{max}$, reversing the situation. This process continues until the last round n and thus the output ranking alternates between $\succ_v$ and $\succ_u$ in each round. As a result, in round n, every voter is satisfied in roughly half of the rounds such that $sat_n(v) \approx \frac{n}{2}$ and $sat_n(u) \approx \frac{n}{2}$. However, since voter v is just a single individual, the quota is $qu_t(v) = \frac{1}{n} \cdot n = 1$, which is much less than $\frac{n}{2}$ as long as $n > 2$. Conversely, for each u, the quota is $qu_t(v) = \frac{n-1}{n} \cdot n = n - 1$, which is much greater than $\frac{n}{2}$. Thus, $sat_n(v) > qu_n(v)$ and $sat_n(u) < qu_n(u)$ for any other voter $u \in N \setminus \{v\}$, which violates simple proportionality. □

Corollary 2.1. *Weighted Borda, weighted Kemeny, and weighted squared Kemeny with myopic weight update functions do not satisfy simple proportionality.*

Proof. This property is implied by Theorem 2, as Borda, Kemeny and Squared Kemeny satisfy continuity. □

Theorem 3. *Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in G_1 submit the ranking $\succ_{G_1}^t$ and all voters in G_2 submit $\succ_{G_2}^t$ for any round t. Then, the Squared Kemeny rule with the constant weight update function satisfies both simple proportionality and perpetual lower quota if the following holds for some round $t \in T$ (WLOG, assume $d_K(\succ_{G_1}^t, \succ_{out}^t) \leq d_K(\succ_{G_2}^t, \succ_{out}^t)$):*

$$d_K(\succ_{G_1}^t, \succ_{G_2}^t) - d_K(\succ_{G_1}^t, \succ_{out}^t) = \left\lfloor (\frac{|G_1|}{|N|} \cdot d_K(\succ_{G_1}^t, \succ_{G_2}^t)) + \frac{1}{2} \right\rfloor \leq \alpha$$

Proof. The property called *2-Rankings-Proportionality* of the Squared Kemeny rule, introduced by Lederer et al. [46], guarantees that the output ranking satisfies the following:

$$d_K(\succ_{G_1}, \succ_{G_2}) - d_K(\succ_{G_1}, \succ_{out}^t) = \left\lfloor (\frac{|G_1|}{|N|} \cdot d_K(\succ_{G_1}, \succ_{G_2})) + \frac{1}{2} \right\rfloor \leq \alpha$$

Analogously it also satisfies the following:

$$d_K(\succ_{G_1}, \succ_{G_2}) - d_K(\succ_{G_2}, \succ_{out}^t) = \left\lfloor (\frac{|G_2|}{|N|} \cdot d_K(\succ_{G_1}, \succ_{G_2})) + \frac{1}{2} \right\rfloor \leq \alpha$$

Since there are only two groups, at most one group can have a smaller KT distance to the output ranking, i.e., it has to agree with more pairs. WLOG, we assume that this group is G_1. If $\alpha \geq d_K(\succ_{G_1}, \succ_{G_2}) - d_K(\succ_{G_1}, \succ_{out}^t) \geq d_K(\succ_{G_1}^t, \succ_{G_2}^t) - d_K(\succ_{G_2}^t, \succ_{out}^t)$ for any round t, it follows that both groups' voters are satisfied in all rounds, as the constant update function does not update the weights such that $\omega(t, v) = 1$ for each round $t \in T$ and voter $v \in N$. Therefore, $sat_o(v) = qu_o(v) = |T|$ for every voter $v \in N$, where $o = |T|$ is the final round, i.e., squared Kemeny satisfies not only simple proportionality but also *perpetual lower quota* (see 3.2.1) under the given condition. □

Proposition 3.1. *Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in G_1 submit the ranking $\succ^t_{G_1}$ and all voters in G_2 submit $\succ^t_{G_2}$ for any round t. Then, the Squared Kemeny rule with any weight update function satisfies both simple proportionality and perpetual lower quota if the following holds for **every** round $t \in T$ (WLOG, assume $d_K(\succ^t_{G_1}, \succ^t_{out}) \leq d_K(\succ^t_{G_2}, \succ^t_{out})$):*

$$d_K(\succ^t_{G_1}, \succ^t_{G_2}) - d_K(\succ^t_{G_1}, \succ^t_{out}) = \left\lfloor (\frac{|G_1|}{|N|} \cdot d_K(\succ^t_{G_1}, \succ^t_{G_2})) + \frac{1}{2} \right\rfloor \leq \alpha$$

Proof. The statement is a generalization of Theorem 3 by allowing any weight update function. Because the weight update rule is no longer constant, weights may vary over time, and the condition must hold for each round $t \in T$ separately. Therefore, it has to be shown that the following condition holds for each $t \in T$ separately (WLOG $d_K(\succ^t_{G_1}, \succ^t_{out}) \leq d_K(\succ^t_{G_2}, \succ^t_{out})$):

$$\alpha \geq d_K(\succ_{G_1}, \succ_{G_2}) - d_K(\succ_{G_1}, \succ^t_{out}) \geq d_K(\succ^t_{G_1}, \succ^t_{G_2}) - d_K(\succ^t_{G_2}, \succ^t_{out})$$

This implies that voters in both groups are satisfied at threshold α in each round t. As a result, every voter $v \in N$ is satisfied in all $|T|$ rounds, regardless of how their weights evolve.

Thus, it follows that $sat_o(v) = qu_o(v) = |N|$ for every voter $v \in N$, where $o = |T|$ is the last round, i.e., squared Kemeny satisfies not only simple proportionality but also *perpetual lower quota* (see 3.2.1).

□

Lemma 3.1. *Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in G_1 submit the ranking $\succ_{G_1}$ and all voters in G_2 submit $\succ_{G_2}$. Then, the output ranking of the Kemeny rule, which is the ranking with the minimal KT-distance, is:*

- $\succ_{out} = \succ_{G_1}$ *if* $|G_1| > |G_2|$,
- $\succ_{out} = \succ_{G_2}$ *if* $|G_1| < |G_2|$,
- *decided by tie-breaking if* $|G_1| = |G_2|$.

Proof. We prove the first case where $|G_1| > |G_2|$. The other two follow symmetrically or via tie-breaking.

Let $\succ'$ be an arbitrary ranking that is not equal to $\succ_{G_1}$, and let:

$$x = d_K(\succ_{G_1}, \succ'), \quad y = d_K(\succ_{G_2}, \succ').$$

The total distance of $\succ'$ to all voters is:

$$D(\succ') = \sum_{v \in G_1} d_K(\succ_{G_1}, \succ') + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ') = |G_1| \cdot x + |G_2| \cdot y.$$

We will construct a new ranking $\succ''$ that reduces the total distance by applying pairwise swaps in $\succ'$ that move it closer to $\succ_{G_1}$ such that $d_K(\succ', \succ'') = 1$ in each step. Consider for instance any pair of alternatives (a, b) for which $a \succ' b$ but $b \succ'' a$, i.e., the pairwise order in $\succ'$ disagrees with $\succ''$. The Kemeny rule selects the ranking with the minimal total distance, so if we can show that this construction leads to $\succ_{G_1}$, we are done. After each such step, there are three cases:

1. **Both groups agree with the swap:** $d_K(\succ_{G_1}, \succ'') = d_K(\succ_{G_1}, \succ') - 1 = x - 1$ and $d_K(\succ_{G_2}, \succ'') = d_K(\succ_{G_2}, \succ') - 1 = y - 1$. This reduces the total distance as follows:

$$\begin{aligned} D(\succ'') &= \sum_{v \in G_1} d_K(\succ_{G_1}, \succ'') + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ'') \\ &= \sum_{v \in G_1} (d_K(\succ_{G_1}, \succ') - 1) + \sum_{u \in G_2} (d_K(\succ_{G_2}, \succ') - 1) \\ &= \sum_{v \in G_1} d_K(\succ_{G_1}, \succ') - |G_1| + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ') - |G_2| \\ &= |G_1| \cdot (x - 1) + |G_2| \cdot (y - 1) < |G_1| \cdot x + |G_2| \cdot y = D(\succ'). \end{aligned}$$

   As $D(\succ'') < D(\succ')$, the ranking $\succ''$ is better than $\succ'$.

2. **Only G_1 benefits:** $d_K(\succ_{G_1}, \succ'') = d_K(\succ_{G_1}, \succ') - 1 = x - 1$ and $d_K(\succ_{G_2}, \succ'') = d_K(\succ_{G_2}, \succ') + 1 = y + 1$. This reduces the total distance as follows:

$$\begin{aligned} D(\succ'') &= \sum_{v \in G_1} d_K(\succ_{G_1}, \succ'') + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ'') \\ &= \sum_{v \in G_1} d_K(\succ_{G_1}, \succ') - |G_1| + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ') + |G_2| \\ &= |G_1| \cdot x - |G_1| + |G_2| \cdot y + |G_2| < |G_1| \cdot x + |G_2| \cdot y = D(\succ'). \end{aligned}$$

   This inequality holds as $|G_1| > |G_2|$ so $\succ''$ is better than $\succ'$.

3. **Only G_2 benefits:** $d_K(\succ_{G_1}, \succ'') = d_K(\succ_{G_1}, \succ') + 1 = x + 1$ and $d_K(\succ_{G_2}, \succ'') = d_K(\succ_{G_2}, \succ') - 1 = y - 1$. This increases the total distance:

$$\begin{aligned} D(\succ'') &= \sum_{v \in G_1} d_K(\succ_{G_1}, \succ'') + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ'') \\ &= \sum_{v \in G_1} d_K(\succ_{G_1}, \succ') + |G_1| + \sum_{u \in G_2} d_K(\succ_{G_2}, \succ') - |G_2| \\ &= |G_1| \cdot x + |G_1| + |G_2| \cdot y - |G_2| > |G_1| \cdot x + |G_2| \cdot y = D(\succ'). \end{aligned}$$

   Since $D(\succ'') > D(\succ')$, the ranking $\succ''$ is worse than $\succ'$, moving us further from the goal.

If we repeatedly apply Step 1 until it is not applicable, we obtain the following total distance for a ranking $\succ'''$ constant $c \geq 0$ indicating the total number of iterations of Step 1:

$$D(\succ''') = |G_1| \cdot (x - c) + |G_2| \cdot (y - c)$$

Then we apply Step 2 until it is not applicable ($x - c$ times) and we end up in an output ranking $\succ_{out}$. Then, we have $D(\succ_{out}) = |G_1| \cdot ((x - c) - (x - c)) + |G_2| \cdot (y - c) - (x - c) = |G_1| \cdot 0 + |G_2| \cdot (y - x)$, i.e., the output ranking corresponds to $\succ_{G_1}$ such that $d_K(\succ_{out}, \succ_{G_1}) = 0$.

The case $|G_2| > |G_1|$ is symmetric and leads to $\succ_{G_2}$ being the minimizer ranking.

If $|G_1| = |G_2|$, then any ranking has the same total distance, and a tie-breaking rule determines the output.

□

Lemma 3.2. *Assume there are two groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in the group G_1 submits $\succ_{G_1}$ and all voters in the group G_2 submits the completely reversed ranking $\succ_{G_2}$ of $\succ_{G_1}$. Then, the output ranking of Borda rule, which is the ranking with the highest Borda score, is:*

- $\succ_{out} = \succ_{G_1}$ *if* $|G_1| > |G_2|$,
- $\succ_{out} = \succ_{G_2}$ *if* $|G_1| < |G_2|$,
- *decided by tie-breaking if* $|G_1| = |G_2|$.

Proof. We prove the first case where $|G_1| > |G_2|$. The other two follow symmetrically or via tie-breaking.

Let $|A| = m$ be the number of alternatives. Assume that all voters in G_1 submit $a_1 \succ_{G_1} a_2 \succ_{G_1} \ldots \succ_{G_1} a_m$. Then, voters in G_2 submit $a_m \succ_{G_2} a_{m-1} \succ_{G_2} \ldots \succ_{G_2} a_1$. The Borda score for the alternatives are as follows:

$$Borda(a_1) = (m - 1) \cdot |G_1| + 0 \cdot |G_2|$$

$$Borda(a_2) = (m-2) \cdot |G_1| + 1 \cdot |G_2|$$

$$\ldots$$

$$Borda(a_m) = 0 \cdot |G_1| + (m-1) \cdot |G_2|$$

As $|G_1| > |G_2|$, the Borda scores are ranked as follows:

$$Borda(a_1) > Borda(a_2) > \ldots > Borda(a_m)$$

The order of the alternatives in the output ranking $\succ_{out}$ corresponds to $\succ_{G_1}$ such that $\succ_{G_1} = \succ_{out}$.

The case $|G_2| > |G_1|$ is symmetric and leads to $\succ_{G_2}$ being the output ranking.

If $|G_1| = |G_2|$, then any ranking has the same Borda score, and a tie-breaking rule determines the output. □

Theorem 4. *Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ such that all voters in the group G_1 submit the ranking $\succ^t_{G_1}$ and all voters in G_2 submit the completely reversed ranking $\succ^t_{G_2}$ of $\succ^t_{G_1}$ in every round t. Then, Borda and Kemeny with the unit cost weight update function satisfy simple proportionality.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 and Lemma 3.2 that $\succ^t_{out} = \succ^t_{G_1}$, since the majority group determines the outcome. Assume that we have a tie-breaking mechanism such that $\succ^t_{out} = \succ^t_{G_1}$ if $x = y$.

Voters in G_1 are satisfied in the first round, since $1 \cdot x \geq 1 \cdot y$. The voters in G_2 are not satisfied so the weights of each voter in G_2 increases by 1 such that the new total weight of the voters in G_2 becomes $y + y = 2 \cdot y$. There are $x + y$ rounds in total and the weights of the unsatisfied voters continue to increase. In the worst case, in the last round, we have $(x + y) \cdot y > x$ so G_2 members will certainly be eventually satisfied in one of the rounds.

As already mentioned, G_1 remains satisfied until there is a z such that $z \cdot y > u \cdot x$ in round $z + u - 1$. After that, G_2 is satisfied until $u \cdot x \geq z \cdot y$ in round $z + u - 1$.

In round $x + y - 1$ (one round before the last one), we have $z + u = x + y$ and $z = x$ and $u = y$. Therefore, $z \cdot y = u \cdot x$. As the total weight of voters in G_1 equals the total weight of voters in G_2, voters in G_1 win by tie-breaking. Now, the number of satisfied rounds of voters in G_1 is x and number of satisfied rounds of G_2 is $y - 1$. Then, we have $(z + 1) \cdot y > u \cdot x$ in the last round so $\succ^{x+y}_{out} = \succ^{x+y}_{G_2}$, i.e., each voter in G_2 is satisfied in the last round. Therefore, the number of satisfied rounds of G_2 is y.

In total, voters in G_1 are satisfied x times and voters in G_2 are satisfied y times. Hence, each group is satisfied in proportion to its size, which proves simple proportionality.

□

Proposition 4.1. *Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ such that all voters in the group G_1 submit the ranking $\succ^t_{G_1}$ and all voters in G_2 submit the ranking $\succ^t_{G_2}$ in every round t. Then, Kemeny with the unit cost weight update function satisfies simple proportionality.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 that $\succ^t_{out} = \succ^t_{G_1}$. Assume that we have a tie-breaking mechanism such that $\succ^t_{out} = \succ^t_{G_1}$ if $x = y$. The rest is analogous to the proof of Theorem 4. Since G_2 accumulates weight across rounds, there will be rounds in which G_2 becomes the majority in weighted terms, and thus gets satisfied. Over $x + y$ rounds, each group will be satisfied exactly in proportion to its size. Therefore, simple proportionality holds.

□

Theorem 5. *Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in the group G_1 submit the ranking $\succ^t_{G_1}$ and all voters in the group G_2 submit the completely reversed ranking $\succ^t_{G_2}$ of $\succ^t_{G_1}$ in every round t. Then, Borda and Kemeny with the perpetual KT weight update function satisfy simple proportionality.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 and Lemma 3.2 that $\succ^t_{out} = \succ^t_{G_1}$, since the majority group determines the outcome. Assume that we have a tie-breaking mechanism such that $\succ^t_{out} = \succ^t_{G_1}$ if $x = y$.

Voters in G_1 are satisfied in the first round, since $1 \cdot x \geq 1 \cdot y$, and the weights of each $v \in G_1$ decreases by $\frac{1}{2}$ such that the new total weight of G_1 becomes $x \cdot \frac{1}{2} = \frac{x}{2}$. There are $x + y$ rounds in total and the weights of the satisfied voters continue to decrease. In the worst case, in the last round, we have $y > \frac{x}{x+y}$ so G_2 members will certainly be eventually satisfied in one of the rounds.

As already mentioned G_1 remains satisfied until there is a u such that $\frac{y}{z} > \frac{x}{u}$ in round $z + u - 1$. After that, G_2 is satisfied until $\frac{y}{z} \leq \frac{x}{u}$ in round $z + u - 1$.

In round $x + y - 1$ (one round before the last one), we have $z + u = x + y$ and $z = x$ and $u = y$. Therefore, $\frac{y}{z} = \frac{x}{u}$. As the total weight of voters in G_1 equals the total weight of voters in G_2, voters in G_1 win by tie-breaking. Now, the number of satisfied rounds of G_1 is x and number of satisfied rounds of G_2 is $y - 1$. Then, we have $\frac{y}{z} > \frac{x}{u+1}$ in the last round so $\succ^{x+y}_{out} = \succ^{x+y}_{G_2}$, i.e., each voter $v \in G_2$ are satisfied in the last round. Therefore, the number of satisfied rounds of G_2 is y.

In total, voters in G_1 are satisfied x times and voters in G_2 are satisfied y times. Hence, each group is satisfied in proportion to its size, which proves simple proportionality.

□

Proposition 5.1. *Assume there are two groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in the group G_1 submits $\succ^t_{G_1}$ and all voters in the group G_2 submits the ranking $\succ^t_{G_2}$ for any round t. Then, Kemeny with the perpetual KT weight update function satisfy simple proportionality if satisfaction threshold $\alpha < d_K(\succ_{G_1}, \succ_{G_2})$.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 that $\succ^t_{out} = \succ^t_{G_1}$. Assume that we have a tie-breaking mechanism such that $\succ^t_{out} = \succ^t_{G_1}$ if $x = y$. The rest is analogous to the proof of Theorem 5. Since G_1 loses weight across rounds, there will be rounds in which G_2 becomes the majority in weighted terms, and thus gets satisfied. Over $x + y$ rounds, each group will be satisfied exactly in proportion to its size. Therefore, simple proportionality holds.

□

Theorem 6. *Assume there are two groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in the group G_1 submits $\succ^t_{G_1}$ and all voters in the group G_2 submits the completely reversed ranking $\succ^t_{G_2}$ of $\succ^t_{G_1}$ for any round t. Then, Borda and Kemeny with a KT-distance-based update function satisfy simple proportionality.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 and Lemma 3.2 that $\succ^t_{out} = \succ^t_{G_1}$, since the majority group determines the outcome. Assume that we have a tie-breaking mechanism such that $\succ^t_{out} = \succ^t_{G_1}$ if $x = y$.

The outcome of the first round is $\succ^t_{out} = \succ_{G_1}$, since $1 \cdot x \geq 1 \cdot y$. The voters in G_2 are not satisfied in the first round, since $d_K(\succ_{out}, \succ_{G_2}) = d_{max}$. The weights of each voter in G_2 increases by d_{max} such that the new total weight of the voters in G_2 becomes $y + d_{max}$. There are $x + y$ rounds in total and the weights of the voters in G_2 continue to increase by d_{max}. In the worst case, in the last round, we have $d_{max} \cdot x + y > x$ so G_2 members will certainly be eventually satisfied in one of the rounds.

As already mentioned, $\succ^1_{out} = \succ_{G_1}$ until there is a z such that $d_{max} \cdot z + y > u \cdot x$ in round $z + u - 1$. After that, G_2 is satisfied until $d_{max} \cdot u + x \geq z \cdot y$ in round $z + u - 1$.

In round $x + y - 1$ (one round before the last one), we have $z + u = x + y$ and $z = x$ and $u = y$. Therefore, $d_{max} \cdot z + y = d_{max} \cdot u + x$. As the total weight of voters in G_1 equals the total weight of voters in G_2, voters in G_1 win by tie-breaking. Now, the number of satisfied rounds of voters in G_1 is x and number of satisfied rounds of G_2 is $y - 1$. Then, we have $d_{max} \cdot (z + 1) + y > d_{max} \cdot u + x$ in the last round so $\succ^{x+y}_{out} = \succ^{x+y}_{G_2}$, i.e., each voter in G_2 is satisfied in the last round. Therefore, the number of satisfied rounds of G_2 is y.

In total, voters in G_1 are satisfied x times and voters in G_2 are satisfied y times. Hence, each group is satisfied in proportion to its size, which proves simple proportionality.

□

Proposition 6.1. *Assume there are two groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and all voters in the group G_1 submits $\succ^t_{G_1}$ and all voters in the group G_2 submits the ranking $\succ^t_{G_2}$ for any round t. Then, Kemeny with a KT-distance-based update function satisfy simple proportionality.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 that $\succ^t_{out} = \succ^t_{G_1}$. Assume that we have a tie-breaking mechanism such that $\succ^t_{out} = \succ^t_{G_1}$ if $x = y$.

The rest is analogous to the proof of Theorem 6. Since G_2 increases its total weight across rounds, there will be rounds in which G_2 becomes the majority in weighted terms, and thus gets satisfied. Over $x + y$ rounds, each group will be satisfied exactly in proportion to its size. Therefore, simple proportionality holds.

$\square$

Theorem 7. *Even if there are two groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and $|N| > 2$, and all voters in the group G_1 submits $\succ_{G_1}^t$ and all voters in the group G_2 submits the completely reversed ranking $\succ_{G_2}^t$ of $\succ_{G_1}^t$ for any round t, Borda and Kemeny with reset-based update functions or myopic weight update functions do not satisfy simple proportionality.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 and Lemma 3.2 that $\succ_{out}^t = \succ_{G_1}^t$, since the majority group determines the outcome. Assume that we have a tie-breaking mechanism such that $\succ_{out}^t = \succ_{G_1}^t$ if $x = y$.

As $x > y$, the output ranking of the first round is $\succ_{out}^1 = \succ_{G_1}$. Since $d_K(\succ_{out}^1, \succ_{G_1}) = 0$, the weights of the voters in G_1 is reset to 0 (for the myopic weight update functions) or 1 (for the reset-based weight update functions). However, $d_K(\succ_{out}^1, \succ_{G_2}) = d_{max}$. Therefore, the weights of the voters in G_2 increase by at least 1 such that G_2 has a higher total weight in round 2:

$$\sum_{v \in G_1} \omega(v, 2) < \sum_{u \in G_2} \omega(u, 2)$$

Therefore, $\succ_{out}^2 = \succ_{G_2}$. However, weights are updated again: Since $d_K(\succ_{out}^2, \succ_{G_2}) = 0$, the weights of the voters in G_2 is reset to 0 or 1. However, $d_K(\succ_{out}^2, \succ_{G_1}) = d_{max}$. Therefore, the weights of the voters in G_1 increase by at least 1 such that G_1 has a higher total weight in round 3.

This process continues until the last round n and thus the output ranking alternates between $\succ_{G_1}$ and $\succ_{G_2}$ in each round. As a result, in the final round $x + y$, every voter is satisfied in roughly half of the rounds. However, as $x > y$, voters in G_1 should be satisfied more in order to satisfy simple proportionallity. This does not happen so simple proportionality does not hold. $\square$

3.2.3 Independence of Uncontroversial Decisions

A TSPF's outcome in every round should ideally not change due to the uncontroversial rounds where all voters unanimously agree on the same ranking. Otherwise, such rounds could be strategically used to manipulate the output rankings in future rounds — for example by embedding additional uncontroversial rounds.

This property requires a formal definition of uncontroversial preferences. This suggests determining a threshold-value, since a round must be classified as controversial or uncontroversial. To this end, we can use the threshold value α that introduced at the beginning of this section to define satisfaction (Equation 3.15) as a situation in which all voters are satisfied with each other's preferences may indicate that the rankings are uncontroversial. However, it is also possible to define a more gradual transition from fully uncontroversial to fully controversial. For this reason, we propose different definitions for TBD and CD:

TBD: A preference profile P is *uncontroversial* if $d_K(\succ_v, \succ_u) \leq \alpha$ for any $v, u \in N$. A TSPF f satisfies *independence of uncontroversial decisions* if the output ranking $\succ_{out}^{t'}$ remains the same for each round $t' \in T$ except the *uncontroversial round* t, where P is uncontroversial, if we ignore the round t such that $T' = T \setminus \{t\}$ is the new set of rounds.

CD: A preference profile P is *ϕ-controversial* if $\phi = \frac{\sum_{v=1}^{n} \sum_{u=v}^{n} d_K(\succ_v, \succ_u)}{|N|}$ for any $v, u \in N$. A TSPF f satisfies *independence of uncontroversial decisions* if the output ranking $\succ_{out}^{t'}$ remains the same for each round $t' \in T$ except the minimal ϕ-controversial (least controversial) round t if we ignore the round t such that $T' = T \setminus \{t\}$ is the new set of rounds.

Theorem 8. *Independence of uncontroversial decisions does not hold for any TSPF satisfying continuity if every voter is satisfied in the uncontroversial round and for each satisfied voter v in round t the weight is multiplied by a constant $c > 0$ s.t. $\omega(t+1, v) = c \cdot \omega(t, v)$, where this constant might change from round to round.*

Proof. Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$, $G_1 \cup G_2 = N$, $|G_1| = g_1$ and $|G_2| = g_2$.

In an arbitrary round t, all voters in the group G_1 submit $\succ_{G_1}^t$ and all voters in the group G_2 submit $\succ_{G_2}^t$. Let ω be the following weight update function with a constant $d \geq 0$:

$$\omega(t+1, v) = \begin{cases} \omega(t, v) + d, & \text{if } d_K(\succ_v^t, \succ_{out}^t) > \alpha, \\ c_t \cdot \omega(t, v), & \text{if } d_K(\succ_v^t, \succ_{out}^t) \leq \alpha. \end{cases}$$

After the uncontroversial round t where all voters $x \in N$ are satisfied, the weight of any voter $x \in N$ is updated as follows:

$\omega(t+1, x) = c_t \cdot \omega(t, x)$

Assume that in round $t+1$, the groups have divergent opinions such that the output ranking satisfies only one of the groups. WLOG we assume that G_1 is satisfied. Therefore, after round $t+1$, the weight of any voter $v \in G_1$ is updated as follows:

$\omega(t+2, v) = c_{t+1} \cdot \omega(t+1, v) = c_{t+1} \cdot c_t \cdot \omega(t, v)$

As voters $u \in G_2$ are not satisfied, their weight is updated as follows:

$\omega(t+2, u) = \omega(t+1, u) + d = c_t \cdot \omega(t, u) + d$

Since $|G_1| = g_1$ and $|G_2| = g_2$, the sum of weights of voters in group G_1 is

$\sum_{v \in G_1} \omega(t+2, v) = c_{t+1} \cdot c_t \cdot \sum_{v \in G_1} \omega(t, v)$

and the sum of weights of voters in group G_2 is

$$\sum_{u\in G_2} \omega(t+2,u) = c_t \cdot \sum_{u\in G_2} (\omega(t,u)+d) = c_t \cdot \sum_{u\in G_2} \omega(t,u) + g_2 \cdot d.$$

Now, we remove the uncontroversial round t from T such that such that $T' = T \setminus \{t\} = \{1, 2 \ldots, t-1, t+1, \ldots, o\}$ is the new set of time steps. Observe that the sums of weights of voters differ if t is removed:

If $v \in G_1$:

$$\sum_{v\in G_1} \omega(t+2,v) = g_1 \cdot c_{t+1} \cdot \sum_{v\in G_1} \omega(t,v)$$

and if $u \in G_2$:

$$\sum_{u\in G_2} \omega(t+2,u) = \sum_{u\in G_2} (\omega(t,u)+d) = \sum_{u\in G_2} \omega(t,u) + g_2 \cdot d.$$

It is obvious that the proportion of total weight of G_1 to the total weight of G_2 is not preserved:

$$\frac{c_{t+1} \cdot c_t \cdot \sum_{v\in G_1} \omega(t,v)}{c_t \cdot \sum_{u\in G_2} \omega(t,u) + g_2 \cdot d} \neq \frac{c_{t+1} \cdot \sum_{v\in G_1} \omega(t,v)}{\sum_{u\in G_2} \omega(t,u) + g_2 \cdot d}$$

For any TSPF that satisfies continuity, there always exist values c_t and d such that the output ranking in round $t+1$ changes depending on whether round t is taken into account or ignored. This happens because the constants c_t and d influence the weights, thereby altering the proportions among voter groups. Specifically, one can choose a sufficiently large c_t and an appropriate d such that either group G_1 or group G_2 dominates the outcome in round $t+1$. Therefore, independence of uncontroversial decisions is not satisfied.

□

Note that this result seems in contrast with a similar result *(Proposition 3)* for the approval setting in an unpublished manuscript by Lackner and Maly [43]. Indeed, we believe that the following counterexample refutes their *Proposition 3*. It should be noted however that this claim has never appeared in a formal publication; we report it merely for the sake of completeness.

Proposition 3 from [43] is stated as follows:

> **Proposition 3** (from Lackner and Maly [43]): *Let $\mathcal{R}$ be a basic WAM. Then, $\mathcal{R}$ satisfies independence of uncontroversial decisions if and only if for every k there is a constant c_k such that for every x that can occur as a weight after k rounds we have $g(x) = c_k x$ for $c_k > 0$.*

In other words, an approval voting based perpetual voting method with an update function — can be compared to a TSPF in our settings — satisfies independence of uncontroversial decisions if and only if the weight of a winning voter can only be changed by multiplying it with a constant. We do not introduce the notation used in this paper. Instead, we try to stick with our notation and use natural language if it is not possible to express it with our notation. The main goal is to discuss the correctness of this result by giving a concrete counterexample.

Although it is claimed that independence of uncontroversial decisions still holds if the weight of a winning voter can only be changed by multiplying it with a constant, we can construct a counterexample that violates independence of uncontroversial decisions even in approval voting settings. In approval voting, the voters do not submit a ranking but an approval set in which the alternatives they support are contained. Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$, $G_1 \cup G_2 = N$ and $|G_1| = 4$, $|G_2| = 12$. Let $A = \{a_1, a_2\}$ be the set of alternatives and ω be the following weight update function:

$$\omega(t+1, v) = \begin{cases} \omega(t, v) + \frac{1}{2}, & \text{if the approval set of } v \text{ is not represented in the outcome,} \\ \frac{1}{2} \cdot \omega(t, v), & \text{if the approval set of } v \text{ is represented in the outcome.} \end{cases}$$

These two possible scenarios, one with and one without the uncontroversial round 1, show that independence of uncontroversial decisions is violated:

1. In round 1, all $x \in N$ approve $\{a_1, a_2\}$ so the first round is uncontroversial. After round 1, the weight of any voter $x \in N$ is updated as follows:

   $\omega(2, x) = \frac{1}{2} \cdot \omega(1, x) = \frac{1}{2} \cdot 1 = \frac{1}{2}$.

   In round 2, all $v \in G_1$ approve $\{a_1\}$ whereas $u \in G_2$ approve $\{a_2\}$. As $|G_1| = 4$, $|G_2| = 12$ and $\omega(2, x) = \frac{1}{2}$ for any voter x, the total weight of voters of G_1 is

   $\sum\limits_{v \in G_1} \omega(2, v) = \frac{1}{2} \cdot 4 = 2$

   and of G_2 is

   $\sum\limits_{u \in G_2} \omega(2, u) = \frac{1}{2} \cdot 12 = 6$.

   Since $\sum\limits_{u \in G_2} \omega(2, u) > \sum\limits_{v \in G_1} \omega(2, v)$, a_2 wins round 2. After round 2, the weight of any voter $u \in G_2$ is updated as follows:

   $\omega(3, u) = \frac{1}{2} \cdot \omega(2, u) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

   And the weight of any voter $v \in G_1$ is updated as follows:

   $\omega(3, u) = \omega(2, u) + \frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1$.

   In round 3, all $v \in G_1$ approve $\{a_1\}$ whereas $u \in G_2$ approve $\{a_2\}$. As $|G_1| = 4$, $|G_2| = 12$, the total weight of voters of G_1 is

$\sum_{v \in G_1} \omega(3, v) = 1 \cdot 4 = 4$

and of G_2 is

$\sum_{u \in G_2} \omega(3, u) = \frac{1}{4} \cdot 12 = 3.$

Since $\sum_{u \in G_2} \omega(2, u) < \sum_{v \in G_1} \omega(2, v)$, a_1 wins round 3.

2. In round 1, all $v \in G_1$ approve $\{a_1\}$ whereas $u \in G_2$ approve $\{a_2\}$. As $|G_2| > |G_1|$ and $\omega(1, x) = 1$ for any x, a_2 is the winner. After round 1, the weight of any voter $u \in G_2$ is updated as follows:

   $\omega(2, u) = \frac{1}{2} \cdot \omega(1, u) = \frac{1}{2} \cdot 1 = \frac{1}{2}.$

   And the weight of any voter $v \in G_1$ is updated as follows:

   $\omega(2, u) = \omega(1, u) + \frac{1}{2} = 1 + \frac{1}{2} = \frac{3}{2}.$

   In round 2, all $v \in G_1$ approve $\{a_1\}$ whereas $u \in G_2$ approve $\{a_2\}$. As $|G_1| = 4$, $|G_2| = 12$, the total weight of voters of G_1 is

   $\sum_{v \in G_1} \omega(2, v) = \frac{3}{2} \cdot 4 = 6$

   and of G_2 is

   $\sum_{u \in G_2} \omega(2, u) = \frac{1}{2} \cdot 12 = 6.$

   Since $\sum_{u \in G_2} \omega(2, u) = \sum_{v \in G_1} \omega(2, v)$, by using an appropriate tie-breaking rule, a_2 wins round 2, which differs from the result of round 3 in Case 1, violating the independence of uncontroversial decisions.

Theorem 9. *Assume that a TSPF satisfies the property that if voter preferences and weights remain unchanged across rounds — that is, for all $t \in T$, $\succ_v^{t+1} = \succ_v^t$ and $\omega(t + 1, v) = \omega(t, v)$ for every voter $v \in N$ — then the output rankings also remain unchanged across rounds, i.e., $\succ_{out}^t = \succ_{out}^{t+1}$ for all t. Independence of uncontroversial decisions holds for any such TSPF, where every voter is satisfied in every uncontroversial round and for each satisfied voter v in round t the weight remains unchanged such that $\omega(t + 1, v) = \omega(t, v)$.*

Proof. Assume that round t is an uncontroversial round for a preference profile P, i.e., $d_K(\succ_v^t, \succ_u^t) \leq \alpha$ for any $v, u \in N$, where N is the set of voters. As every voter is satisfied in every uncontroversial round, each voter $v \in N$ is satisfied in uncontroversial round t such that $d_K(\succ_v^t, \succ_{out}^t) \leq \alpha$. After this round, the weight remains the same for each voter v:

$$\omega(t + 1, v) = \omega(t, v)$$

Since we assume that every voter is satisfied with the output ranking in an uncontroversial round, the weight remains the same for each voter in round $t + 1$ after an uncontroversial round t. If we ignore the round t, the preferences and weights for other rounds $t' \in T' \setminus \{t\}$

do not change, as round t does not change the weights. Since the given TSPF satisfies the property that if voter preferences and weights remain unchanged across rounds, the output rankings also remain unchanged across rounds, the output rankings do not change either. Therefore, independence of uncontroversial decisions is satisfied.

□

Corollary 9.1. *Assume that a TSPF satisfies the property that if voter preferences and weights remain unchanged across rounds — that is, for all $t \in T$, $\succ_v^{t+1} = \succ_v^t$ and $\omega(t+1, v) = \omega(t, v)$ for every voter $v \in N$ — then the output rankings also remain unchanged across rounds, i.e., $\succ_{out}^t = \succ_{out}^{t+1}$ for all t. Independence of uncontroversial decisions holds for any such TSPF with the unit cost weight update function, where every voter is satisfied in every uncontroversial round.*

Proof. If a voter $v \in N$ with a preference ranking $\succ_v^t$ in round $t \in T$ is satisfied with the output ranking $\succ_{out}^t$ such that $d_K(\succ_v^t, \succ_{out}) \leq \alpha$, the unit cost weight update function leaves the weight of v the same:

$$\omega(t+1, v) = \omega(t, v)$$

Since we assume that every voter is satisfied with the output ranking, Theorem 9 implies that any TSPF satisfying the given property above with the unit cost weight update function is independent of uncontroversial decisions. □

Proposition 9.1. *Assume that a TSPF satisfies the property that if voter preferences and weights remain unchanged across rounds — that is, for all $t \in T$, $\succ_v^{t+1} = \succ_v^t$ and $\omega(t+1, v) = \omega(t, v)$ for every voter $v \in N$ — then the output rankings also remain unchanged across rounds, i.e., $\succ_{out}^t = \succ_{out}^{t+1}$ for all t. Independence of uncontroversial decisions holds for any such TSPF with the constant weight update function.*

Proof. Assume that round t is an uncontroversial round for a preference profile P, i.e., $d_K(\succ_v^t, \succ_u^t) \leq \alpha$ for any $v, u \in N$, where N is the set of voters. After this round, the weight remains the same for each voter v, as the constant weight update function keeps the weights fixed over time:

$$\omega(t+1, v) = \omega(t, v)$$

If we ignore the round t, the preferences and weights for other rounds $t' \in T'\{t\}$ do not change, as round t does not change the weights. Since the given TSPF satisfies the property that if voter preferences and weights remain unchanged across rounds, the output rankings also remain unchanged across rounds, the output rankings do not change either. Therefore, independence of uncontroversial decisions is satisfied. □

Theorem 10. *Independence of uncontroversial decisions does not hold for any TSPF satisfying continuity for each satisfied voter v in round t a constant $c > 0$ is added to the weight s.t. $\omega(t+1, v) = \omega(t, v) + c$, where this constant might change from round to round.*

Proof. Assume there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$, $G_1 \cup G_2 = N$, $|G_1| = g_1$ and $|G_2| = g_2$.

In an arbitrary round t, all voters in the group G_1 submit $\succ^t_{G_1}$ and all voters in the group G_2 submit $\succ^t_{G_2}$. Let ω be the following weight update function with a constant $d \geq 0$:

$$\omega(t+1, v) = \begin{cases} \omega(t, v) + d, & \text{if } d_K(\succ^t_v, \succ^t_{out}) > \alpha, \\ c_t + \omega(t, v), & \text{if } d_K(\succ^t_v, \succ^t_{out}) \leq \alpha. \end{cases}$$

After the uncontroversial round t where all voters $x \in N$ are satisfied, the weight of any voter $x \in N$ is updated as follows:

$\omega(t+1, x) = c_t + \omega(t, x)$

Assume that in round $t+1$, the groups have divergent opinions such that the output ranking satisfies only one of the groups. WLOG we assume that G_1 is satisfied. Therefore, after round $t+1$, the weight of any voter $v \in G_1$ is updated as follows:

$\omega(t+2, v) = c_{t+1} + \omega(t+1, v) = c_{t+1} + c_t + \omega(t, v)$

As voters $u \in G_2$ are not satisfied, their weight is updated as follows:

$\omega(t+2, u) = \omega(t+1, u) + d = c_t + d + \omega(t, u)$

Since $|G_1| = g_1$ and $|G_2| = g_2$, the sum of weights of voters in group G_1 is

$\sum\limits_{v \in G_1} \omega(t+2, v) = g_1 \cdot (c_{t+1} + c_t) + \sum\limits_{v \in G_1} \omega(t, v)$

and the sum of weights of voters in group G_2 is

$\sum\limits_{u \in G_2} \omega(t+2, u) = g_2 \cdot (c_t + d) + \sum\limits_{u \in G_2} \omega(t, u).$

Now, we remove the uncontroversial round t from T such that such that $T' = T \setminus \{t\} = \{1, 2 \ldots, t-1, t+1, \ldots, o\}$ is the new set of time steps. Observe that the sums of weights of voters differ if t is removed:

If $v \in G_1$:

$\sum\limits_{v \in G_1} \omega(t+2, v) = g_1 \cdot c_{t+1} + \sum\limits_{v \in G_1} \omega(t, v)$

and if $u \in G_2$:

$\sum\limits_{u \in G_2} \omega(t+2, u) = g_2 \cdot d + \sum\limits_{u \in G_2} \omega(t, u).$

It is obvious that the proportion of total weight of G_1 to the total weight of G_2 is not preserved:

$$\frac{g_1 \cdot (c_{t+1} + c_t) + \sum\limits_{v \in G_1} \omega(t, v)}{g_2 \cdot (c_t + d) + \sum\limits_{u \in G_2} \omega(t, u)} \neq \frac{g_2 \cdot c_{t+1} + \sum\limits_{v \in G_1} \omega(t, v)}{g_2 \cdot d + \sum\limits_{u \in G_2} \omega(t, u)}$$

For any TSPF that satisfies continuity, there always exist values c_t and d such that the output ranking in round $t+1$ changes depending on whether round t is taken into account or ignored. This happens because the constants c_t and d influence the weights, thereby altering the proportions among voter groups. Specifically, one can choose sufficiently large c_t and d such that either group G_1 or group G_2 dominates the outcome in round $t+1$. Therefore, independence of uncontroversial decisions is not satisfied. □

Theorem 11. *Independence of uncontroversial decisions does not hold for any TSPF satisfying continuity with the KT-based or myopic weight update functions.*

Proof. Let t be an uncontroversial round such that

$$\forall\, u, v \in N: \quad d_K(\succ_u^t, \succ_v^t) \leq \alpha,$$

but not all every preference lists are identical — so some voters disagree slightly (but still within α).

Let $v, u \in N$ be two voters whose preferences in round t differ (so that $d_K(\succ_v^t, \succ_u^t) > 0$). WLOG, assume that we use KT and myopic KT weight update functions. After round t, the weights are updated as follows for each $w \in N$ for KT weight update function:

$$\omega(t+1, w) = \omega(t, w) + d_K(\succ_w^t, \succ_{out}^t)$$

and updated as follows for any $v \in N$ for myopic KT weight update function:

$$\omega(t+1, w) = d_K(\succ_w^t, \succ_{out}^t)$$

The crucial point is that because $\succ_v^t \neq \succ_u^t$, the increments $d_K(\succ_v^t, \succ_{out}^t)$ and $d_K(\succ_u^t, \succ_{out}^t)$ differ by at least one. If we ignore round t, this weight updates do not happen and the weight of each voter stays the same for each voter $w \in N$ such that

$$\omega(t+1, w) = \omega(t, w)$$

Hence,

$$(\omega(t, v) + d_K(\succ_v^t, \succ_{out}^t)) - (\omega(t, u) + d_K(\succ_u^t, \succ_{out}^t)) \;\neq \omega(t, v) - \omega(t, u)$$

In particular, the weight gap between v and u at the start of round $t+1$ is strictly different if we ignore round t than the case in which we do not ignore round t.

Because the underlying SPF is continuous, a sufficiently large weight advantage for one voter can force their individual ranking to become the group outcome. If we ignore t, this does not happen and another output ranking is produced in round $t+1$.

This shows that inserting or removing an uncontroversial round t can change the outcome in round $t+1$. Therefore, independence of uncontroversial decisions fails for every KT-distance-based or myopic update rule.

□

Proposition 11.1. *Independence of uncontroversial decisions does not hold for any TSPF satisfying continuity with the perpetual KT weight update function.*

Proof. Let t be an uncontroversial round such that

$$\forall\, u, v \in N : \quad d_K(\succ_u^t, \succ_v^t) \leq \alpha,$$

Let $v, u \in N$ be two voters whose weights in round t differ (so that $\omega(t, v) \neq \omega(t, u)$). WLOG, assume that we use the unit cost reset weight update function. After round t, the weights are updated as follows for each satisfied $w \in N$ with $d_K(\succ_{out}^t, \succ_w^t) \leq \alpha$:

$$\omega(t+1, w) = \frac{\omega(t+1, w)}{\omega(t+1, w) + 1}$$

and updated as follows for each unsatisfied $w \in N$ with $d_K(\succ_{out}^t, \succ_w^t) \leq \alpha$:

$$\omega(t+1, w) = \omega(t, w)$$

Now, in round $t+1$, there are voter with a decreased weight but also voters whose weights remain unchanged, since not all voters have to be satisfied with the output ranking $\succ_{out}^t$. Because the underlying SPF is continuous, a sufficiently large weight advantage for one voter forces their individual ranking to become the group outcome. If we remove round t, we cancel this weight update such that this does not happen. This shows that inserting or removing an uncontroversial round t can change the outcome in round $t+1$. Therefore, independence of uncontroversial decisions fails for the perpetual KT weight update function. □

Proposition 11.2. *Independence of uncontroversial decisions does not hold for any TSPF satisfying continuity with a reset-based weight update function.*

Proof. Let t be an uncontroversial round such that

$$\forall\, u, v \in N : \quad d_K(\succ_u^t, \succ_v^t) \leq \alpha,$$

Let $v, u \in N$ be two voters whose weights in round t differ (so that $\omega(t, v) \neq \omega(t, u)$). WLOG, assume that we use the unit cost reset weight update function. After round t, the weights are updated as follows for each satisfied $w \in N$ with $d_K(\succ_{out}^t, \succ_w^t) \leq \alpha$:

$$\omega(t+1, w) = 1$$

and updated as follows for each unsatisfied $w \in N$ with $d_K(\succ_{out}^t, \succ_w^t) \leq \alpha$:

$$\omega(t+1, w) = \omega(t, w) + 1$$

Now, in round $t+1$, there are voter with an increased weight but also voters whose weights remain unchanged, since not all voters have to be satisfied with the output ranking $\succ_{out}^t$.

Because the underlying SPF is continuous, a sufficiently large weight advantage for one voter forces their individual ranking to become the group outcome. If we remove round t, we cancel this weight update such that this does not happen. This shows that inserting or removing an uncontroversial round t can change the outcome in round $t+1$. Therefore, independence of uncontroversial decisions fails for every reset-based weight update function. □

3.2.4 Bounded Dry Spells

Now, we continue with the final axiomatic property examined in this thesis: *bounded dry spells*. This property formalizes the intuitive idea that every voter should experience satisfaction within a limited number of rounds. In other words, no voter should be excluded from receiving satisfactory outcomes for too long. It captures temporal fairness by placing an upper bound on how long a voter can go without being satisfied.

For this property, we provide two definitions, tailored to the two satisfaction models introduced earlier (TBD and CD):

TBD: A voter $v \in N$ has a dry spell of length ℓ if there exists $t \leq |T| - \ell$ such that $sat_t(v) = sat_{t+\ell}(v)$, i.e., voter v's satisfaction does not increase in the last ℓ rounds.

CD: A voter $v \in N$ has a dry spell of length ℓ if there exists $t \leq |T| - \ell$ such that $sat_t(v) + \min_{k \in T} sat_k(v) \leq sat_{t+\ell}(v)$, i.e., voter v's satisfaction does not increase more than its round with the least satisfaction level in the last ℓ rounds.

Let $d : \mathbb{N} \to \mathbb{N}$ be a function. A TSPF has *bounded dry spells* of d if there is no voter with a dry spell of length $d(|N|)$ for any preference profile.

Theorem 12. *Assume that a TSPF satisfies the property that if voter preferences remain unchanged across rounds — that is, for all $t \in T$, $\succ_v^{t+1} = \succ_v^t$ for every voter $v \in N$ — then the output rankings also remain unchanged across rounds, i.e., $\succ_{out}^t = \succ_{out}^{t+1}$ for all t. No such TSPF with the constant weight update function has bounded dry spells.*

Proof. Since the weight update function is constant, $\omega(t, v) = \omega(t + \ell, v)$ for all $t \in T$, $\ell \geq 0$, and $v \in N$. Moreover, due to the TSPF's property, the output ranking remains fixed across all rounds if preferences do not change, i.e., $\succ_{out}^t = \succ_{out}^{t+\ell}$ for all ℓ. Now, suppose there exists a round $t \in T$ with an unsatisfied voter $v \in N$. Additionally, assume that preferences do not change in the following rounds. Since the output ranking does not change in any of the rounds, voter v stays unsatisfied in any upcoming round $t + \ell$, i.e., the dry spell is unbounded. □

Corollary 12.1. *Weighted Borda, weighted Kemeny and weighted squared Kemeny with the constant weight update function has no bounded dry spell guarantee.*

Proof. For these rules, the property holds that if voter preferences remain unchanged across rounds, then the output rankings also remain unchanged across rounds. Accordingly, a TSPF based on one of these SPFs also satisfies this property. It follows from 12 that such a TSPF with the constant weight update function cannot have bounded dry spells. Therefore, if there is a round $t \in T$ with an unsatisfied voter $v \in N$, this voter remains unsatisfied throughout the following rounds, i.e., the dry spell is unbounded. ☐

Theorem 13. *No TSPF satisfying continuity with a myopic weight update function has a bounded dry spell guarantee.*

Proof. Let $v \in N$ be a voter with preference ranking $\succ_v^t$ in round $t \in T$, $C_1, C_2 \subseteq N$ coalitions such that $C_1 \cap C_2 = \emptyset$ and $C_1 \cup C_2 \cup \{v\} = N$. In addition, let A be the set of alternatives. We show that v has unbounded dry spells.

Assume that $P_{C_1} \in \mathcal{L}(A)^{|C_1|}$ is a preference profile, where every voter $x \in C_1$ submit the same ranking $\succ_{C_1}^t$ in round $t \in T$. Likewise, assume a profile $P_{C_2} \in \mathcal{L}(A)^{|C_2|}$, where every $u \in C_2$ submit the same ranking $\succ_{C_2}^t$ in round $t \in T$. Assume further that $|C_1| = |C_2| + 1 + \ell$, where ℓ is large enough such that the output ranking is $\succ_{out}^t = \succ_{C_1}$ for any round t, which holds by the continuity of the TSPF.

In round 1, $\succ_{out}^1 = \succ_{C_1}^1$ and suppose that $d_K(\succ_{C_2}^1, \succ_{out}^1) = d_{max}$ and $d_K(\succ_v^1, \succ_{out}^1) = d_{max}$, i.e., $\succ_{C_1}^1$ is the completely reversed ranking of $\succ_{C_2}^1 = \succ_v^1$. After round 1, the weights are updated according to a myopic weight update function. WLOG, assume that we use *myopic KT weight update function* such that $\omega(2, u) = d_{max}$ for any $u \in C_2 \cup \{v\}$ and $\omega(2, x) = 0$ for any $x \in C_1$. Assume that $C_2 = 1 + k$, where k is large enough such that the output ranking is $\succ_{out}^t = \succ_{C_2}^t$ if $\sum_{x \in C_1} \omega(t, x) = 0$ for any round t, which holds by the continuity of the TSPF.

In round 2, $\succ_{out}^2 = \succ_{C_2}^2$ and suppose that $d_K(\succ_{C_1}^2, \succ_{out}^2) = d_{max}$ and $d_K(\succ_v^2, \succ_{out}^2) = d_{max}$, i.e., $\succ_{C_2}^2$ is the completely reversed ranking of $\succ_{C_1}^2 = \succ_v^2$. After round 2, the weights are updated again such that $\omega(3, x) = d_{max}$ for any $x \in C_1 \cup \{v\}$ and $\omega(3, u) = 0$ for any $u \in C_2$.

This pattern continues until the final round $o = |T|$. At each round, the output ranking is dominated by the coalition whose members currently have weight d_{max}, and that coalition's ranking is the reverse of v's current preferences. Thus, in every round t, v's satisfaction is zero, i.e., $sat_t(v) = 0$. Therefore, the dry spell of v can be extended indefinitely, implying that no bounded dry spell guarantee exists under a TSPF satisfying continuity with a myopic weight update function.

☐

Theorem 14. *No TSPF satisfying continuity with a KT-distance-based weight update function has bounded dry spells.*

Proof. Let $v \in N$ be a voter with preference ranking $\succ_v^t$ in round $t \in T$, $C_1, C_2 \subseteq N$ coalitions such that $C_1 \cap C_2 = \emptyset$ and $C_1 \cup C_2 \cup \{v\} = N$. In addition, let A be the set of alternatives. We show that v has unbounded dry spells.

Assume that $P_{C_1} \in \mathcal{L}(A)^{|C_1|}$ is a preference profile, where every voter $x \in C_1$ submit the same ranking $\succ^t_{C_1}$ in round $t \in T$. Likewise, assume a profile $P_{C_2} \in \mathcal{L}(A)^{|C_2|}$, where every $u \in C_2$ submit the same ranking $\succ^t_{C_2}$ in round $t \in T$. Assume further that $|C_1| = |C_2| + 1 + \ell$, where ℓ is large enough such that the output ranking is $\succ^t_{out} = \succ_{C_1}$ for any round t, which holds by the continuity of the TSPF.

In round 1, $\succ^1_{out} = \succ^1_{C_1}$ and suppose that $d_K(\succ^1_{C_2}, \succ^1_{out}) = d_{max}$ and $d_K(\succ^1_v, \succ^1_{out}) = d_{max}$, i.e., $\succ^1_{C_1}$ is the completely reversed ranking of $\succ^1_{C_2} = \succ^1_v$. After round 1, the weights are updated according to a KT-distance-based weight update function. WLOG, assume that we use *KT weight update function* such that $\omega(2, u) = 1 + d_{max}$ for any $u \in C_2 \cup \{v\}$ and $\omega(2, x) = 1$ for any $x \in C_1$. Now, assume that $\sum_{u \in C_2} \omega(2, u) = 1 + d_{max} + \sum_{x \in C_1} \omega(2, x) + k$, where k is large enough such that the output ranking is $\succ^t_{out} = \succ_{C_2}$ for round 2, which holds by the continuity of the TSPF.

In round 2, $\succ^2_{out} = \succ^2_{C_2}$ and suppose that $d_K(\succ^2_{C_1}, \succ^2_{out}) = d_{max}$ and $d_K(\succ^2_v, \succ^2_{out}) = d_{max}$, i.e., $\succ^2_{C_2}$ is the completely reversed ranking of $\succ^2_{C_1} = \succ^2_v$. After round 2, the weights are updated again such that $\omega(3, x) = 1 + d_{max}$ for any $x \in C_1$, $\omega(3, v) = 1 + 2 \cdot d_{max}$ and $\omega(3, u) = 1 + d_{max}$ for any $u \in C_2$.

This pattern continues until the final round $o = |T|$, with the following total weights, where, WLOG, we assume number of rounds $|T|$ to be even for simplicity:

$$\begin{aligned}
\omega(o, v) &= 1 + (o - 1) \cdot d_{\max} \\
&< \sum_{u \in C_2} \omega(o, u) + \sum_{x \in C_1} \omega(o, x) \\
&< |C_2| \cdot \left(1 + d_{\max} \cdot \left(\frac{o}{2} - 1\right)\right) + |C_1| \cdot \left(1 + d_{\max} \cdot \frac{o}{2}\right) \\
&= (|C_1| + |C_2|) \cdot \left(\left(\frac{o}{2} + \left(\frac{o}{2} - 1\right)\right) \cdot d_{\max}\right) + 2 \\
&= (|C_1| + |C_2|) \cdot (o - 1) \cdot d_{\max} + 2
\end{aligned}$$

v's weight is never large enough to dominate in any round so, at each round, the output ranking is dominated by one of the coalitions, having a ranking that is the reverse of v's current preferences. Thus, in every round t, v's satisfaction is zero, i.e., $sat_t(v) = 0$. Therefore, the dry spell of v can be extended indefinitely, implying that no bounded dry spell guarantee exists under a TSPF satisfying continuity with a KT-distance-based weight update function. □

Theorem 15. *No TSPF satisfying continuity with the unit cost weight update function has bounded dry spells.*

Proof. The proof builds on the same idea as the proof of Theorem 14.

Let $v \in N$ be a voter with preference ranking $\succ^t_v$ in round $t \in T$, $C_1, C_2 \subseteq N$ coalitions such that $C_1 \cap C_2 = \emptyset$ and $C_1 \cup C_2 \cup \{v\} = N$. In addition, let A be the set of alternatives. We show that v has unbounded dry spells.

Assume that $P_{C_1} \in \mathcal{L}(A)^{|C_1|}$ is a preference profile, where every voter $x \in C_1$ submit the same ranking $\succ^t_{C_1}$ in round $t \in T$. Likewise, assume a profile $P_{C_2} \in \mathcal{L}(A)^{|C_2|}$, where every $u \in C_2$ submit the same ranking $\succ^t_{C_2}$ in round $t \in T$. Assume further that $|C_1| = |C_2| + 1 + \ell$, where ℓ is large enough such that the output ranking is $\succ^t_{out}=\succ_{C_1}$ for any round t, which holds by the continuity of the TSPF.

In round 1, $\succ^1_{out}=\succ^1_{C_1}$ and suppose that $d_K(\succ^1_{C_2}, \succ^1_{out}) = d_{max}$ and $d_K(\succ^1_v, \succ^1_{out}) = d_{max}$, i.e., $\succ^1_{C_1}$ is the completely reversed ranking of $\succ^1_{C_2}=\succ^1_v$. After round 1, the weights are updated according to the unit cost weight update function such that $\omega(2, u) = 2$ for any $u \in C_2 \cup \{v\}$ and $\omega(2, x) = 1$ for any $x \in C_1$. Now, assume that $\sum_{u \in C_2} \omega(2, u) = 2 + \sum_{x \in C_1} \omega(2, x) + k$, where k is large enough such that the output ranking is $\succ^t_{out}=\succ_{C_2}$ for round 2, which holds by the continuity of the TSPF.

In round 2, $\succ^2_{out}=\succ^2_{C_2}$ and suppose that $d_K(\succ^2_{C_1}, \succ^2_{out}) = d_{max}$ and $d_K(\succ^2_v, \succ^2_{out}) = d_{max}$, i.e., $\succ^2_{C_2}$ is the completely reversed ranking of $\succ^2_{C_1}=\succ^2_v$. After round 2, the weights are updated again such that $\omega(3, x) = 2$ for any $x \in C_1$, $\omega(3, v) = 3$ and $\omega(3, u) = 2$ for any $u \in C_2$.

This pattern continues until the final round $o = |T|$, with the following total weights, where, WLOG, we assume number of rounds $|T|$ to be even for simplicity:

$$\begin{aligned}
\omega(o, v) = o &< \sum_{u \in C_2} \omega(o, u) + \sum_{x \in C_1} \omega(o, x) \\
&< |C_2| \cdot \left(1 + \left(\frac{o}{2} - 1\right)\right) + |C_1| \cdot \left(1 + \frac{o}{2}\right) \\
&= (|C_1| + |C_2|) \cdot \left(\frac{o}{2} + \frac{o}{2}\right) = (|C_1| + |C_2|) \cdot o
\end{aligned}$$

v's weight is never large enough to dominate in any round so, at each round, the output ranking is dominated by one of the coalitions, having a ranking that is the reverse of v's current preferences. Thus, in every round t, v's satisfaction is zero, i.e., $sat_t(v) = 0$. Therefore, the dry spell of v can be extended indefinitely, implying that no bounded dry spell guarantee exists under a TSPF satisfying continuity with a KT-distance-based weight update function. □

Theorem 16. *Borda and Kemeny with a myopic weight update function or a reset-based weight update function has a bounded dry spell guarantee of length 2 if there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and $G_1 \cup G_2 = N$, where all voters in G_1 submit $\succ_{G_1}$ and all voters in G_2 submit $\succ_{G_2}$ in every round t.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 and Lemma 3.2 that $\succ^t_{out}=\succ_{G_1}$, since the majority group determines the outcome. Assume that we have a tie-breaking mechanism such that $\succ^t_{out}=\succ_{G_1}$ if $x = y$.

As $x > y$, the output ranking of the first round is $\succ_{out}^{1} = \succ_{G_1}$. Since $d_K(\succ_{out}^{1}, \succ_{G_1}) = 0$, the weights of the voters in G_1 is reset to 0 (for the myopic weight update functions) or 1 (for the reset-based weight update functions). However, $d_K(\succ_{out}^{1}, \succ_{G_2}) = d_{max}$. Therefore, the weights of the voters in G_2 increase by at least 1 such that G_2 has a higher total weight in round 2:

$$\sum_{v \in G_1} \omega(v, 2) < \sum_{u \in G_2} \omega(u, 2)$$

Therefore, $\succ_{out}^{2} = \succ_{G_2}$. However, weights are updated again: Since $d_K(\succ_{out}^{2}, \succ_{G_2}) = 0$, the weights of the voters in G_2 is reset to 0 or 1. However, $d_K(\succ_{out}^{2}, \succ_{G_1}) = d_{max}$. Therefore, the weights of the voters in G_1 increase by at least 1 such that G_1 has a higher total weight in round 3.

This process continues until the last round n and thus the output ranking alternates between $\succ_{G_1}$ and $\succ_{G_2}$ in each round. Therefore, for each voter, the satisfaction increases every second round, i.e., there is a dry spell guarantee of length 2. $\square$

Proposition 16.1. *Kemeny with a myopic weight update function or a reset-based weight update function has a bounded dry spell guarantee of length 2 if there are two disjoint groups $G_1 \subseteq N$ and $G_2 \subseteq N$ with $G_1 \cap G_2 = \emptyset$ and $G_1 \cup G_2 = N$, where all voters in G_1 submit $\succ_{G_1}$ and all voters in G_2 submit $\succ_{G_2}$ in an arbitrary round t.*

Proof. Let $|G_1| = x$, $|G_2| = y$. WLOG, assume $x > y$. We know from Lemma 3.1 that $\succ_{out}^{t} = \succ_{G_1}$. Assume that we have a tie-breaking mechanism such that $\succ_{out}^{t} = \succ_{G_1}$ if $x = y$.

The rest is analogous to the proof of Theorem 16. The output ranking alternates between $\succ_{G_1}$ and $\succ_{G_2}$ in each round. Therefore, for each voter, the satisfaction increases every second round, i.e., there is a dry spell guarantee of length 2. $\square$

3.3 Complexity-Theoretic Considerations

In real-world scenarios, it is likely that we deal with large sets of voters, agents, and time steps. For example, the music streaming service introduced in Section 1.1 could have millions of users submitting their preferences over thousands of songs. Therefore, we cannot ignore the scalability aspect of our model. In particular, our TSPFs have mainly three sources of complexity: SPF, weight update function, and the number of rounds in total.

The chosen SPF has the greatest impact on the computational complexity of a TSPF. Determining a Kemeny winner is known to be Θ_2^P-complete [38] and — since Squared Kemeny is a similar rule based on squared KT distance — it is also likely to be Θ_2^P-complete, making both rules computationally intractable for large instances. In contrast, Borda has a much lower complexity. It runs in time quadratic in the number of voters and alternatives, since each voter's preference list must be gone through, and scores are assigned accordingly. The RSD rule, on the other hand, selects a random voter's ranking as the output and thus operates in constant time in the worst case. Hence, both Borda and RSD are in P.

The weight update function is another potential bottleneck, although it is less severe. Updating weights requires either computing the set of satisfied voters (e.g., in the unit cost, unit cost reset, or perpetual KT update functions), or directly calculating values (e.g., in KT-based or myopic update functions). In both cases, computing the KT distance is necessary. A naive implementation of KT distance computation requires time quadratic in the number of alternatives. Since we must compute this for each voter, the overall complexity of a weight update step is $\mathcal{O}(|N| \cdot |A|^2)$, where $|N|$ is the number of voters and $|A|$ is the number of alternatives.

The third source of complexity, the number of rounds $|T|$, is relatively manageable. A TSPF proceeds iteratively: in each round, the SPF is applied and then the corresponding weight update function is executed. This is repeated for all $|T|$ rounds. While a large number of rounds may impact runtime, we cannot reduce it without altering the input data. Therefore, the more promising approach to scalability is to develop or use more efficient SPFs and weight update functions.

Overall, the complexity of a TSPF (given the update functions we consider) is $O(|T| \cdot |N| \cdot |A|^2 \cdot C_{SPF})$, where C_{SPF} denotes the complexity of the selected SPF. The temporal component increases complexity only linearly in the number of rounds, and weight updates — at least with the update functions that we use — are computationally manageable. While it is still important to study rules like Kemeny and Squared Kemeny for their desirable properties that we have mentioned above in the beginning of the chapter, simpler rules such as Borda rule are much more tractable and thus more suitable for real-world applications of our model.

CHAPTER 4

Simulations

In this chapter, we evaluate the practical behavior of the temporal social preference functions (TSPFs) introduced in Chapters 3. For this purpose, we implemented all proposed SPFs, weight update functions, and initialization strategies.

Our goal is to assess how different TSPFs — each consisting of a unique combination of an SPF, a weight update function, and an initialization strategy — perform on average in simulations, and to measure proportionality metrics that are difficult to capture through axiomatic analysis due to the complexity of the problem.

The code used for these experiments is publicly available at `https://github.com/kaanunalan/pifra-experiments`. Although the instances are generated at runtime, they are reproducible due to fixed random seeds. Nevertheless, the exact instances used in our simulations are also uploaded. In addition, the results are provided as a CSV file, while plots are available separately.

4.1 Setup

We implemented our system in Python3 and tested it with version Python 3.10.12. For data generation, we used the library *PrefSampling* [18].

We use Euclidean model for data generation. Voters and alternatives are represented as points placed randomly in a two-dimensional Euclidean space. The data is generated as a uniformly distributed cube (actually a square, since we work in a two-dimensional Euclidean space) with center point at the origin in the coordinate space and width of 1 in both dimensions. Rankings are based on Euclidean distances: A voter prefers the alternative the most to which their least Euclidean distance is the least. After each round, all voters and candidates are resampled.

Arguably, a Euclidean space serves as a good approximation of our intended application scenario of recommender systems (see Chapter 5). Even though recommendations and stakeholder interests may evolve from round to round —allowing voters (e.g., representing fairness concerns) to be regenerated every round— it is reasonable to assume a certain structure. For example, two related fairness concerns —such as location fairness (e.g., avoiding recommendations biased toward the US) and nationality fairness (e.g., diversifying beyond US citizens)— are expected to be located near each other and to have similar rankings. In case of the music streaming platform example in the introduction (1.1), two voters with similar attributes (e.g., classical music listeners from Europe) are expected to be located near each other and to have similar rankings.

In our experiment, randomness is required for data generation and execution of the weighted RSD rule. To ensure reproducibility, we use fixed random seeds at both levels. As each data set has a unique seed, the results are reproducible. We simulate 30 different instances, each consisting of 50 voters, 5 candidates, and 30 rounds. Since we generate 30 rounds of data for each of the 30 instances, we create 900 sets of 50 voters and 5 candidates.

The workflow of the experiment is as follows: First we generate the data. Then, we process each instance — i.e., each preference profile — through all possible combinations of SPFs, weight update functions, and initialization strategies, which together form TSPFs. This procedure is repeated for 3 different satisfaction threshold values. Each TSPF returns a result: a list of 30 output rankings corresponding to the total number of rounds, which is then analyzed by a result processor using the predefined metrics. After the analysis, we visualize the values obtained.

Due to the large number of instances, we compute average values and generate scatter plots for each combination of metric, initialization strategy, and satisfaction threshold. In each plot, the x-axis represents the update functions and the y-axis shows the corresponding values. Different point colors indicate different SPFs.

We continue with the TSPFs that we studied. Each TSPF consists of an SPF, a weight update function, and an initialization strategy (see Chapter 3 for details). Additionally, some weight update functions' behavior depends on the threshold value determining whether a voter is satisfied. We list thus each of these that we used in our simulations.

First, we list the SPFs (given with the corresponding string names in the code):

- Weighted Borda ("borda") (3.1)
- Weighted Kemeny ("kemeny") (3.2)
- Weighted Squared Kemeny ("sq-kemeny") (3.3)
- Weighted Random Serial Dictatorship ("weighted-rsd")

We continue with the weight update functions (given with the corresponding string names in the code):

- Constant weight update function ("constant") (3.4)
- Myopic KT weight update function ("myopic-kt") (3.5)
- Myopic squared KT weight update function ("myopic-sq-kt") (3.6)
- KT weight update function ("kt") (3.7)
- Squared KT weight update function ("sq-kt") (3.8)
- Unit cost weight update function ("unit-cost") (3.11)
- KT reset weight update function ("kt-reset") (3.9)
- Unit cost reset weight update function ("unit-cost-reset") (3.10)
- Perpetual KT weight update function ("perpetual-kt") (3.12)
- Myopic KT maximum value weight update function ("myopic-kt-special-voter") (3.13)
- KT sum weight update function ("kt-special-voter") (3.14)

Next, we list the studied weight initialization strategies (see 3.1.1). Again, provided with their corresponding string names in the code:

- equal weights ("equal")
- special voter's weight is about 25% of the total weight ("special-voter-25-percent")

Finally, we studied the following satisfaction threshold values, denoted by α (note that for 5 alternatives — as used in our instances — the maximum KT distance is 10):

- $\alpha = 0$
- $\alpha = 3$
- $\alpha = 7$

In total, there are 4 social preference functions, 11 weight update functions, 2 weight initialization strategies, and 3 satisfaction threshold values. The total number of TSPFs are 88 and each of them are tested with 3 satisfaction values, i.e., we test 30 instances with 264 different configurations.

4.2 Metrics

In this section, we introduce the metrics used to assess the various proportionality properties of the TSPFs that we study. In total, we define 8 metrics. Two of them, perpetual lower quota compliance and Gini influence coefficient, are adapted from Lackner [42]. We evaluate them for all combinations of TSPFs, satisfaction threshold values, and instances. Subsequently, we compute the average value for each metric across the instances. Given the large number of configurations, this averaging allows for a good visualization with scatter plots. Note that median values are less suitable in this context, as they ignore the influence of outliers, which might be informative in our setting because we are also interested in edge cases.

4.2.1 Perpetual Lower Quota Compliance

Perpetual lower quota compliance measures the proportion of voters whose satisfaction up to a given round is at least as high as the cumulative support their preferences have received from other voters. In other words, a voter is considered not underrepresented if they have been satisfied in more rounds than the total number of times other voters have supported similar preferences in the current round.

The value is normalized by the total number of voters. Values close to 0 indicate that even if a voter's preference ranking receives support from other voters, the satisfaction of the voter is low, i.e., the voter is underrepresented. In contrast, values closer to 1 imply that the voter is satisfied more often than their preferences are supported by others, i.e., the voter has proportional representation or even more.

$$compl = \frac{1}{o \cdot |N|} \sum_{t=1}^{o} |\{v \in N \mid sat_t(v) \geq \lfloor qu_t(v) \rfloor\}|$$

4.2.2 Gini Influence Coefficient

Gini influence coefficient measures how voter influence is distributed and whether discrepancies between individual levels of influences are large. If the value is closer to 1, there is a higher level of inequality, i.e., a few voters have much larger influence on the outcome compared to other voters. Conversely, if it is closer to 0, the influence of each voter on the outcome is more balanced. This indicates a lower level of inequality.

To calculate the Gini influence coefficient, we first define the *influence* of voter $v \in N$ up to round o and provide two different definitions.

We begin with the threshold-based definition (TBD), which we have already agreed to use throughout the thesis. In this definition, $\mathbb{1}_{d_K(\succ_v^t, \succ_{out}^t) \leq \alpha}$ denotes the indicator function that is 1 if $d_K(\succ_v^t, \succ_{out}^t) \leq \alpha$, and 0 otherwise:

$$infl_o(v) = \sum_{t=1}^{o} \frac{\mathbb{1}_{d_K(\succ_v^t, \succ_{out}^t) \leq \alpha}}{|\{u \in N \mid d_K(\succ_u^t, \succ_{out}^t) \leq \alpha\}|}$$

For the sake of completeness, it is worth noting that the continuous definition (CD) may also be considered —at least for simulation purposes — despite our decision to adopt the threshold-based definition (TBD) throughout this thesis. In future work, results obtained using the CD could be compared to those derived from the TBD to evaluate potential differences in influence measurement. Below, we present the continuous definition:

$$infl_o(v) = \sum_{t=1}^{o} \frac{d_{max} - d_K(\succ_v^t, \succ_{out}^t)}{\sum_{u \in N} d_{max} - d_K(\succ_u^t, \succ_t^*)}$$

Next, we introduce *average influence*, which we need to normalize the Gini influence coefficient. Let a be the average influence:

$$a = \frac{\sum_{v \in N} infl_o(v)}{|N|}$$

Now, we define the Gini influence coefficient for round o:

$$gini_o = \frac{1}{2a \cdot |N|^2} \sum_{u \in N} \sum_{v \in N} |infl_o(u) - infl_o(v)|$$

4.2.3 Average KT Distance

Average KT distance per voter measures the average Kendall-tau distance for a voter $v \in N$ in the entire voting process after the final round $o = |T|$. A value closer to the maximum KT distance d_{max} (introduced as Formula 3.17) indicates that the voter's preferences significantly deviate from the output rankings. Conversely, a lower value indicates less deviation between the voter's rankings and the outcomes. The definition is as follows:

$$avg_{d_K}(v) = \frac{\sum_{t=1}^{o} d_K(\succ_v^t, \succ_{out}^t)}{|T|}$$

Overall average KT distance computes the average of the individual average KT distances of all voters up to round $o = |T|$. Higher values suggest higher disagreement between voters' preferences and the outcomes of the TSPF. Lower values indicate general alignment of preference rankings and output rankings. The definition is as follows:

$$avg_{d_K}(N) = \frac{1}{|N|} \sum_{v \in N} avg_{d_K}(v)$$

4.2.4 Average Squared KT Distance

Average squared KT distance per voter measures the average squared Kendall-tau distance for a voter $v \in V$ in the entire voting process after the final round $o = |T|$. Squaring the distances penalizes larger deviations more heavily, making this metric more sensitive to extreme mismatches between preferences and outcomes. Again, a value closer to the maximum squared KT distance d^2_{max} indicates that the voter's preferences significantly deviate from the output rankings, probably also with more extreme deviations than in average KT distance case. Conversely, a lower value indicates less deviation between the voter's rankings and the outcomes. The definition is as follows:

$$avg_{d_K}{}^2(v) = \frac{\sum_{t=1}^{o} d_K(\succ_v^t, \succ_{out}^t)^2}{|T|}$$

Overall average squared KT distance computes the average of the individual average squared KT distances of all voters up to round $o = |T|$. Higher values suggest higher disagreement between voters' preferences and the outcomes of the TSPF, probably also more extreme than in average KT distance case. Lower values indicate general alignment of preference rankings and output rankings. The definition is as follows:

$$avg_{d_K}{}^2(N) = \frac{1}{|N|} \sum_{v \in N} avg_{d_K}{}^2(v)$$

4.2.5 Average Spearman Footrule Distance

Average Spearman footrule distance per voter measures the average Spearman footrule distance for a voter $v \in N$ in the entire voting process after the final round $o = |T|$. A higher value indicates that the voter's preferences significantly deviate from the output rankings and contain a high number of far displacements between the rankings. Conversely, a lower value indicates less deviation between the voter's rankings and the outcomes. The definition is as follows:

$$avg_{d_S}(v) = \frac{\sum_{t=1}^{o} d_S(\succ_v^t, \succ_{out}^t)}{|T|}$$

Overall Average Spearman's Footrule Distance computes the average of the individual average Spearman's footrule distances of all voters up to round $o = |T|$. Higher values suggest higher number of displacements of alternatives between voters' preferences and the outcomes of the TSPF. Lower values indicate general alignment of preference rankings and output rankings. Average Spearman's footrule distance is in general higher than the average KT distance, as not only the number of swaps but also the distance between displaced alternatives play a role. The definition is as follows:

$$avg_{d_S}(N) = \frac{1}{|N|} \sum_{v \in N} avg_{d_S}(v)$$

4.2.6 Egalitarian KT Distance

Egalitarian KT Distance finds the voter with the highest average KT distance after the final round $o = |T|$, which is the least satisfied voter according to KT distance. A lower value suggests that even the voter who agrees the least with the output rankings is relatively well represented. Conversely, a higher value indicates greater dissatisfaction for at least one voter in the system. The definition is as follows:

$$min_{d_K}(N) = \min_{v \in N} avg_{d_K}(v)$$

4.2.7 Standard Deviation of KT Distance

Standard Deviation of KT Distance measures how much the average KT distances vary across voters. This provides insights how equal or unequal voters are represented. The definition is as follows:

$$SD_{avg_{d_K}} = \sqrt{\frac{1}{|N|} \sum_{v \in N} \Big(\mathrm{avg}_{d_K}(v) - \mathrm{avg}_{d_K}(N) \Big)^2}$$

4.2.8 Total Number of Satisfaction for Special Voter

Total Number of Satisfaction for Special Voter measures the satisfaction of the special voter $v_n \in N$ in the final round $o = |T|$, reflecting in how many rounds the special voter is satisfied.

Values close to 0 indicate more than dissatisfaction: they signal a lack of proportionality. For proportional representation, at least a level as high as the cumulative support the special voter's preferences have received from other voters should be guaranteed (see perpetual lower quota compliance in 4.2.1). On the other hand, a value closer to o suggests that the special voter is overrepresented. Whether this is desirable depends on the specific application scenario. If a special voter is intended to receive more representation than others, then a total number of satisfaction closer to o may be considered positive. The definition corresponds to the TBD of satisfaction in Chapter 3.2:

$$sat_o(v) = |\{\ell \leq o : d_K(\succ_v^\ell, \succ_{out}^\ell) \leq \alpha\}|$$

4.3 Hypotheses

In the experiment, we study the following hypotheses:

- H1 Weighted RSD does not perform well in any of the results.
- H2 Weighted Squared Kemeny has a better Gini influence coefficient and worse perpetual lower quota compliance than others.

H3 Constant update function is worse than the others.

H4 Myopic KT-distance-based weight update functions are worse in every metric than KT-distance-based weight update functions.

H5 Unit cost is worse in Gini influence coefficient than unit reset.

H6 Initial weight does not have a big influence (except for constant weight update function), as round after round everything is updated.

H7 Spearman Footrule and KT distances behave similarly and consistently favor weighted Kemeny, followed by weighted Borda, weighted Squared Kemeny, and weighted RSD.

H8 Weighted Squared Kemeny has a better average squared KT-distance than weighted Borda and weighted Kemeny.

H9 KT sum weight update function performs better than myopic KT maximum value weight update function in satisfying the special voter.

H10 Both special-voter-sensitive functions are worse than all other update functions in every metric except the total number of satisfaction for special voter.

4.4 Results

In this section, we present the results of our experiments and evaluate them with respect to the hypotheses defined earlier. First, we interpret the outcomes. Then, we examine each hypothesis individually and assess whether the experimental data supports or contradicts it.

In Figure 4.1, the scatter plots show the average KT distances over time with different threshold levels and weight initialization strategies. Most update functions (except the *constant* and *kt-special-voter*) yield consistently low KT distances (approximately 3.5-3.6) for the first three SPFs (*borda*, *kemeny*, *sq-kemeny*). These functions (e.g., *myopic-kt*, *kt, unit-cost*, etc.) appear to be effective in preserving ranking accuracy according to the KT distance. We observe high KT distances for *weighted-rsd* consistently (above 4.2).

In addition, the update functions *constant* (only with the initialization strategy *special-voter-25-percent*) and *kt-special-voter* increase the KT distance of *borda*, *kemeny*, *sq-kemeny* significantly, where *kemeny* is affected the most, followed by *borda. sq-kemeny* is relatively stable for all update functions. Surprisingly, the combination of Perpetual KT and weighted Borda delivers outliers with a much higher KT distance for threshold values 3 and 7. As a general trend, myopic weight update functions *myopic-kt* and *myopic-sq-kt* have slightly higher KT distances than their non-myopic counterparts *kt* and *sq-kt.*

In Figure 4.2, the scatter plots show the average Spearman footrule distances over time with different threshold levels and weight initialization strategies. The general trends correspond to the average KT distance case. Most update functions (except the *constant* and *kt-special-voter*) yield consistently low footrule distances (approximately 6.0-6.2) for the first three SPFs (*borda*, *kemeny*, *sq-kemeny*). These functions (e.g., *myopic-kt*, *kt, unit-cost*, etc.) appear to be effective in preserving ranking accuracy according to the Spearman footrule distance. We observe high footrule distances for *weighted-rsd* consistently (above 4.2).

In addition, the update functions *constant* (only with the initialization strategy *special-voter-25-percent*) and *kt-special-voter* increase the footrule distance of *borda*, *kemeny*, *sq-kemeny* significantly, where *kemeny* is affected the most. *sq-kemeny* is relatively stable for all update functions. Surprisingly, the combination of Perpetual KT and weighted Borda delivers outliers with a much higher footrule distance for threshold values 3 and 7. As a general trend, myopic weight update functions *myopic-kt* and *myopic-sq-kt* have slightly higher footrule distances than their non-myopic counterparts *kt* and *sq-kt*.

In Figure 4.3, the scatter plots show the average squared KT distances over time with different threshold levels and weight initialization strategies. Although the general trends corresponds to the average KT distance or average Spearman footrule distance, the biggest difference is that *sq-kemeny* is always at least slightly better than the other SPFs. Most update functions (except the *constant* and *kt-special-voter*) yield consistently low squared KT distances (approximately 15-17) for the first three SPFs (*borda*, *kemeny*, *sq-kemeny*). These functions (e.g., *myopic-kt*, *kt, unit-cost*, etc.) appear to be effective in preserving ranking accuracy according to the squared KT distance. We observe higher squared KT distances for *weighted-rsd* consistently (above 4.2).

In addition, the update functions *constant* (only with the initialization strategy *special-voter-25-percent*) and *kt-special-voter* increase the KT distance of *borda*, *kemeny*, *sq-kemeny* significantly, where *kemeny* is affected the most, followed by *borda*. *sq-kemeny* is relatively stable for all update functions. Like in average KT distance and average Spearman footrule distance, the combination of Perpetual KT and weighted Borda delivers outliers with a much higher squared KT distance for threshold values 3 and 7. As a general trend, myopic weight update functions *myopic-kt* and *myopic-sq-kt* have slightly higher squared KT distances than their non-myopic counterparts *kt* and *sq-kt*.

In Figure 4.4, the scatter plots show the egalitarian KT distance over time with different threshold levels and weight initialization strategies. The general trends corresponds to the squared KT distance. The SPF *sq-kemeny* outperforms others consistently. Most update functions (except the *constant* and *kt-special-voter*) yield consistently low squared KT distances (approximately 4.3-4.7) for the first three SPFs (*borda*, *kemeny*, *sq-kemeny*). These functions (e.g., *myopic-kt*, *kt, unit-cost*, etc.) appear to be effective in having a relatively well represented least satisfied voter according to the KT distance. We observe higher egalitarian KT distances for *weighted-rsd* consistently (above 5.3).

In addition, the update functions *constant* (only with the initialization strategy *special-voter-25-percent*) and *kt-special-voter* increase the KT distance of *borda*, *kemeny*, *sq-kemeny* significantly, where *kemeny* is affected the most, followed by *borda*. *sq-kemeny* is relatively stable for all update functions. Like in other cases, the combination of Perpetual KT and weighted Borda delivers outliers with a much higher squared KT distance for threshold values 3 and 7. As a general trend, myopic weight update functions *myopic-kt* and *myopic-sq-kt* have slightly higher egalitarian KT distances than their non-myopic counterparts *kt* and *sq-kt*.

In Figure 4.5, the scatter plots show the standard deviation of KT distance over time with different threshold levels and weight initialization strategies. The general trends corresponds to the squared KT distance. The SPF *sq-kemeny* has always a lower standard deviation than others (mostly lower than 0.3). Most update functions (except the *constant* and *kt-special-voter*) yield consistently low squared KT distances (approximately 0.2-0.4) for the first three SPFs (*borda*, *kemeny*, *sq-kemeny*). These functions (e.g., *myopic-kt*, *kt, unit-cost*, etc.) appear to be effective in having a relatively low standard deviation of KT distance. We observe — except a few exceptions — higher standard deviations for *weighted-rsd* consistently (above 5.3).

In addition, the update functions *constant* (only with the initialization strategy *special-voter-25-percent*) and *kt-special-voter* increase the standard deviation of *borda*, *kemeny*, *sq-kemeny* significantly, where *kemeny* is affected the most, followed by *borda*. *sq-kemeny* is relatively stable for all update functions. Like in other cases, the combination of Perpetual KT and weighted Borda delivers outliers with a much higher standard deviation for threshold values 3 and 7. As a general trend, myopic weight update functions *myopic-kt* and *myopic-sq-kt* have slightly higher standard deviation than their non-myopic counterparts *kt* and *sq-kt*.

In Figure 4.6, the scatter plots show the Gini influence coefficient with different threshold levels and weight initialization strategies. For threshold 0, *weighted-rsd* surprisingly performs best at the same level with *kemeny*, as satisfying any voter is difficult at this threshold. However, *sq-kemeny* is the most unequal SPF. The SPF *borda* shows moderate equality.

As the threshold increases to 3, *kemeny* clearly becomes the most fair, followed closely by *borda*. *weighted-rsd* now becomes the worst as the increased threshold allows more consistent satisfaction of preferences. For threshold 7, *sq-kemeny* outperforms all others, since it naturally aims to satisfy a broader range of voters, which aligns with the more inclusive threshold (maximal KT distance $d_{max} = 10$ for 5 alternatives). We also note that *kt-special-voter* and *constant* update functions lead to increased inequality, particularly when paired with *borda* or *kemeny*. Finally, Gini influence coefficient decreases as the threshold value increases, since it becomes easier to satisfy more voters, leading to a more equal distribution of influence.

In Figure 4.7, the scatter plots show the perpetual lower quota compliance with different threshold levels and weight initialization strategies. For threshold 0, *weighted-rsd* surprisingly performs best, even though its average satisfaction is poor in other metrics. For *constant* and *kt-special-voter*, which are weight update functions that usually do not adapt well for a proportional outcome, *kemeny* follows *weighted-rsd* closely. This happens, as satisfying any voter is difficult at this threshold if a rule prioritizes proportionality but *weighted-rsd* satisfy at least one voter per round. In contrast, *sq-kemeny* performs poorly at threshold 0 since it emphasizes proportionality and almost never satisfies a voter fully, having exactly the same preference ranking.

As the threshold increases to 3, *kemeny*, *borda*, *sq-kemeny* increasingly fulfill the lower quota requirement and take the lead from *weighted-rsd*. *sq-kemeny* still trails compared to *kemeny* and *borda* but improves slowly as the satisfaction threshold allows for more voter groups to be satisfied together. At threshold 7, *sq-kemeny* becomes the most consistent at satisfying lower quota and satisfies almost every voter. Conversely, *weighted-rsd* fails the most, as its random selection of a voter does not align with the maximization of the number of satisfied voters. Update functions *kt-special-voter* or *constant* behave slightly worse for *kemeny* and *borda*. The combination of *borda* with *perpetual-kt* leads again to worse results. However, in general, perpetual lower quota compliance increases as the threshold value increases, since it becomes easier to satisfy more voters, leading to better voter representation.

In Figure 4.8, the scatter plots show the total number of satisfaction for the special voter with different threshold levels and weight initialization strategies. Apparently, *kt-special-voter* satisfies the special voter in almost every round. Only with *sq-kemeny* at threshold value 0, the special voter's satisfaction is less than 5. Among the other update functions, *constant* stands out when used with the weight initialization strategy *special-voter-25-percent*, since the unchanged initial weight favors the special voter. All other update functions perform more or less the same, showing fewer satisfied rounds with *weighted-rsd*, but satisfaction increases with the rising threshold, peaking at threshold 7, where *borda* and sq-kemeny are the maximum value of 30 with any weight update function. Interestingly, although *myopic-kt-special-voter* is designed as special-voter-sensitive weight update function, the differences between it and, for example, *myopic-kt* are not noticeable. Note that some points which seem to be missing are tied exactly with another rule (point) such that they are visually covered by the other point. In these cases, when not all rules are tied, *borda* is usually tied with *sq-kemeny*, and *kemeny* with *weighted-rsd*.

H1

Weighted RSD is by far the worst performer in average KT distance, average Spearman's footrule distance, average squared KT distance, egalitarian KT distance, and standard deviation of KT distance. It also has the highest Gini influence coefficient and lowest perpetual lower quota compliance at threshold values 3 and 7. Only at threshold value 0, Gini influence coefficient is the lowest and perpetual lower quota is the highest. This distortion is caused by the low threshold value, which can satisfy a voter only if the output ranking exactly matches the preference ranking of the voter. While other SPFs aim to aggregate diverse voter preferences into a single output ranking that is reasonably acceptable for as many voters as possible, output rankings do not often correspond exactly to (at least) one of the preference rankings of voters. However, weighted RSD always selects one of the preference rankings as output ranking, creating a misleading exception. Consequently, weighted RSD does not perform well in any of the results, and we confirm H1.

H2

Weighted Squared Kemeny is expected to produce more proportional rankings than other SPFs, as it penalizes the extreme deviations heavier. This should enhance equality, lowering the Gini influence coefficient, probably at the cost of perpetual lower quota compliance, since fewer voters are satisfied. However, the results tell a different story. The Gini influence coefficient is the lowest only for threshold value 7 (nearly 0, meaning full equality), as it can be seen from the perpetual lower quota compliance values that are mostly 1.00 or close to 1.00.

For threshold value 0, the Gini influence coefficient is the highest and perpetual lower quota compliance is the lowest, as it is very rare that a voter's preference ranking is fully represented in the output ranking, due to the proportional nature of Squared Kemeny. Threshold value 3 yields better results than weighted RSD but is still worse than weighted Kemeny and weighted Borda in most cases.

Therefore, we conclude that Hypothesis H2 stating that weighted Squared Kemeny has a better Gini influence coefficient and worse perpetual lower quota compliance than others is only partially supported by our results.

H3

Constant weight update function is a more or less average weight update function in average KT distance, average Spearman footrule distance, average squared KT distance, egalitarian KT distance, standard deviation of KT distance, or Gini influence coefficient when the initialization strategy is equal weights. However, the resulting values become significantly less favorable, if the special voter's weight is about 25% of the total weight. Performance notably gets worse for weighted Borda, weighted Kemeny, and weighted Squared Kemeny. This can be explained by the fact that the weights are never adapted so the output ranking remains unchanged for each round. For perpetual lower quota compliance, there is no consistent drop in value but in one of the most representative case — threshold 3 – weighted Borda and weighted Kemeny both yield one of the worst values with constant weight update function compared to other weight update functions, and this happens for both initialization strategies, although it is more obvious when special voter is initialized with 25% of the total weight. Since all metrics except one show noticeable differences, we conclude that Hypothesis H3 is true: Constant weight update function is worse than the others.

H4

Myopic KT-distance-based weight update functions, such as the myopic KT and myopic squared KT weight update function, are worse in every metric except Gini influence coefficient and perpetual lower quota than KT-distance-based weight update functions like the KT and squared KT weight update function. This holds if we leave out the weighted RSD rule, as this rule chooses a preference ranking randomly and is therefore not representative. In particular, the myopic KT weight update function generally has worse results than the KT weight update function and the myopic squared KT weight update function performs worse than its squared KT counterpart, although the differences are mostly minimal.

Regarding Gini influence coefficient, the differences are quite small and sometimes the myopic KT-distance-based weight update functions are even slightly better than the KT-distance-based weight update functions. Thus, it is not easy to identify a clear trend. It is an exceptable result, as the influence is still distributed with myopic KT-distance-based functions due to the weight updates.

For perpetual lower quota compliance, it is still the case that KT-distance-based weight update functions consistently have higher values than their myopic counterparts in general. The squared KT weight update function in particular has a significant advantage over the myopic squared KT weight update function. The only case, where the myopic KT update function performs better than the KT update function is with threshold 3 and when the special voter gets about 25% of the total weight under Kemeny rule. This is the case due to two factors:

1. Threshold 3 is generally a suitable value for perpetual lower quota compliance to identify real tendencies, as threshold 0 is overly restrictive and threshold 7 is too easy to satisfy.
2. The myopic function only considers weights from the previous round, thereby forgetting the initial weight advantage of the special voter and contributing to improved overall representability.

Squared KT distance does not capture this detail because its much larger values lead to stronger weight updates. In conclusion, Hypothesis H4 is confirmed with the exception of the Gini influence coefficient.

H5

Unit cost weight update function increases the weight of each unsatisfied voter by 1 after each round. Unit cost reset weight update functions additionally reset the weight of satisfied voters to 1, introducing a redistribution effect. Again, we ignore weighted RSD. As the satisfied voters lose all their weight and are reset to 1, it is expected to distribute the influences better and therefore yield a lower Gini influence coefficient, which is confirmed in our results with two exceptions: For equally initialized Kemeny with threshold 3 and Squared Kemeny with threshold 0, where the special voter is initialized with about 25% of the total weight. The latter can be ignored, as the difference is very low and it is not easy to satisfy any voter with Squared Kemeny and threshold 0, and resetting the weights of the satisfied voters does not make it better. This might lead to misleading results. The former is however more important but the real cause is an open question. Therefore, further simulations with new instances and different distributions are required to find the real cause. However, we can still partially confirm H5.

H6

Initial weights play the biggest role when used with the constant weight update function. As the constant weight update function does not adapt weights over time, output ranking stays the same for all rounds, tending to favor the voter that has a disproportionately large weight, namely the special voter. In particular, the weighted Kemeny rule tends to satisfy the voter group with the highest weight, which often includes the special voter. Weighted Borda occasionally exhibits similar behavior, whereas Squared Kemeny maintains some degree of proportionality, making it less sensitive to such initialization strategies.

Myopic KT-distance-based functions, including the myopic KT maximum value weight update function, forget the rounds before the last round so they are not affected at all after the second round. Other update functions are influenced in the beginning but the effect of the initial weights disappear as time progresses. For this reason, there might be some small differences between the myopic weight update functions and others in favor the myopic weight update functions, which are neglectable in comparison to the big difference with the constant weight update function. Therefore, we conclude that initial weight does not have a significant impact (except for constant weight update function), as weights are continuously updated each round. This confirms Hypothesis H6.

H7

The plots in Figure 4.1 and Figure 4.2 demonstrate that all the rules behave similarly with respect to both average KT distance and average Spearman footrule distance. This suggests that the results do not seem to depend on the specific distance measure we use. However, further research is warranted to substantiate this observation. Even though many rules aim to optimize for KT distance, weighted Kemeny is the best in both distance measures, followed by weighted Borda, weighted squared Kemeny, and weighted RSD. Therefore, both distance measures behave similarly, confirming H7.

H8

The plots in Figure 4.1 and Figure 4.3 indicate that the weighted Squared Kemeny rule is has the lowest average squared KT distance, followed by weighted Borda and weighted Kemeny for all TSPFs and threshold levels. It is expected to have such a result confirming that Squared Kemeny is designed to minimize the squared KT distance. This distance measure penalizes larger disagreements more heavily. Compared to Borda, Kemeny usually tends to reflect the preferences of majority voters more strongly (see 3.1) compared to Borda so weighted Kemeny has a higher average squared KT distance than weighted Borda. To sum up, weighted Squared Kemeny has a better average squared KT distance than weighted Borda and weighted Kemeny, confirming H8.

H9

The KT sum weight update function satisfies the special voter almost in every configuration across all rounds, whereas myopic KT maximum value weight update function does not have a significant advantage over non-special-voter-sensitive weight update functions. This shows that it does not suffice to receive the maximum weight in a each round, as in myopic functions, weights from the past rounds except the last round are ignored. The maximum weight alone cannot guarantee a total number of satisfaction above average, as other voters may have weights close to or equal to that of the special voter. Therefore, KT sum weight update function performs better than myopic KT maximum value weight update function in satisfying the special voter. In conclusion, we confirm H9.

H10

The myopic KT maximum value weight update function is not the best in perpetual lower quota compliance, especially at threshold value 3, as shown in Figure 4.7. In contrast, the KT sum weight update function performs worse than others in every metric except for the Gini influence coefficient — only for weighted Squared Kemeny — perpetual lower quota compliance —except Squared Kemeny — and at threshold 0. As previously discussed in other hypotheses, threshold 0 often produces misleading results; thus, this case is not a reliable exception. In summary, the myopic KT maximum value update function is not worse than all other update functions but the KT sum weight update function is. This partially confirms H10.

4.5 Discussion

We find that the TSPFs mostly behave in accordance with our hypotheses. Accordingly, we could confirm H1, H2, H3, H6, H7, H8, and H9. However, some unexpected outcomes prevented full confirmation of H4, H5, and H10.

If we compare the experimental results with the axiomatic results, axiomatic results have limited power to reflect practical behavior, as many complex scenarios are not covered. Still, there is no contradiction between theoretical and empirical results: The theoretical insights apply to restricted settings, while the simulations represent more complex scenarios.

Due to limited resources, simulations were conducted on 30 instances with 30 rounds each, generated using uniform distributions. If we extend these simulations by new instances, more rounds, and different distributions, we can further confirm our results or identify new patterns. Additionally, it might also be an interesting approach to test the results with the continous definition of the Gini influence coefficient and compare them with the results of the threshold-based version.

Moreover, we resampled the voters and candidates after each round. It might be interesting and could change the behavior of certain update functions if updates on preferences were restricted. This might increase the discrepancy between some update functions, as the randomness of resampling in our setting might hide the insufficient weight adjustments of some rules. For example, the constant weight update function initialized with equal weights shows average performance, as voter preferences change a lot even though the weights do not change. If this effect disappears, the comparability of weight update functions might become much easier.

Two open questions regarding the results remain: The TSPF resulting from the combination of weighted Borda and the perpetual KT weight update function produces outlier results, especially at threshold levels 3 and 7. The second question is, identifying the degree of weight dominance at which a special-voter-sensitive function starts to favor the special voter more without worsening the other metrics too much. As we showed with our two example special-voter-sensitive update functions, balancing these effects is a non-trivial challenge.

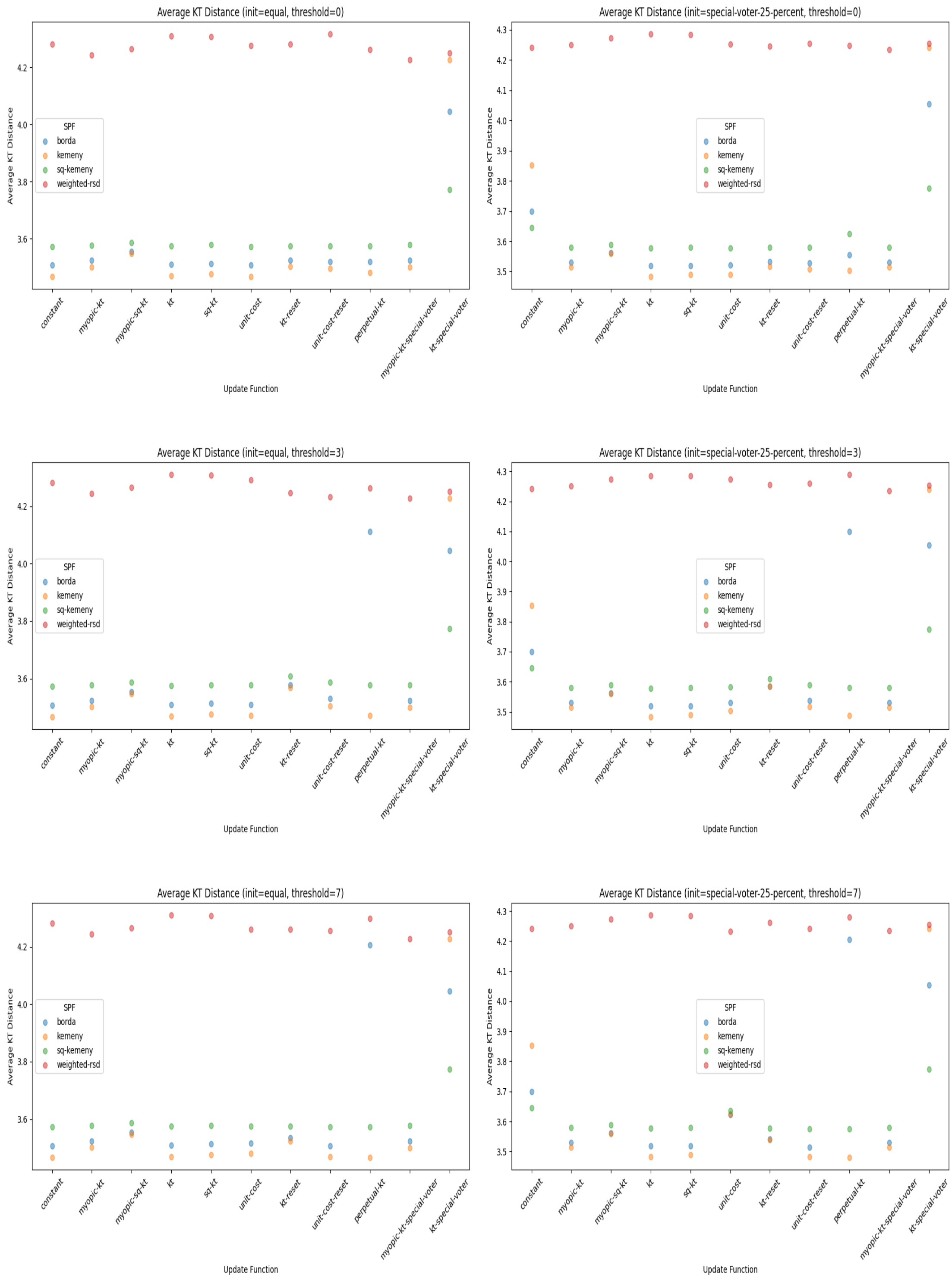

Figure 4.1: Average KT distances over time with different threshold levels and weight initialization strategies.

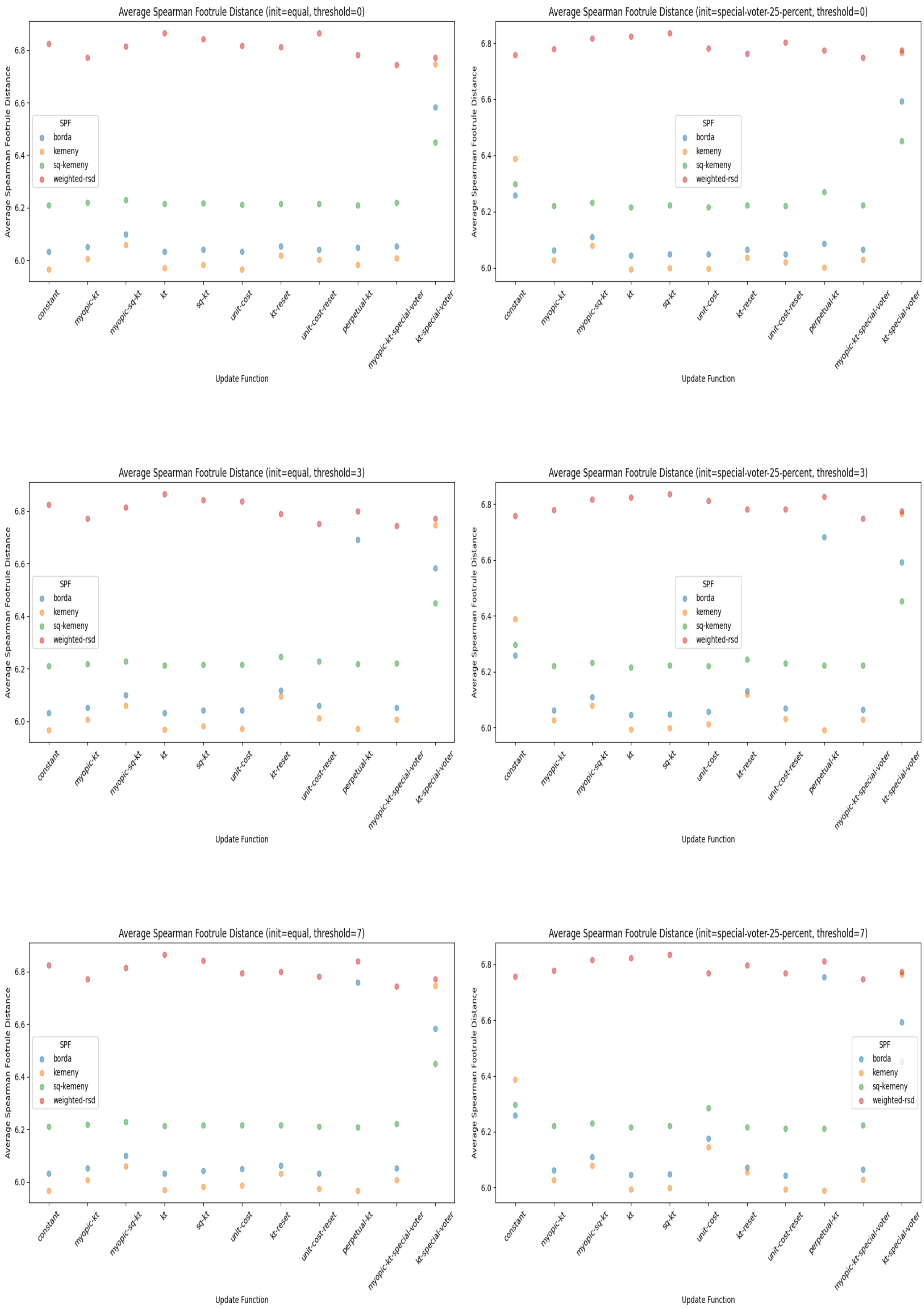

Figure 4.2: Average Spearman footrule distances over time with different threshold levels and weight initialization strategies.

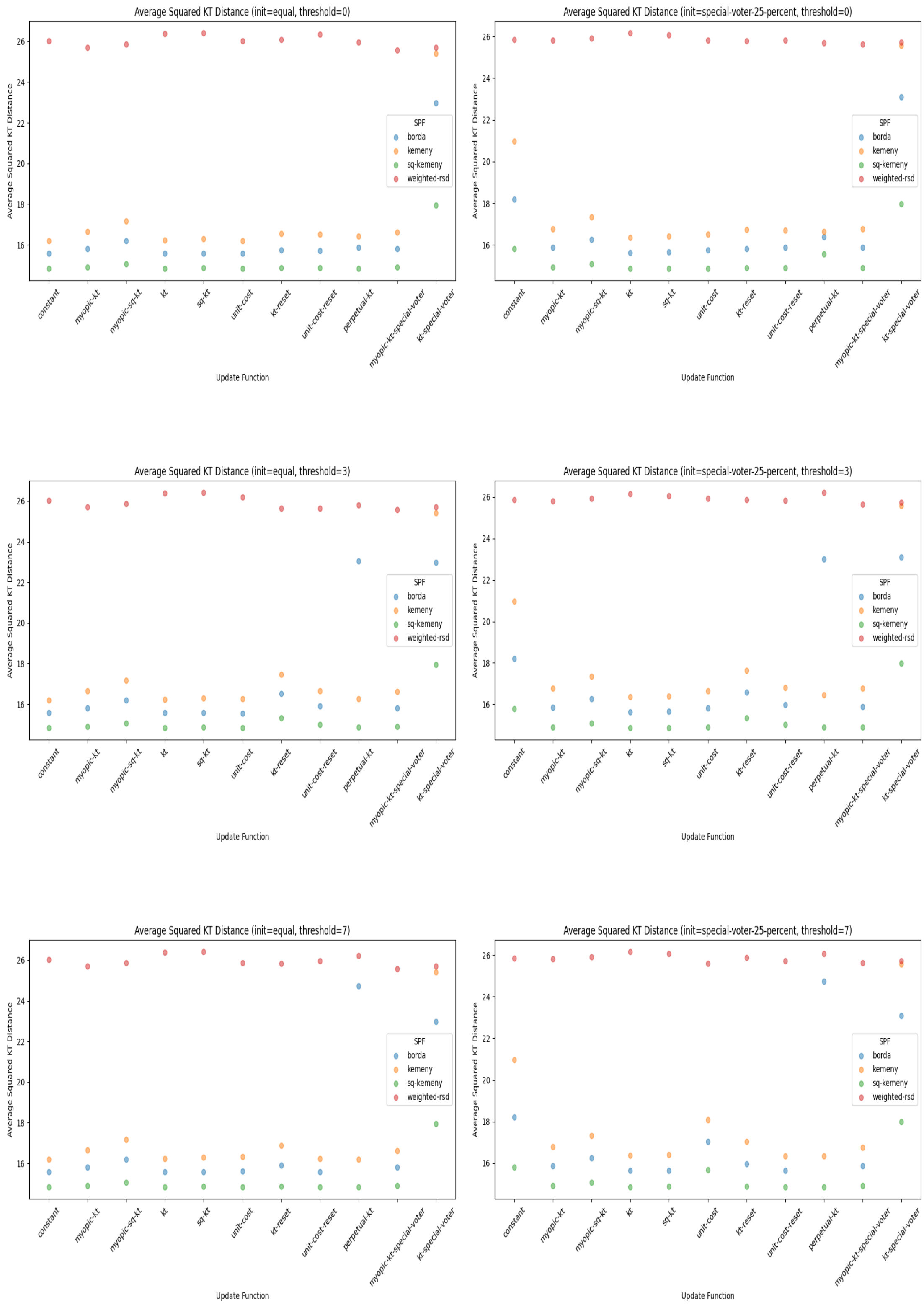

Figure 4.3: Average squared KT distances over time with different threshold levels and weight initialization strategies.

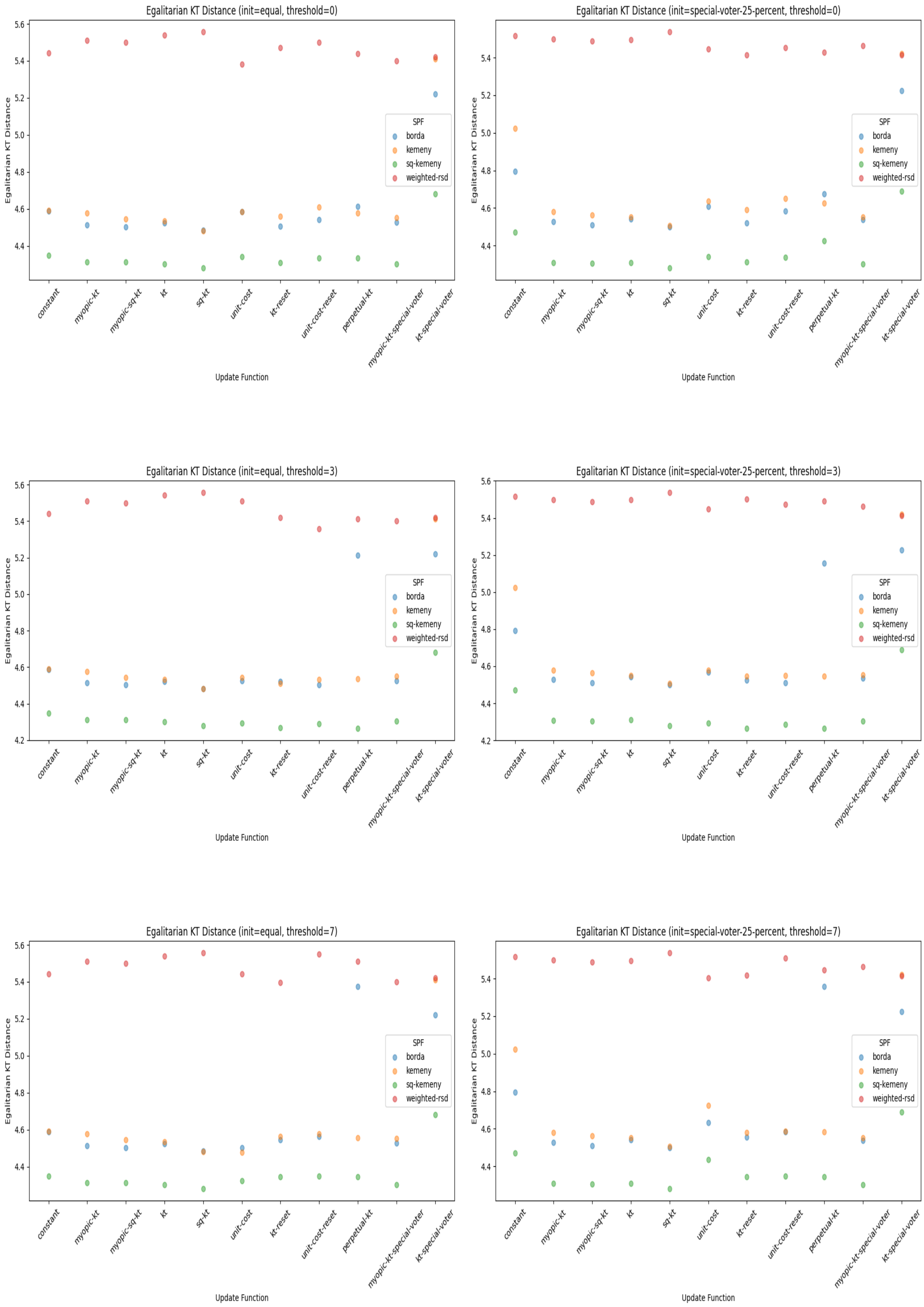

Figure 4.4: Average egalitarian KT distances over time with different threshold levels and weight initialization strategies.

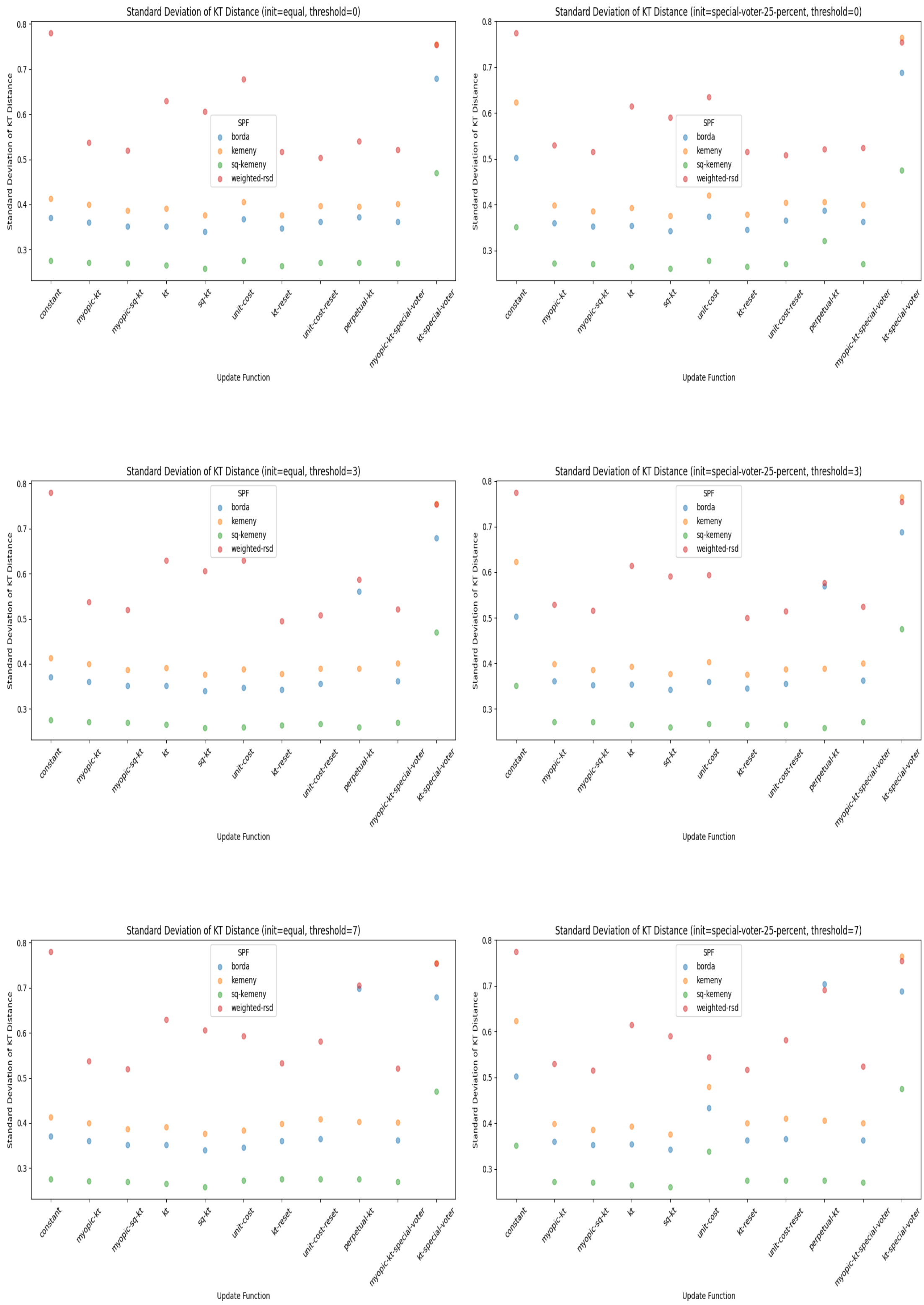

Figure 4.5: Average standard deviations of KT distances over time with different threshold levels and weight initialization strategies.

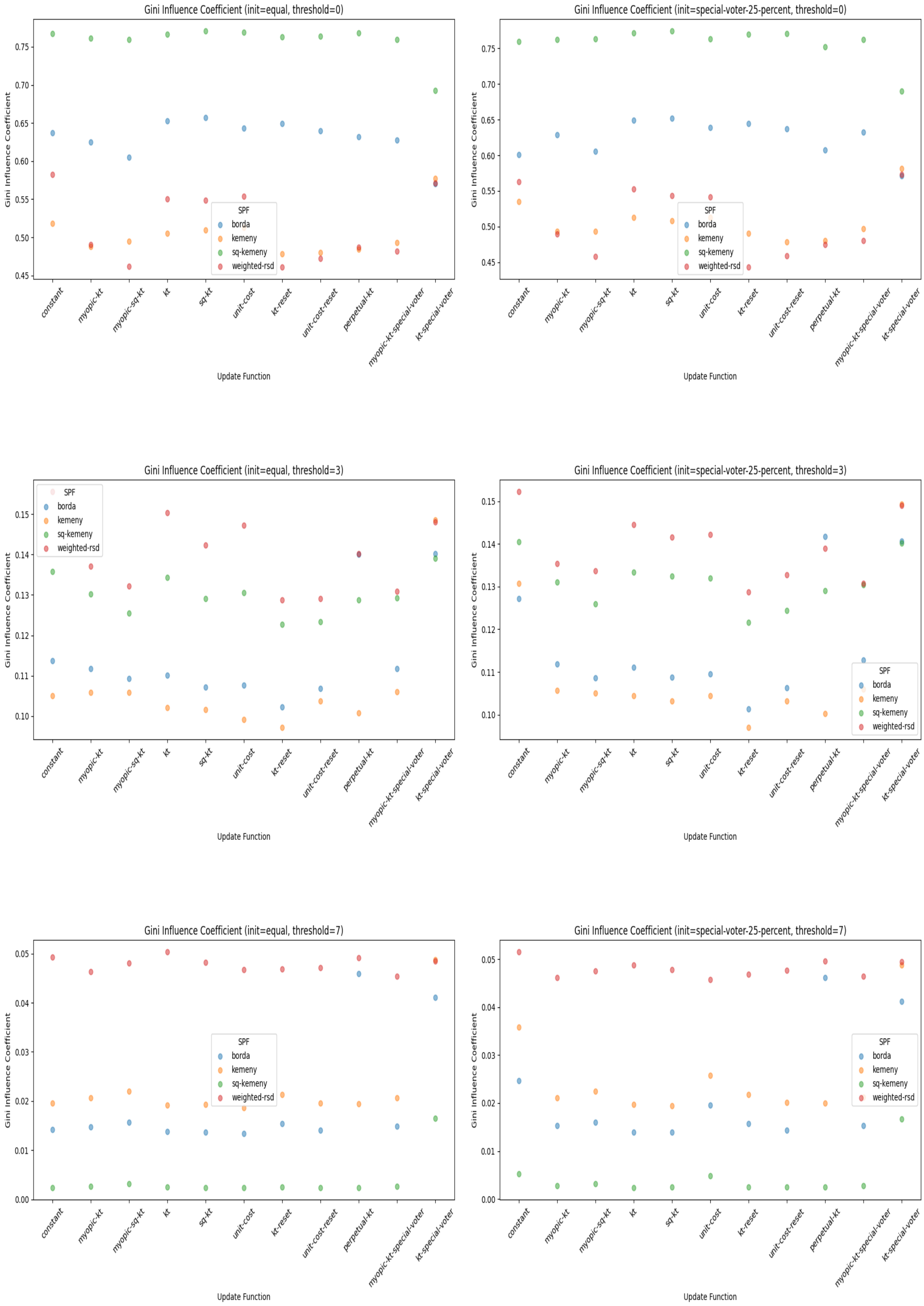

 Figure 4.6: Average Gini influence coefficients with different threshold levels and weight initialization strategies.

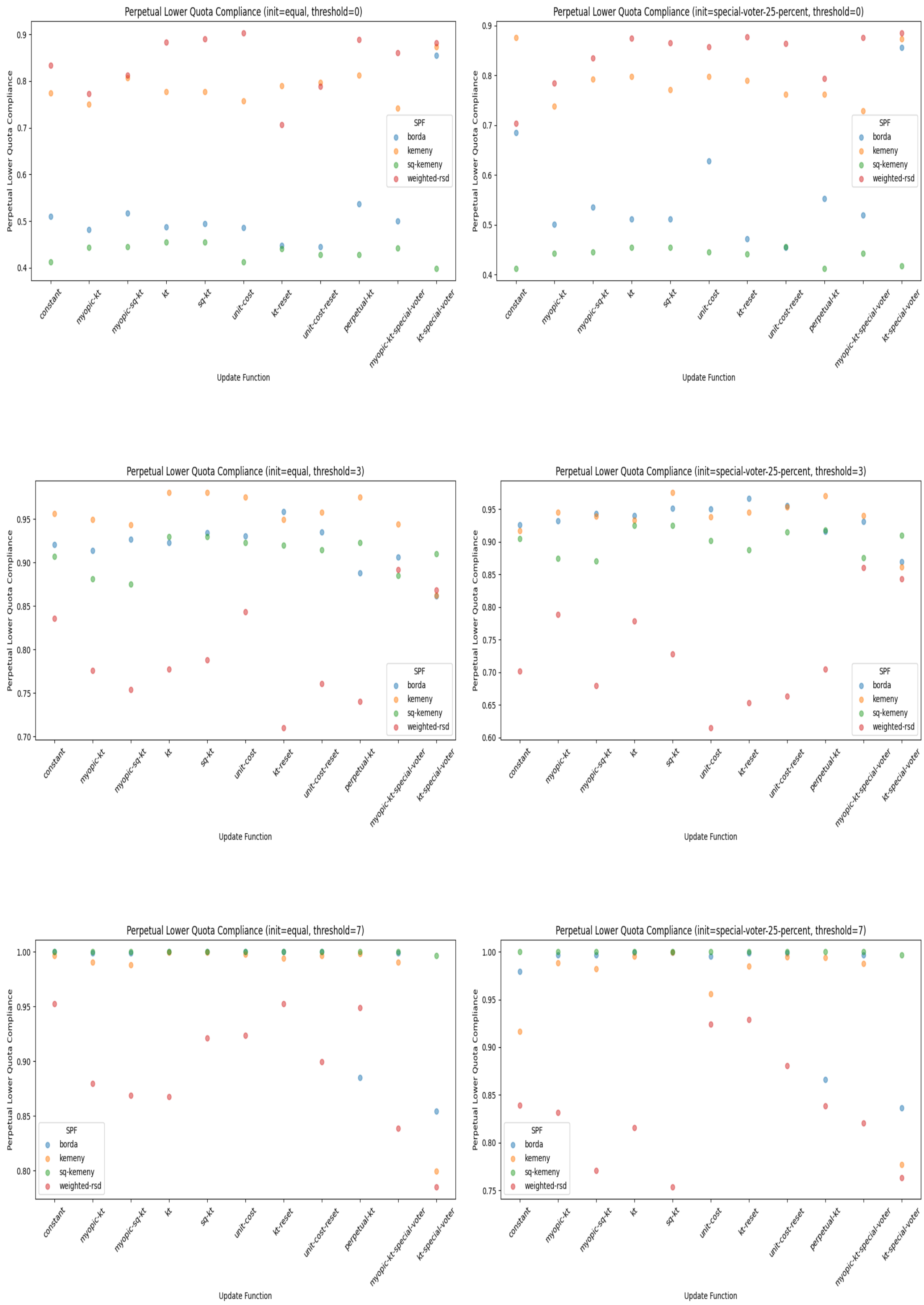

Figure 4.7: Average perpetual lower quota compliance values with different threshold levels and weight initialization strategies.

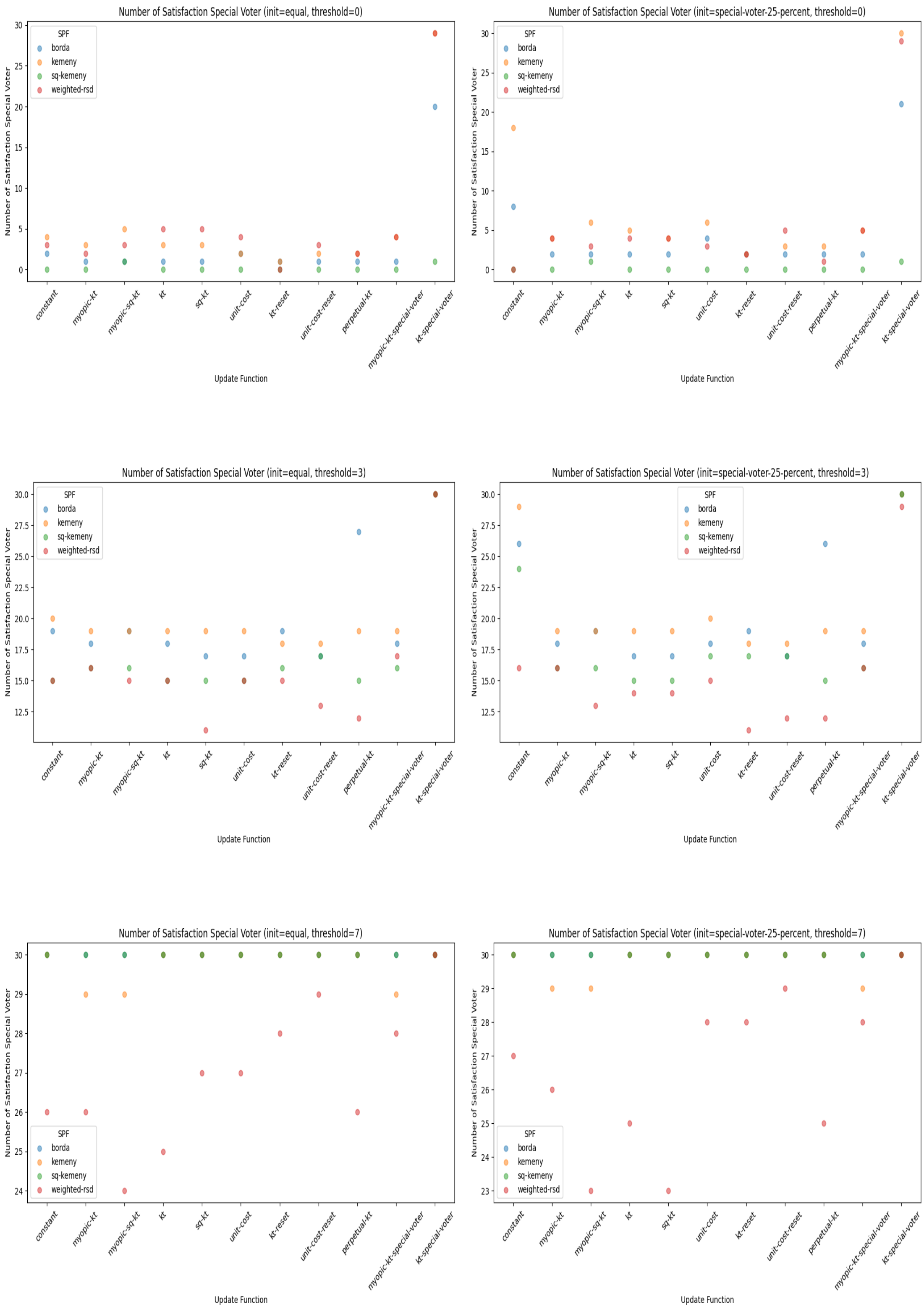

Figure 4.8: Average total number of satisfied rounds for the special voter with different threshold levels and weight initialization strategies.

CHAPTER 5

Potential Application: Recommender Systems

A *recommender system* is a tool that suggests items to a user. In the age of big data and information overload, recommender systems support users in making decisions across many areas such as which news to read, which products to buy, and which films to watch [53]. This chapter is concerned with a possible way of using our model of *perpetual individualized fair rank aggregation* to design a recommender system.

Fairness and achievement of objectives in recommender systems are complex and challenging problems [27], as there are diverse goals and fairness concerns, each relevant to different stakeholders, and these interests sometimes cannot be aligned [2, 1]. Moreover, fairness should be addressed over time in dynamic settings, where the preferences of stakeholders or their interactions with the system and items may evolve over time [3, 36]. In addition, recommender systems emphasize individualized recommendations for each user. Therefore, the problem of multi-stakeholder fair recommendation can be modeled by our framework.

In Figure 5.1, we illustrate how our framework can be applied to the domain of recommender systems. In this scenario, the *special voter* represents the initial recommendation list generated by a recommender system algorithm. In addition to this base recommendation, the system considers multiple voters, each modeling a different fairness concern or stakeholder objective, which we refer to as *fairness agents* in the following. These agents submit their own preference rankings over items, aiming to influence the final recommendation in ways that reflect their specific goals.

For instance, one agent may represent a *popularity bias correction* mechanism, promoting underrepresented items that are typically ranked low due to lack of popularity. Another agent might focus on *provider fairness*, ensuring that items from different producers (e.g.,

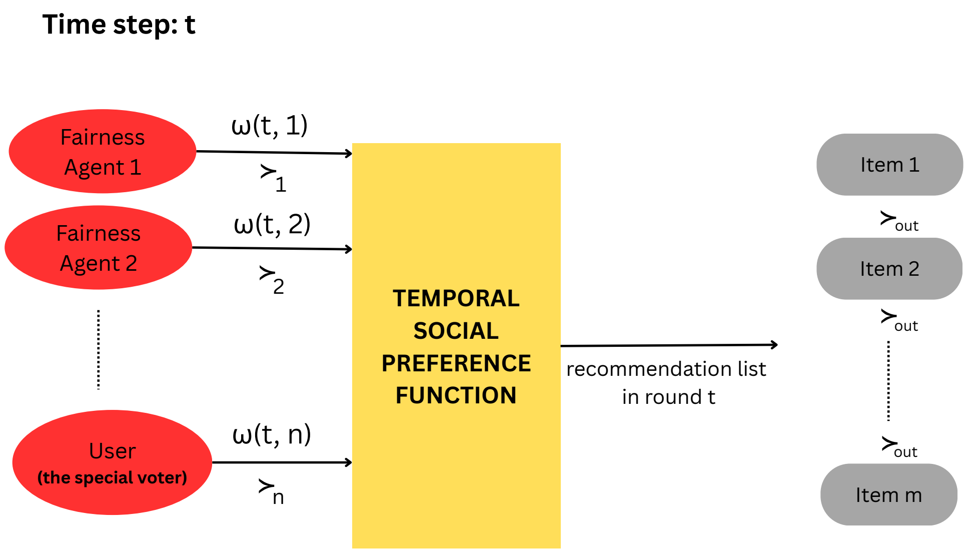

Figure 5.1: Perpetual individualized fair rank aggregation in the context of multi-stakeholder fair recommendation.

small vendors or minority creators) are fairly represented in the recommendation. Yet another agent might represent *content diversity*, favoring a wider variety of item types.

A TSPF produces a recommendation list based on the input preferences. After generating a final recommendation list (i.e., the aggregated ranking in our model) in a specific round, the preferences and weights of agents are updated. Agents that are more satisfied in one round — meaning their preference rankings resemble the recommendation list the most and accordingly their preferred items were ranked near the top — may see their influence reduced in the next round. Conversely, unsatisfied agents may be given more weight to correct for underrepresentation over time. User preferences can also evolve: new users may arrive, or a user's recent interaction behavior may shift their inferred preference profile.

The special voter's weight (representing the base recommender's influence) can be controlled to ensure a balance between fairness and accuracy. For example, we may enforce a lower bound on the user's weight to prevent fairness agents from fully overriding individualization. This is especially important in most cases, at least where relevance to the user is a primary business goal, such as e-commerce platforms aiming to optimize for user engagement or conversions. However, in contexts like public-service media or educational content delivery, it may be desirable that fairness concerns occasionally dominate — e.g., by pushing informative but less-clicked content — even at the cost of short-term accuracy. This is aligned with the proportionality principles introduced in Chapters 3 and 4 such as *simple proportionality*, *bounded dry spells*, and *perpetual lower* *quota (compliance).*

Moreover, the special voter might attempt to manipulate outcomes by submitting rankings that align with those of fairness agents, in order to shift weight distributions in future rounds. This is especially relevant when fairness agents agree on similar rankings — e.g., when multiple fairness goals point to the same items, or when the fairness agents — just like the special voter — try to find opportunities to manipulate strategically over several rounds. In such cases, the property of *independence of uncontroversial items* becomes particularly important, as it helps to prevent coordinated manipulation of the system by exploiting consensus.

This setup reflects a dynamic, long-term balancing process over multiple interests. Depending on the application and concrete implementation of the TSPF, different fairness agents may need to have stronger or weaker influence in different rounds. Our model enables this kind of adaptive trade-off, where fairness objectives are considered proportionally over time rather than enforced in each separate recommendation.

A promising future direction would be to allow fairness agents to take the user's current preferences into account before generating their own rankings. This would make it possible to design agents that respond more contextually. For instance, they can adapt their goals based on the user's behavioral patterns derived by the user's current preferences, which enables more nuanced and effective fairness interventions.

CHAPTER 6

Conclusion & Future Work

We conclude the thesis with a brief summary of the main contributions and a discussion of future work. In this thesis, we extended the formalism of *perpetual voting* to settings where both the input and output are preference rankings. For this purpose, we introduced the concept of *temporal social preference functions* (TSPFs), each consisting of a *social preference function* (SPF), a *weight update function*, and a *weight initialization strategy*. Accordingly, we introduced various SPFs, weight update functions, and two weight initialization strategies.

We analyzed axiomatically the properties of the TSPFs in Chapter 3, in particular the *simple proportionality*, *independence of uncontroversial decisions*, and *bounded dry spells*. These properties serve different purposes: simple proportionality gives proportionality guarantees, independence of uncontroversial decisions is important to avoid manipulative behavior, and bounded dry spells ensures that each voter is guaranteed to be represented in some round.

Our model is more expressive compared to the approaches based on approval ballots, as it allows rankings instead of approvals. However, this comes at a price: Complexity of the model increases and it becomes more difficult to yield formal results. We still managed to obtain some interesting results. We showed that independence of uncontroversial decisions cannot be satisfied in the general case, even when the weights of satisfied voters are multiplied by constants other than one. Only if the weights of satisfied voters remain unchanged — for example, under the *unit cost weight update function* — and every voter is satisfied with the outcome, the property is satisfied. Furthermore, we proved that simple proportionality holds in some restricted settings, such as for weighted Kemeny with certain weight update functions, when voters are partitioned into two groups where all group members have the same preference ranking. Additionally, we demonstrated that many weight update functions do not guarantee bounded dry spells.

We also evaluated the model experimentally, simulating all TSPFs under various parameter settings (see Chapter 4). Our results show that the choice of parameters — not only the SPF, weight update function, and initialization strategy but also the *satisfaction threshold* α —plays a significant role. While the effect of the weight initialization strategy diminishes over time, this happens fastest for *myopic weight update functions*, which ignore information from earlier rounds. However, an SPF or a weight update function has more permanent long-term effects.

In our experiments, we observed that the differences between weight update functions are only slightly reflected in the results. For instance, in almost all configurations and metrics, it is the case that *myopic weight update functions* perform worse than their non-myopic KT-distance-based counterparts in general but the difference is not substantial. This may be due to the fact that evolving preferences introduce some randomness, which reduces the importance of weights — an interesting observation to be examined for future research. It is, however, evident that the *constant weight update function* and the *KT sum weight update function* do not deliver good results. Thus, the choice of weight update function remains a decisive factor in designing a TSPF that ensures fairness and proportionality, although the differences are not always as noticeable as they typically are in the choice of SPFs.

Among the studied SPFs, *weighted Borda*, *weighted Kemeny*, and *weighted Squared Kemeny* performed significantly better than *weighted Random Serial Dictatorship*. Additionally, there are some small differences: *Weighted Kemeny* provided a good overall representation for voters but *weighted Squared Kemeny* favored more equal satisfaction across voters. For a compromise, *weighted Borda* seems to be a good choice that offers competitive results with lower computational complexity. This makes it attractive for potential real-world application scenarios.

In Chapter 5, we sketched a concept of how to apply our model in *multi-stakeholder fair recommender systems*. The task of developing such systems and adapting theoretical models to practical implementations is both crucial and challenging, and requires various research efforts from different disciplines.

Several theoretical and practical questions remain open. Future research could explore new TSPFs by defining alternative SPFs, weight update mechanisms, and initialization strategies that we did not study in this thesis. These TSPFs can be examined both axiomatically and experimentally.

Additional axioms from the social choice literature —such as *justified representation* [10] — can be adapted to our setting. Simpler and more flexible notions of proportionality may also be defined to capture more general or complex voter profiles.

However, even within our settings, there are many open questions. For example, it remains unclear whether *weighted Squared Kemeny* combined with *reset-based weight update functions* has a bounded dry spell guarantee.

Finally, the experimental results also suggest several directions of further investigations.

First, this experiment can be repeated with different numbers of rounds or different preference distributions to check which effects these factors have on the outcome and to confirm our own results. Moreover, we observed unexpected behavior from the combination of *weighted Borda* with *perpetual KT update function*, producing outliers. The reason for this unexpected behavior could be studied. Understanding these interactions more deeply may lead to better design of perpetual rank aggregation systems. Lastly, one could study new *special-voter-sensitive functions* to identify the threshold at which increasing the special voter's weight starts to favor the special voter more without worsening the other metrics too much. The *special-voter-sensitive update functions* we implemented demonstrate that balancing these effects is a non-trivial challenge.

Overview of Generative AI Tools Used

Generative AI tools were used to obtain high-level feedback regarding the written text. Specifically, Chat-GPT variant 4o and Copilot with Microsoft Edge (no specific model variant provided) were asked to "correct and refine the text if necessary" on existing text only. Their output was not directly used to change the text but served as feedback. The tools were accessed only in July 2025.

List of Figures

Bibliography

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Augusto Pizzato. Multistakeholder recommendation: Survey and research directions. *User Model. User Adapt. Interact.*, 30(1):127–158, 2020.

[2] Himan Abdollahpouri and Robin Burke. Multistakeholder recommender systems. In Francesco Ricci, Lior Rokach, and Bracha Shapira, editors, *Recommender Systems Handbook*. Springer, 3rd edition, 2020.

[3] Amanda Aird, Cassidy All, Paresha Farastu, Elena Stefancova, Joshua Sun, Nicholas Mattei, and Robin Burke. Exploring social choice mechanisms for recommendation fairness in SCRUF. *CoRR*, 2023.

[4] Amanda Aird, Paresha Farastu, Joshua Sun, Elena Stefancova, Cassidy All, Amy Voida, Nicholas Mattei, and Robin Burke. Dynamic fairness-aware recommendation through multi-agent social choice. *Trans. Recomm. Syst.*, 3(2):21:1–21:35, 2025.

[5] Amanda Aird, Elena Stefancova, Cassidy All, Amy Voida, Martin Homola, Nicholas Mattei, and Robin Burke. Social choice for heterogeneous fairness in recommendation. In Tommaso Di Noia, Pasquale Lops, Thorsten Joachims, Katrien Verbert, Pablo Castells, Zhenhua Dong, and Ben London, editors, *Proceedings of the 18th ACM Conference on Recommender Systems, RecSys 2024, Bari, Italy, October 14-18, 2024*, pages 1096–1101. ACM, 2024.

[6] Sanjeev Arora and Boaz Barak. *Computational Complexity - A Modern Approach*. Cambridge University Press, 2009.

[7] Kenneth Arrow. *The Possibility of a Universal Social Welfare Function*. RAND Corporation, Santa Monica, CA, 1948.

[8] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

[9] Kenneth J. Arrow. *Social Choice and Individual Values*. Cowles Foundation, New Haven, CT, 1st edition, 1951.

[10] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. Justified representation in approval-based committee voting. *Soc. Choice Welf.*, 48(2):461–485, 2017.

[11] Haris Aziz, Patrick Lederer, Dominik Peters, Jannik Peters, and Angus Ritossa. Committee monotonicity and proportional representation for ranked preferences. In Itai Ashlagi and Aaron Roth, editors, *Proceedings of the 26th ACM Conference on Economics and Computation, EC 2025, Stanford University, Stanford, CA, USA, July 7-10, 2025*, page 896. ACM, 2025.

[12] Haris Aziz and Nisarg Shah. Participatory budgeting: Models and approaches. In Tamás Rudas and Gábor Péli, editors, *Pathways Between Social Science and Computational Social Science: Theories, Methods, and Interpretations*, pages 215–236. Springer, 2021.

[13] Michel Balinski and Rida Laraki. *Majority Judgment: Measuring, Ranking, and Electing*. MIT Press, New Haven, CT, 2010.

[14] Francesco Barile, Tim Draws, Oana Inel, Alisa Rieger, Shabnam Najafian, Amir Ebrahimi Fard, Rishav Hada, and Nava Tintarev. Evaluating explainable social choice-based aggregation strategies for group recommendation. *User Model. User Adapt. Interact.*, 34(1):1–58, 2024.

[15] Péter Biró. Applications of matching models under preferences. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 20, pages 345–373. AI Access, 2017.

[16] Duncan Black. On the rationale of group decision-making. *Journal of Policital Economy*, 56(1):23–34, 1948.

[17] Gustavo Adrian Bodanza, Fernando Tohmé, and Marcelo Auday. Collective argumentation: A survey of aggregation issues around argumentation frameworks. *Argument Comput.*, 8(1):1–34, 2017.

[18] Niclas Boehmer, Piotr Faliszewski, Lukasz Janeczko, Andrzej Kaczmarczyk, Grzegorz Lisowski, Grzegorz Pierczynski, Simon Rey, Dariusz Stolicki, Stanislaw Szufa, and Tomasz Was. Guide to numerical experiments on elections in computational social choice. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, Jeju, South Korea, August 3-9, 2024*, pages 7962–7970. ijcai.org, 2024.

[19] Sylvain Bouveret. Social choice on the web. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 20, pages 387–402. AI Access, 2017.

[20] Steven J. Brams and Peter C. Fishburn. Voting procedures. In Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, editors, *Handbook of Computational Social Choice and Welfare*, volume 1, pages 173–236. Elsevier, 2002.

[21] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. Introduction to computational social choice. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors, *Handbook of Computational Social Choice*, pages 1–20. Cambridge University Press, 2016.

[22] Martin Caminada and Gabriella Pigozzi. On judgment aggregation in abstract argumentation. *Auton. Agents Multi Agent Syst.*, 22(1):64–102, 2011.

[23] Nikhil Chandak, Shashwat Goel, and Dominik Peters. Proportional aggregation of preferences for sequential decision making. In Michael J. Wooldridge, Jennifer G. Dy, and Sriraam Natarajan, editors, *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, pages 9573–9581. AAAI Press, 2024.

[24] Yann Chevaleyre, Ulle Endriss, Jérôme Lang, and Nicolas Maudet. A short introduction to computational social choice. In Jan van Leeuwen, Giuseppe F. Italiano, Wiebe van der Hoek, Christoph Meinel, Harald Sack, and Frantisek Plásil, editors, *SOFSEM 2007: Theory and Practice of Computer Science, 33rd Conference on Current Trends in Theory and Practice of Computer Science, Harrachov, Czech Republic, January 20-26, 2007, Proceedings*, volume 4362 of *Lecture Notes in Computer Science*, pages 51–69. Springer, 2007.

[25] Franz Dietrich. A generalised model of judgment aggregation. *Soc. Choice Welf.*, 28(4):529–565, 2007.

[26] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif. Intell.*, 77(2):321–358, 1995.

[27] Michael D. Ekstrand, Anubrata Das, Robin Burke, and Fernando Diaz. Fairness in information access systems. *Found. Trends Inf. Retr.*, 16(1-2):1–177, 2022.

[28] Ulle Endriss. Computational social choice: Prospects and challenges. In Elisabeth Giacobino and Rolf Pfeifer, editors, *Proceedings of the 2nd European Future Technologies Conference and Exhibition, FET 2011, Budapest, Hungary, May 4-6, 2011*, volume 7 of *Procedia Computer Science*, pages 68–72. Elsevier, 2011.

[29] Ulle Endriss. Social choice theory as a foundation for multiagent systems. In Jörg P. Müller, Michael Weyrich, and Ana L. C. Bazzan, editors, *Multiagent System Technologies - 12th German Conference, MATES 2014, Stuttgart, Germany, September 23-25, 2014. Proceedings*, volume 8732 of *Lecture Notes in Computer Science*, pages 1–6. Springer, 2014.

[30] Ulle Endriss. Judgment aggregation. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors, *Handbook of Computational Social Choice*, pages 399–426. Cambridge University Press, 2016.

[31] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Multiwinner voting: A new challenge for social choice theory. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 2, pages 27–47. AI Access, 2017.

[32] Luis Sánchez Fernández, Edith Elkind, Martin Lackner, Norberto Fernández García, Jesús Arias-Fisteus, Pablo Basanta-Val, and Piotr Skowron. Proportional justified representation. In Satinder Singh and Shaul Markovitch, editors, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*, pages 670–676. AAAI Press, 2017.

[33] Marc Fleurbaey and Maurice Salles. *A Brief History of Social Choice and Welfare Theory*, pages 1–16. Springer, 2021.

[34] Rupert Freeman, Seyed Majid Zahedi, and Vincent Conitzer. Fair and efficient social choice in dynamic settings. In Carles Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, pages 4580–4587. ijcai.org, 2017.

[35] David Gale and Lloyd Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

[36] Yingqiang Ge, Shuchang Liu, Ruoyuan Gao, Yikun Xian, Yunqi Li, Xiangyu Zhao, Changhua Pei, Fei Sun, Junfeng Ge, Wenwu Ou, and Yongfeng Zhang. Towards long-term fairness in recommendation. In Liane Lewin-Eytan, David Carmel, Elad Yom-Tov, Eugene Agichtein, and Evgeniy Gabrilovich, editors, *WSDM '21, The Fourteenth ACM International Conference on Web Search and Data Mining, Virtual Event, Israel, March 8-12, 2021*, pages 445–453. ACM, 2021.

[37] Georges-Théodule Guilbaud. Les théories de l'intérêt général et le problème logique de l'agrégation. *Économie Appliquée*, 5:501–584, 1952.

[38] Edith Hemaspaandra, Holger Spakowski, and Jörg Vogel. The complexity of Kemeny elections. *Theor. Comput. Sci.*, 349(3):382–391, 2005.

[39] Jonas Israel and Markus Brill. Dynamic proportional rankings. *Soc. Choice Welf.*, 64(1):221–261, 2025.

[40] John G. Kemeny. Mathematics without numbers. *Daedalus*, 88(4):577–591, 1959.

[41] Bettina Klaus, David F. Manlove, and Francesca Rossi. Matching under preferences. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors, *Handbook of Computational Social Choice*, pages 333–355. Cambridge University Press, 2016.

[42] Martin Lackner. Perpetual voting: Fairness in long-term decision making. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 2103–2110. AAAI Press, 2020.

[43] Martin Lackner and Jan Maly. Perpetual voting: The axiomatic lens. *CoRR*, abs/2104.15058, 2021.

[44] Martin Lackner and Jan Maly. Proportional decisions in perpetual voting. In Brian Williams, Yiling Chen, and Jennifer Neville, editors, *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, pages 5722–5729. AAAI Press, 2023.

[45] Jean-François Laslier and M. Remzi Sanver, editors. *Introduction to the Handbook on Approval Voting*. Springer, 2010.

[46] Patrick Lederer, Dominik Peters, and Tomasz Was. The squared Kemeny rule for averaging rankings. In Dirk Bergemann, Robert Kleinberg, and Daniela Sabán, editors, *Proceedings of the 25th ACM Conference on Economics and Computation, EC 2024, New Haven, CT, USA, July 8-11, 2024*, page 755. ACM, 2024. Full version available at `https://doi.org/10.48550/arXiv.2404.08474`.

[47] Judith Masthoff and Amra Delic. Group recommender systems: Beyond preference aggregation. In Francesco Ricci, Lior Rokach, and Bracha Shapira, editors, *Recommender Systems Handbook*, pages 381–420. Springer, 3rd edition, 2020.

[48] Christos H. Papadimitriou. *Computational complexity*. Addison-Wesley, 1994.

[49] David C. Parkes and Ariel D. Procaccia. Dynamic social choice with evolving preferences. In Marie desJardins and Michael L. Littman, editors, *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, July 14-18, 2013, Bellevue, Washington, USA*, pages 767–773. AAAI Press, 2013.

[50] David M. Pennock, Eric Horvitz, and C. Lee Giles. Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering. In Henry A. Kautz and Bruce W. Porter, editors, *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on on Innovative Applications of Artificial Intelligence, July 30 - August 3, 2000, Austin, Texas, USA*, pages 729–734. AAAI Press / The MIT Press, 2000.

[51] Dominik Peters, Grzegorz Pierczynski, and Piotr Skowron. Proportional participatory budgeting with additive utilities. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Advances*

in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual, pages 12726–12737, 2021.

[52] Gabriella Pigozzi, Alexis Tsoukiàs, and Paolo Viappiani. Preferences in artificial intelligence. *Ann. Math. Artif. Intell.*, 77(3-4):361–401, 2016.

[53] Francesco Ricci, Lior Rokach, and Bracha Shapira. Recommender systems: Techniques, applications, and challenges. In Francesco Ricci, Lior Rokach, and Bracha Shapira, editors, *Recommender Systems Handbook*. Springer, 3rd edition, 2020.

[54] Markus Schulze. Free riding and vote management under proportional representation by the single transferable vote. 2011.

[55] Piotr Skowron, Martin Lackner, Markus Brill, Dominik Peters, and Edith Elkind. Proportional rankings. In Carles Sierra, editor, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, pages 409–415. ijcai.org, 2017.

[56] Moshe Tennenholtz. Transitive voting. In *Proceedings of the 5th ACM Conference on Electronic Commerce (ACM-EC-2004)*, pages 230–231. ACM, 2004.

[57] William Thomson. Fair allocation rules. In Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, editors, *Handbook of Computational Social Choice and Welfare*, volume 2, pages 393–506. Elsevier, 2011.

[58] William Thomson. Introduction to the theory of fair allocation. In Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors, *Handbook of Computational Social Choice*, pages 333–355. Cambridge University Press, 2016.

[59] Thi Ngoc Trang Tran, Müslüm Atas, Alexander Felfernig, Viet Man Le, Ralph Samer, and Martin Stettinger. Towards social choice-based explanations in group recommender systems. In George Angelos Papadopoulos, George Samaras, Stephan Weibelzahl, Dietmar Jannach, and Olga C. Santos, editors, *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization, UMAP 2019, Larnaca, Cyprus, June 9-12, 2019*, pages 13–21. ACM, 2019.