# The Squared Kemeny Rule for Averaging Rankings

PATRICK LEDERER, Technical University of Munich, Germany
DOMINIK PETERS, CNRS, LAMSADE, Université Paris Dauphine - PSL, France
TOMASZ WĄS, CNRS, LAMSADE, Université Paris Dauphine - PSL, France

Manuscript: February 2024

For the problem of aggregating several rankings into one ranking, Kemeny [1959] proposed two methods: the *median rule* which selects the ranking with the smallest total swap distance to the input rankings, and the *mean rule* which minimizes the *squared* swap distances to the input rankings. The median rule has been extensively studied since and is now known simply as *Kemeny's rule*. It exhibits majoritarian properties, so for example if more than half of the input rankings are the same, then the output of the rule is the same ranking.

We observe that this behavior is undesirable in many rank aggregation settings. For example, when we rank objects by different criteria (quality, price, etc.) and want to aggregate them with specified weights for the criteria, then a criterion with weight 51% should have 51% influence on the output instead of 100%. We show that the Squared Kemeny rule (i.e., the mean rule) behaves this way, by establishing a bound on the distance of the output ranking to any input rankings, as a function of their weights. Furthermore, we give an axiomatic characterization of the Squared Kemeny rule, which mirrors the existing characterization of the Kemeny rule but replaces the majoritarian Condorcet axiom by a proportionality axiom. Finally, we discuss the computation of the rule and show its behavior in a simulation study.

Contents



Authors' addresses: Patrick Lederer, Technical University of Munich, Germany, ledererp@in.tum.de; Dominik Peters, CNRS, LAMSADE, Université Paris Dauphine - PSL, France, dominik.peters@lamsade.dauphine.fr; Tomasz Wąs, CNRS, LAMSADE, Université Paris Dauphine - PSL, France, tomasz.was@lamsade.dauphine.fr.

## 1 INTRODUCTION

Many search engines allow users to sort results by several criteria. For example, websites such as booking.com and expedia.com allow users to sort hotels by their price, their review score, or their distance to the city center. They also offer sorting by combinations of these criteria (for example "best reviewed and lowest price", see Figure 1). More generally, they could allow users to specify weights over the different ways to sort alternatives (e.g. 60% price, 30% reviews, 10% location) and then present an aggregated ranking of hotels.

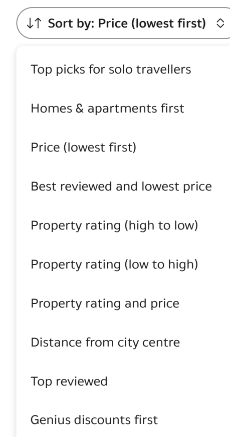


Fig. 1. Sort options on booking.com.

The task of combining several rankings (possibly with weights) into one ranking is known as *rank aggregation*. The best-known and most frequently discussed rule for this problem is *Kemeny's rule* [Kemeny, 1959, Kemeny and Snell, 1960], which minimizes the total distance to all input rankings. To be precise, let $A$ be a set of $m$ alternatives (e.g., hotels), and let $\mathcal{R}$ be the set of rankings (linear orders) on $A$. For two rankings $\succ, \rhd \in \mathcal{R}$, the *swap distance* (or Kendall-tau distance) between them is the number of pairs of alternatives on which they disagree: $\mathrm{swap}(\succ, \rhd) = |\{\{x, y\} \subseteq A : x \succ y \text{ and } y \rhd x\}|$. A *profile* $R$ is a function that assigns to each linear order $\succ$ a weight $R(\succ) \geqslant 0$, with weights summing to 1. To aggregate the rankings in a profile into a collective ranking, we use a *social preference function* (*SPF*) $f$, which selects for each profile $R$ a set $f(R) \subseteq \mathcal{R}$ of output rankings (ideally just one ranking, but there may be ties). Finally, Kemeny's rule is the SPF that selects the rankings which minimize the average swap distance to the input:

$$\mathrm{Kemeny}(R) = \arg\min_{\rhd \in \mathcal{R}} \textstyle\sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd).$$

Kemeny's rule is an attractive SPF for several reasons: it is the maximum likelihood estimator (MLE) of the Mallow's $\phi$ model [Young, 1988] which makes it a good fit for epistemic problems where we wish to discover a ground truth ranking given noisy estimates. Further, the rule is axiomatically characterized by a Condorcet-style axiom [Young and Levenglick, 1978]; in particular, when the weight of some ranking exceeds 50%, then the output of Kemeny's rule will be that ranking.

While this latter property is desirable in epistemic and electoral settings, for a hotel booking website it is disqualifying. In the above example of a user wishing to sort hotels by 60% price, 30% reviews, 10% location, the Kemeny output would just be the ranking by price, with review scores and location having no influence. Instead, what the user wants is a ranking where a hotel can compensate a lower position in the price ranking by a high position in the reviews and location rankings. Thus, we need a rule that faithfully follows the desired weighting, and makes use of all the information instead of ignoring low-weight criteria.

As it turns out, exactly such a rule was proposed under the name "mean rule" by Kemeny [1959] and Kemeny and Snell [1960] in the same articles that introduced Kemeny's rule. Our name for this SPF is the *Squared Kemeny rule* as it is obtained by squaring the distances in the objective function:

$$\mathrm{SqK}(R) = \arg\min_{\rhd \in \mathcal{R}} \textstyle\sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd)^2.$$

The Squared Kemeny rule appears to have been almost entirely ignored in the subsequent literature. Young and Levenglick [1978, p. 290] quickly dismiss it, writing that "Kemeny left the problem of which solution to choose unresolved. But from the standpoint of collective decision-making there is ample reason to prefer the median, since it turns out that the median consensus leads to a Condorcet method, while the mean does not." We believe that this dismissal was too quick.

The contribution of this paper is showing that the Squared Kemeny rule is well-suited to perform rank aggregation when we wish each input ranking to be reflected in the output ranking, to an

| Price | 90% | 80% | 70% | 60% | 50% | 40% | 30% | 20% | 10% | Score |
|---|---|---|---|---|---|---|---|---|---|---|
| La Quinta | Graduate | Graduate | Graduate | Graduate | Graduate | Graduate | Graduate | Graduate | Graduate | Graduate |
| Graduate | La Quinta | La Quinta | La Quinta | NHH | NHH | NHH | H. Marcel | H. Marcel | H. Marcel | The Study |
| NHH | NHH | NHH | NHH | La Quinta | H. Marcel | H. Marcel | NHH | The Study | The Study | H. Marcel |
| Omni | Omni | H. Marcel | H. Marcel | H. Marcel | La Quinta | The Study | The Study | NHH | NHH | NHH |
| H. Marcel | H. Marcel | Omni | The Study | The Study | The Study | La Quinta | La Quinta | La Quinta | Omni | Omni |
| The Study | The Study | The Study | Omni | Omni | Omni | Omni | Omni | Omni | La Quinta | La Quinta |

Fig. 2. Squared Kemeny when mixing two criteria.

extent that is proportional to its weight.[1] Such proportionality notions have recently attracted significant attention in voting, in particular in the settings of multi-winner voting and participatory budgeting [see, e.g., Aziz et al., 2017, Lackner and Skowron, 2023, Peters et al., 2021]. Following the approach in that literature, we could formalize proportionality, for example, by saying that a ranking that makes up $\alpha\%$ of the weight should agree with the output ranking on at least roughly $\alpha\%$ pairwise comparisons. While Squared Kemeny does not satisfy this in general, we will see that it does in important special cases, and that it satisfies an approximate version in general. More generally, we will show that Squared Kemeny behaves more like an average and thus is more responsive to changes in its input than the Kemeny rule, which behaves more like a median.

### Rank Aggregation With Two Criteria

To understand how the Squared Kemeny rule behaves, it is instructive to consider the problem of aggregating just two different rankings with different weights. Let us again consider a hotel booking example. In Figure 2, we show a ranking of 6 hotels offered on booking.com in New Haven, Connecticut, for the nights 8–11 July 2024 (accessed 5 February 2024). At the very left, the hotels are ranked by price, and at the very right, they are ranked by average user score. The figure shows the output of Squared Kemeny (with ties broken consistently) when the two rankings are given different weights; the top row shows the weight given to the price ranking.

We see that Squared Kemeny smoothly interpolates between the two rankings. Indeed, the price and score rankings differ on exactly 10 pairwise comparisons, and going through the rankings from left to right, we see that in each step one pairwise swap is performed. Thus, for example, when the price ranking has weight 70% (and the score ranking has weight 30%), the Squared Kemeny ranking agrees with the price ranking on 7 of the 10 disagreement comparisons, and it agrees with the score ranking on 3 of 10.

This is true in general. We formalize this by saying that Squared Kemeny satisfies *2-Rankings-Proportionality* (2RP). This axiom says that for every profile $R$ containing just two rankings $\succ_1, \succ_2$ with positive weight that disagree on $d = \mathrm{swap}(\succ_1, \succ_2)$ pairwise comparisons, we have

$$\rhd \in \mathrm{SqK}(R) \iff d - \mathrm{swap}(\succ_i, \rhd) \in \mathrm{round}(R(\succ_i) \cdot d) \text{ for both } i = 1 \text{ and } i = 2,$$

where $\mathrm{round}(z)$ is the set of one or two integers closest to $z \in \mathbb{R}$. Thus, for profiles with two input rankings, the Squared Kemeny rule chooses all "mean rankings" $\rhd$ where the number of pairwise agreements between $\rhd$ and the input rankings is proportional to their weights.

Note that the Kemeny rule behaves very differently on profiles with two rankings – it just outputs the input ranking with higher weight.

[1]Proportionality can be seen as a fairness notion with respect to voters (who have a guaranteed amount of influence on the output ranking). The rank aggregation literature has studied distinct notions of fairness for candidates that come labelled as belonging to protected groups [Chakraborty et al., 2022, Kuhlman and Rundensteiner, 2020, Wei et al., 2022]

Our main result is an axiomatic characterization of the Squared Kemeny rule that uses the same axioms as the famous characterization of the Kemeny rule by Young and Levenglick [1978], but replaces their Condorcet axiom by the 2RP axiom. We also impose a mild continuity axiom.

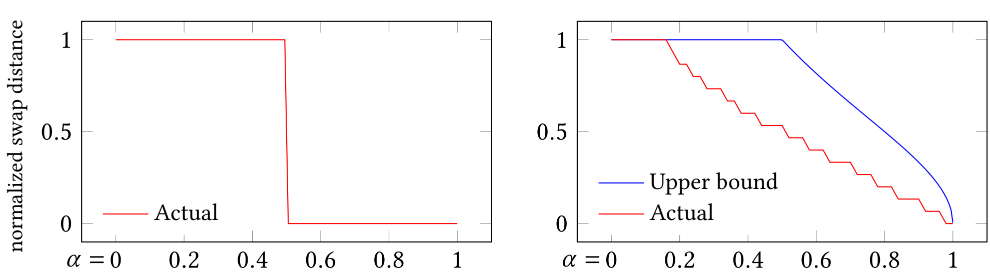
Theorem 3.2. *An SPF satisfies neutrality, reinforcement, continuity, and 2RP if and only if it is the Squared Kemeny rule.*

Neutrality is a standard symmetry condition. The reinforcement axiom is a consistency or convexity axiom, which says that if a ranking $\rhd$ is selected at two different profiles $R_1$ and $R_2$, i.e., $\rhd \in f(R_1) \cap f(R_2)$, then $\rhd$ is also selected for all convex combinations $\lambda R_1 + (1-\lambda)R_2$ of the two profiles ($\lambda \in (0,1)$), and that in addition $f(\lambda R_1 + (1-\lambda)R_2) = f(R_1) \cap f(R_2)$. This is a classic axiom that has been used in many characterizations in social choice [e.g., Fishburn, 1978, Lackner and Skowron, 2021, Myerson, 1995, Young, 1975].

Our proof operates within the space $\mathbb{Q}^{m!}$ of (generalized) profiles and uses reinforcement in a standard way to obtain separating hyperplanes between the regions of profiles where some particular output ranking is selected. However, unlike Young and Levenglick [1978], we cannot pass to a lower-dimensional space of majority margins. Instead, we characterize the hyperplanes by using 2RP to construct profiles where the rule chooses rankings that form a single-crossing path, which allows us to deduce that the hyperplanes encode the Squared Kemeny rule.

### General Proportionality Guarantee

The 2RP axiom applies only to profiles that contain two different rankings. Can we say anything similar about Squared Kemeny in the general case? Inspired by the literature on proportionality in multi-winner voting [Lackner and Skowron, 2023], we will consider groups of rankings and will bound the maximum distance of the output ranking to the group as a function of the group's size.




(a) Kemeny

(b) Squared Kemeny

Fig. 3. The maximum swap distance (normalized to $[0,1]$) between the output of the Kemeny or Squared Kemeny rule to an input ranking, as a function of the weight $\alpha$ of the input ranking, for $m = 6$ alternatives.

Our first question asks how large the swap distance can be between an input ranking $\succ$ with weight $\alpha \in [0,1]$ and the output ranking $\rhd$ of some rule $f$, as a function of $\alpha$. For Kemeny, this is easy to specify: for $\alpha < \frac{1}{2}$, Kemeny might output the reverse ranking of $\succ$ (if this ranking has weight more than $\frac{1}{2}$), and so the distance can be as large as $\binom{m}{2} = \frac{m(m-1)}{2}$ (the highest possible swap distance). When $\alpha > \frac{1}{2}$, the distance is guaranteed to be 0 by the Condorcet property of Kemeny.

For Squared Kemeny, we expect the bound to be smoother, giving some guaranteed influence to rankings with weights below $\frac{1}{2}$. This is indeed the case. We can compute the exact worst-case bound provided by Squared Kemeny for fixed small numbers $m$ of alternatives (see Figure 3), and we can see that it is approximately linear in $\alpha$, except for small $\alpha$. We can also prove a theoretical upper bound that holds for all $m$: the maximum distance of the Squared Kemeny output to an input ranking with weight $\alpha$ is at most $\sqrt{(1-\alpha)/\alpha} \cdot \binom{m}{2}$ (Theorem 4.1). This bound implies that Squared Kemeny will never output the reverse ranking of an input ranking that has weight more than $\frac{1}{2}$. For large $m$, we prove another bound that shows this even for rankings with weight more than $\frac{1}{4}$.

Our second question is about giving a similar type of guarantee not to a single ranking, but to a group of rankings. For example, there could be many similar rankings in the profile that each have a small weight, but which collectively have a significant weight. We want to show that the Squared Kemeny outcome cannot be too far away from those rankings, on average. Theorem 4.2 establishes a bound that applies to all groups of rankings with total weight $\alpha$ (whether these rankings are similar or not), guaranteeing that the Squared Kemeny output has an average distance of at most $\sqrt{1/(4\alpha)} \cdot \binom{m}{2} + o(m^{1.5})$ to the rankings in the group, where the lower-order term vanishes quickly.

### Empirical Analysis

We complement our theoretical analysis with results from simulations to better understand how the Squared Kemeny rule compares to the Kemeny rule. Since we need to compute the outcomes of the rules, we discuss their computational complexity in Section 5. In Section 6.1, we then perform a detailed analysis of an example of using the rank aggregation rules to rank cities according to a mixture of three criteria (GDP per capita, air quality, and sunniness).

Next, we analyze in Section 6.2 the rankings chosen by Kemeny and Squared Kemeny on random data. For example, we sample Euclidean profiles, where criteria and alternatives correspond to points in 2D space, and rankings are induced by sorting the alternatives by their distance to each criterion. In more detail, for the example in Figure 4, we sampled 100 profiles with 40 rankings each, where 75% of the input rankings come from a Gaussian in the lower left corner and the other 25% from a Gaussian in the upper right corner. We then computed the output rankings of the Kemeny rule (red diamonds) and of the Squared Kemeny rule (green squares), and embed these rankings in the same Euclidean space. We observe in Figure 4 that the Kemeny rule is located within the larger of the two voter clusters, while the Squared Kemeny rule interpolates between the two clusters.

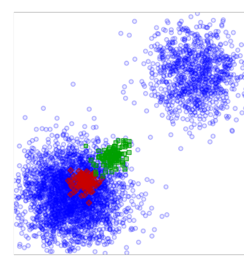


Fig. 4. Euclidean profiles with Kemeny (red) and Squared Kemeny (green) ranking locations shown

Finally, in Section 6.3 we revisit our quantitative worst-case proportionality bounds from an average-case perspective, and experimentally investigate the distance between an output ranking and a group of input rankings, as a function of the total group weight. The results confirm our theoretical predictions: the Kemeny rule ensures only that large groups are satisfied, while the Squared Kemeny rule caters to all group sizes.

### Applications of Proportional Rank Aggregation

Hotel booking websites are just one example where it makes sense to give users fine-grained control over how to sort items and where proportional rank aggregation methods such as Squared Kemeny are desirable. Further examples are lists of products in e-commerce (ranking by cost, rating, delivery time, etc.), newsfeeds of social networks, and database display applications in general.

There are also less technical applications, such as producing university rankings. These rankings are usually a result of aggregating rankings for several criteria (such as student satisfaction, % of students employed after graduating, research output). These rankings could be weighted and then be used to produce an aggregate ranking via Squared Kemeny as all criteria should be taken into account. Similarly, one could produce rankings of cities by livability or suitability for remote work.

In all previous examples, the input rankings are *criteria* (which tend to be objective) and the whole setup is essentially single-agent. However, there are also compelling multi-agent applications, where we can think of the input rankings as *votes*. An example might be a university hiring committee needing to rank applicants. In such a scenario, each committee member can provide their personal ranking. The output should be a ranking instead of a single winning candidate, because we do not know which candidates will accept the job offer. Proportionality may be desirable in this context to ensure that the output ranking reflects the diverse interests of the university department. Other multi-agent examples are groups of friends wanting to produce rankings of favorite music, restaurants, or travel destinations, a context in which a majoritarian method seems out of place.

In our example applications based on criteria, one may object that many criteria are numerical in nature (e.g., hotel price and average user rating). Using only the induced ranking throws away information, and taking a weighted average of the underlying numbers may produce a better result. The advantage of the rank aggregation approach is that it does not require the aggregator to decide on how to normalize the numerical values of different criteria. Normalization can be a difficult task without principled solutions – consider that hotel prices are in the hundreds while user ratings are between 0 and 10, and that for prices, lower is better, while for ratings, higher is better. Rank aggregation sidesteps these problems, and it is also more robust to outlier values.

## 2 THE MODEL

Let $A$ be a finite set of $m \geqslant 2$ alternatives. A *ranking* $\succ$ is a complete, transitive, and anti-symmetric binary relation on $A$. We denote by $\mathcal{R}$ the set of all rankings on $A$. In this paper, we study the problem of aggregating rankings into a collective ranking. To this end, we define a *(ranking) profile* $R$ to be a function from $\mathcal{R}$ to weights in $[0, 1] \cap \mathbb{Q}$ such that $\sum_{\succ \in \mathcal{R}} R(\succ) = 1$. More intuitively, a ranking profile specifies for each ranking a weight; these weights may, e.g., arise from a user assigning importance to different criteria (as in the hotel example in the introduction) or represent the fraction of voters who report a ranking in an election. The restriction that the weights are rational is to ensure compatibility with electorate settings (where discrete voters report preference relations) and does not affect our results. We denote the set of all rankings profiles $R$ by $\mathcal{R}^*$. For a profile $R$, we write $\text{supp}(R) = \{\succ \in \mathcal{R} : R(\succ) > 0\}$ for the set of rankings with positive weight.

Given a profile $R$, we want to derive an aggregate ranking. A *social preference function (SPF)* does this, being a function $f : \mathcal{R}^* \to 2^{\mathcal{R}} \setminus \{\emptyset\}$ that for every profile $R \in \mathcal{R}^*$ returns a non-empty set of chosen rankings $f(R) \subseteq \mathcal{R}$. To help distinguish between input and output rankings, we will follow a convention of denoting the input rankings by $\succ$ and the output rankings by $\rhd$. For two rankings $\succ, \rhd \in \mathcal{R}$, their *swap distance* is the number of pairs of alternatives which they order differently, i.e., $\text{swap}(\succ, \rhd) = |\{a, b \in A : a \succ b \text{ and } b \rhd a\}|$.

We focus on two SPFs in this paper. The *Kemeny* rule is defined as the set of rankings minimizing the (weighted) average swap distance between the input rankings and the output ranking, i.e., $\text{Kemeny}(R) = \arg\min_{\rhd \in \mathcal{R}} \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \text{swap}(\succ, \rhd)$. The *Squared Kemeny* rule is defined analogously, but with the swap distance squared, i.e., $\text{SqK}(R) = \arg\min_{\rhd \in \mathcal{R}} \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \text{swap}(\succ, \rhd)^2$. For a ranking $\rhd$, we let $C_{\text{SqK}}(R, \rhd) = \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \text{swap}(\succ, \rhd)^2$ be its *Squared Kemeny cost* in $R$.

## 3 AXIOMATIC ANALYSIS

We begin by analyzing the Squared Kemeny rule from an axiomatic perspective. First, we demonstrate in Section 3.1 that this rule indeed behaves like an average for special profiles (profiles where only two rankings have positive weight as well as single-crossing profiles). This means that the Squared Kemeny rule is proportional on these profiles, and we use this insight to characterize this SPF in Section 3.2. Finally, we also consider standard properties such as efficiency, participation, and strategyproofness, and check which of them are satisfied by the Squared Kemeny rule in Section 3.3.

### 3.1 2-Rankings-Proportionality and Single-Crossing Profiles

We start by showing that the Squared Kemeny rule is proportional in some special cases by defining two properties that formalize what it means to "proportionally" aggregate rankings on important classes of profiles. Our first property concerns cases where the rule must "average" two rankings with specified weights, just like in the introduction's hotel example. It requires that the output ranking must agree with the input rankings on a proportional number of pairwise comparisons.

*2-Rankings-Proportionality.* An SPF $f$ satisfies *2-Rankings-Proportionality* (2RP) if, for all profiles $R$ with $\text{supp}(R) = \{\succ_1, \succ_2\}$ for two rankings $\succ_1$ and $\succ_2$ with $d = \text{swap}(\succ_1, \succ_2)$, it holds that

$$f(R) = \{\rhd \in \mathcal{R} \colon d - \text{swap}(\succ_i, \rhd) \in \text{round}(R(\succ_i) \cdot d) \text{ for } i \in \{1, 2\}\},$$

or equivalently,

$$f(R) = \{\rhd \in \mathcal{R} \colon \text{ swap}(\succ_i, \rhd) \in \text{round}((1 - R(\succ_i)) \cdot d) \text{ for } i \in \{1, 2\}\},$$

where $\text{round}(z)$ denotes the set of closest integers to $z$.[2] Less formally, this means that the higher the weight of $\succ_1$ (resp. $\succ_2$), the closer the output rankings are to $\succ_1$ (resp. $\succ_2$).

For example, suppose that $\text{swap}(\succ_1, \succ_2) = 10$, so the two rankings in $R$ disagree on 10 pairwise comparisons and that $R(\succ_1) = 30\%$. Then the output ranking $\rhd$ should agree with $\succ_1$ on $30\% \cdot 10 = 3$ of those disagreement pairs, and disagree with $\succ_1$ on $(1 - 30\%) \cdot 10 = 7$ of the disagreement pairs.

The Squared Kemeny rule satisfies 2RP, and in fact it satisfies a stronger property about *single-crossing* profiles. A sequence of rankings $\succ_0, \dots, \succ_n \in \mathcal{R}$ is called *single-crossing* if for every pair of alternative $a, b \in A$ with $a \succ_0 b$,

$$\text{there exists } i \in \{0, \dots, n\} \text{ such that } a \succ_0 b, \dots, a \succ_i b \text{ and } b \succ_{i+1} a, \dots, b \succ_n a.$$

Thus, scanning the rankings from left to right, the relative positions of every pair of alternatives cross at most once. We say that a sequence is *maximal single-crossing* if every pair of alternatives crosses exactly once (which implies that $n = \binom{m}{2}$ and that $\succ_0$ and $\succ_n$ are reverse rankings). As an example, the rankings shown in Figure 5 form a maximal single-crossing sequence.

| 30% | 0% | 30% | 0% | 10% | 0% | 0% | 0% | 0% | 0% | 30% |
|---|---|---|---|---|---|---|---|---|---|---|
| $\succ_0$ | $\succ_1$ | $\succ_2$ | $\succ_3$ | $\succ_4$ | $\succ_5$ | $\succ_6$ | $\succ_7$ | $\succ_8$ | $\succ_9$ | $\succ_{10}$ |
| $a$ | $b$ | $b$ | $b$ | $b$ | $c$ | $c$ | $c$ | $d$ | $d$ | $e$ |
| $b$ | $a$ | $c$ | $c$ | $c$ | $b$ | $d$ | $d$ | $c$ | $e$ | $d$ |
| $c$ | $c$ | $a$ | $d$ | $d$ | $d$ | $b$ | $e$ | $e$ | $c$ | $c$ |
| $d$ | $d$ | $d$ | $a$ | $e$ | $e$ | $e$ | $b$ | $b$ | $b$ | $b$ |
| $e$ | $e$ | $e$ | $e$ | $a$ | $a$ | $a$ | $a$ | $a$ | $a$ | $a$ |

Fig. 5. A single-crossing profile, on which Squared Kemeny outputs $\succ_4$ (the average of the voter locations) and Kemeny outputs $\succ_2$ (the median).

We say that a *profile* $R$ is *single-crossing* if the rankings in $\text{supp}(R)$ can be arranged in a single-crossing sequence. On single-crossing profiles, there is a natural definition of what it means to be an average, because each input ranking is associated with a location $0, \dots, n$ in a one-dimensional space; thus the output ranking should be at the weighted average of these locations. For example, the profile in Figure 5 has 30% of the voters each in locations 0, 2, and 10, with the remaining 10% in location 4. This gives an average location of $0.3 \cdot (0 + 2 + 10) + 0.1 \cdot 4 = 4$. Thus, the "average" ranking for this profile is $\succ_4$, and indeed this is the output ranking of Squared Kemeny. In contrast, the Kemeny rule takes the median location, which is location 2, and so $\succ_2$ is the Kemeny output.[3]

To formalize this behavior, let us say that $R$ is *compatible* with a maximal single-crossing sequence $\succ_0, \dots, \succ_n$ if $\text{supp}(R) \subseteq \{\succ_0, \dots, \succ_n\}$. We can now state an axiom specifying what it means to proportionally aggregate rankings on a single-crossing profile.

[2] For $k \in \mathbb{Z}$, $\text{round}(k + x) = \{k\}$ if $x \in [0, 0.5)$, $\text{round}(k + 0.5) = \{k, k + 1\}$, and $\text{round}(k + x) = \{k + 1\}$ if $x \in (0.5, 1]$.

[3] This is known as the *representative voter theorem* [Rothstein, 1991] which states that the majority relation of a single-crossing profile (and hence the Kemeny ranking) coincides with the input ranking of the median voter.

*Single-Crossing-Proportionality.* An SPF $f$ satisfies *Single-Crossing-Proportionality* if for every single-crossing profile $R$ we have that a ranking $\rhd$ is in $f(R)$ if and only if there exists a maximal single-crossing sequence $\succ_0, \dots, \succ_n$ compatible with $R$ and $\rhd = \succ_i$ for $i \in \mathrm{round}(\sum_{j=1}^{n} R(\succ_j) \cdot j)$.

It is straightforward to check that Single-Crossing-Proportionality implies 2RP, because any profile in which only 2 rankings occur is single-crossing.

Theorem 3.1. *The Squared Kemeny rule satisfies Single-Crossing-Proportionality and 2RP.*

Proof. Since Single-Crossing-Proportionality implies 2RP, it is sufficient to prove that Squared Kemeny satisfies the former.

For this, we fix a single-crossing profile $R$ and an arbitrary maximal single-crossing sequence $\succ_0, \dots, \succ_n$ compatible with $R$. Also, let $\rhd \in \mathcal{R}$ be an arbitrary ranking and define $d = \mathrm{swap}(\succ_0, \rhd)$. Because of the triangle inequality, we can now compute that

$$\begin{aligned} \mathrm{swap}(\succ_i, \rhd) &\geqslant \mathrm{swap}(\succ_0, \rhd) - \mathrm{swap}(\succ_0, \succ_i) = d - i, \quad \text{for } i \in [d], \text{ and} \\ \mathrm{swap}(\succ_i, \rhd) &\geqslant \mathrm{swap}(\succ_0, \succ_i) - \mathrm{swap}(\succ_0, \rhd) = i - d, \quad \text{for } i \in [n] \setminus [d]. \end{aligned} \tag{1}$$

We will first show that if $\rhd$ does not belong to any maximally single-crossing sequence compatible with $R$, then at least one of Inequalities (1) for $i \in [n]$ such that $\succ_i \in \mathrm{supp}(R)$ is strict. Assume otherwise, i.e., $\mathrm{swap}(\succ_i, \rhd) = |d - i|$ for every $i \in [n]$ such that $\succ_i \in \mathrm{supp}(R)$. If $R(\succ_d) > 0$, this implies that $\rhd = \succ_d$, so $\rhd$ belongs to the maximal single-crossing sequence $\succ_0, \dots, \succ_n$. Thus, let us assume that $R(\succ_d) = 0$. Then, let $i \in \{0, \dots, d-1\}$ be maximal and $j \in \{d+1, \dots, n\}$ minimal such that $\succ_i, \succ_j \in \mathrm{supp}(R)$. By adding the respective equalities sidewise, we obtain that

$$\mathrm{swap}(\succ_i, \rhd) + \mathrm{swap}(\succ_j, \rhd) = (d - i) + (j - d) = j - i = \mathrm{swap}(\succ_i, \succ_j).$$

By [Elkind et al., 2022, Proposition 4.6], this implies that $\succ_i$, $\rhd$, and $\succ_j$ form a single-crossing sequence and $\rhd$ is between $\succ_i$ and $\succ_j$. This means that for every pair of alternatives $x, y \in A$ such that $x \rhd y$, it holds that $y \succ_i x$ implies $x \succ_j y$ and $y \succ_j x$ implies $x \succ_i y$. Consequently, the sequence $\succ_0, \dots, \succ_i, \rhd, \succ_j, \dots, \succ_n$ is single-crossing, which contradicts the assumption that $\rhd$ does not belong to any maximally single-crossing sequence compatible with $R$.

Therefore, if $\rhd$ does not belong to any maximally single-crossing sequence compatible with $R$, we can show that there will be a ranking in such set, i.e., $\succ_d$, for which $C_{\mathrm{SqK}}(R, \rhd) > C_{\mathrm{SqK}}(R, \succ_d)$. Indeed, from Inequalities (1) and the fact that one of them is strict, we get that

$$C_{\mathrm{SqK}}(R, \rhd) = \textstyle\sum_{i=0}^{n} R(\succ_i) \cdot \mathrm{swap}(\succ_i, \rhd)^2 > \sum_{i=0}^{n} R(\succ_i) \cdot (i - d)^2 = C_{\mathrm{SqK}}(R, \succ_d).$$

Thus, $\rhd \notin \mathrm{SqK}(R)$ and we know that the only rankings selected by the Squared Kemeny rule belong to some maximally single-crossing sequence compatible with $R$.

Then, take an arbitrary $\rhd \in \mathrm{SqK}(R)$ and again denote $d = \mathrm{swap}(\succ_0, \rhd)$. This means that $C_{\mathrm{SqK}}(R, \rhd) = \sum_{i=0}^{n} R(\succ_i) \cdot (i - d)^2$. Taking the derivative with respect to $d$, we get that $C_{\mathrm{SqK}}(R, \rhd)$ is minimized when $\sum_{i=0}^{n} 2R(\succ_i) \cdot (d - i) = 0$, which is equivalent to $\sum_{i=0}^{n} R(\succ_i) \cdot i = d$. Since $d$ has to be an integer and a quadratic function grows symmetrically from its minimum (note that a convex combination of quadratic functions is still a quadratic function), the rankings $\succ_d$ for $d \in \mathrm{round}(\sum_{i=0}^{n} R(\succ_i) \cdot i)$ have the lowest Squared Kemeny cost among the rankings in $\succ_0, \dots, \succ_n$.

It remains to show that for $\succ_d$ selected in this way, the cost $C_{\mathrm{SqK}}(R, \succ_d)$ is the same no matter which maximally single-crossing sequence we have chosen at the beginning. To this end, take two arbitrary maximally single-crossing sequences $\succ_0, \dots, \succ_n$ and $\succ'_0, \dots, \succ'_n$, both compatible with $R$. Observe that there is a linear function $\ell(x) = ax + b$, with $a \in \{1, -1\}$ such that $\succ_i = \succ'_{\ell(i)}$ for every $\succ_i \in \mathrm{supp}(R)$. Then, by the linearity of the mean, if $\succ_d$ minimizes $C_{\mathrm{SqK}}(R, \succ_j)$ among $\succ_0, \dots, \succ_n$, then $\succ'_{\ell(d)}$ minimizes $C_{\mathrm{SqK}}(R, \succ'_j)$ among $\succ'_0, \dots, \succ'_n$. Furthermore, we have that $C_{\mathrm{SqK}}(R, \succ_d) = \sum_{i=0}^{n} R(\succ_i) \cdot (i - d)^2 = \sum_{i=0}^{n} R(\succ_i) \cdot (\ell(i) - \ell(d))^2 = C_{\mathrm{SqK}}(R, \succ_{\ell(d)})$, which concludes the proof. □

### 3.2 Characterization of the Squared Kemeny Rule

We will next present our characterization of the Squared Kemeny rule, which combines 2RP with three standard properties, namely neutrality, reinforcement, and continuity.

*Neutrality.* Neutrality is a mild symmetry condition that precludes a rule from depending on the names of candidates. An SPF $f$ is *neutral* if $f(\tau(R)) = \{\tau(\rhd) : \rhd \in f(R)\}$ for all profiles $R \in \mathcal{R}^*$ and permutations $\tau : A \to A$. Here, we denote by $\rhd' = \tau(\rhd)$ the ranking defined by $\tau(x) \rhd' \tau(y)$ if and only if $x \rhd y$ for all $x, y \in A$ and $R' = \tau(R)$ is the profile defined by $R'(\tau(\succ)) = R(\succ)$ for all $\succ \in \mathcal{R}$.

*Reinforcement.* Reinforcement is a classic axiom in social choice theory which describes that if some outcomes are chosen for two profiles, then precisely these common outcomes should be chosen in a convex combination of these profiles. An SPF $f$ satisfies *reinforcement* if for all profiles $R, R' \in \mathcal{R}^*$ with $f(R) \cap f(R') \neq \emptyset$, we have $f(\lambda R + (1 - \lambda)R') = f(R) \cap f(R')$ for all $\lambda \in (0, 1) \cap \mathbb{Q}$.

*Continuity.* Continuity requires that a group of rankings with sufficient weight can overrule any other set of rankings and thus determine the outcome. Formally, an SPF $f$ is *continuous* if for all profiles $R, R' \in \mathcal{R}^*$, there is a scalar $\lambda \in (0, 1) \cap \mathbb{Q}$ such that $f(\lambda R + (1 - \lambda)R') \subseteq f(R)$.

The above three axioms that have been frequently used in social choice theory to characterize scoring rules in various contexts [e.g., Lackner and Skowron, 2021, Lederer, 2023, Myerson, 1995, Skowron et al., 2019, Young, 1975]. Moreover, Kemeny's rule has been characterized as the unique SPF satisfying neutrality, reinforcement, and a Condorcet axiom [Young and Levenglick, 1978].

We can now state our characterization result.

THEOREM 3.2. *An SPF satisfies neutrality, reinforcement, continuity, and 2RP if and only if it is the Squared Kemeny rule.*

Like other reinforcement-based characterizations of SPFs, our proof is quite involved, and thus we defer it to Appendix A. Using well-known techniques, the proof uses reinforcement to divide the space of profiles into convex regions where a particular ranking is chosen by the SPF, and then uses 2RP in combination with the other axioms to characterize the boundaries (hyperplanes) of these regions. As for the independence of the axioms, we do not know if neutrality or continuity can be dropped from the characterization. Without 2RP, the Kemeny rule satisfies the remaining axioms. Without reinforcement, the rule that agrees with the Squared Kemeny rule on all profiles in which two rankings jointly have more than 90% of the weight, and returns the set of all rankings $\mathcal{R}$ for all other profiles, satisfies the remaining axioms.

To give more insights into the proof of Theorem 3.2, we will show a weaker statement that is still of interest. To this end, we introduce the family of *ranking scoring functions*. These are analogues of scoring rules in voting and are defined based on a *cost function* $c\colon \mathcal{R} \times \mathcal{R} \to \mathbb{R}$ that assigns to each pair of rankings $\succ, \rhd \in \mathcal{R}$ a cost $c(\succ, \rhd)$. Intuitively, we interpret the term $c(\succ, \rhd)$ as the disutility that the outcome ranking $\rhd$ would give to a voter with a preference order $\succ$. The ranking scoring function $f_c$ based on $c$ returns the rankings with minimal total cost: $f_c(R) = \arg\min_{\rhd \in \mathcal{R}} \sum_{\succ \in \mathcal{R}} R(\succ) \cdot c(\succ, \rhd)$ for every profile $R$. For example, $f_{\text{swap}}$ is the Kemeny rule and $f_{\text{swap}^2}$ is the Squared Kemeny rule.

It is straightforward to check that every ranking scoring function satisfies reinforcement and continuity. The class of (neutral) ranking scoring functions was introduced by Conitzer et al. [2009], who conjectured that this class is in fact characterized by neutrality, reinforcement, and continuity.

We now give a proof that the only ranking scoring function that satisfies the 2RP axiom is the Squared Kemeny rule. The proof uses a very similar strategy to the proof of our full axiomatic characterization (Theorem 3.2) but avoids some of its overhead. Note that this version of the characterization does not require neutrality.

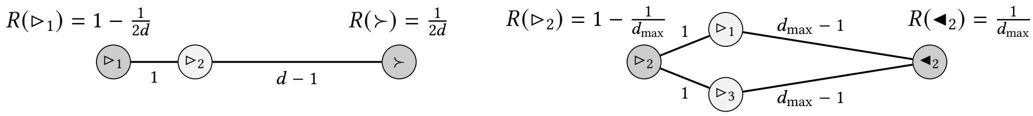



Fig. 6. An illustration of Steps 1 (left) and 2 (right) of the proof of Theorem 3.2.

Theorem 3.3. *A ranking scoring function satisfies 2RP if and only if it is the Squared Kemeny rule.*

Proof. We already showed that the Squared Kemeny rule satisfies 2RP (Theorem 3.1), so we need to show that it is the only ranking scoring function with that property. Hence, let $f$ denote a ranking scoring function that satisfies 2RP and let $c$ be its cost function. Without loss of generality, we can assume for every ranking $\succ \in \mathcal{R}$ that $c(\succ, \succ) = 0$ because adding a constant to a cost function, even a different constant for each first argument, does not affect the outcomes of a ranking scoring function. For a profile $R$ and a ranking $\rhd$, we define $C(R, \rhd) = \sum_{\succ \in \mathcal{R}} R(\succ) \cdot c(\succ, \rhd)$. We will show that $c$ is proportional to the Squared Kemeny cost function $c_{\text{SqK}}(\succ, \rhd) = \text{swap}(\succ, \rhd)^2$, which implies that $f$ is the Squared Kemeny rule. We prove this in three steps: first, we will show that for every pair of rankings $\rhd_1, \rhd_2 \in \mathcal{R}$ with swap distance 1, the difference in the costs of $\rhd_1$ and $\rhd_2$ with respect to any other ranking is proportional to the difference in their Squared Kemeny costs, i.e., there exists a constant $\alpha_{\rhd_1,\rhd_2} > 0$ such that $c(\succ, \rhd_1) - c(\succ, \rhd_2) = \alpha_{\rhd_1,\rhd_2}(c_{\text{SqK}}(\succ, \rhd_1) - c_{\text{SqK}}(\succ, \rhd_2))$ for every $\succ \in \mathcal{R}$ (Step 1). Next, we will prove that these constants are equal if two such pairs of rankings intersect, i.e., $\alpha_{\rhd_1,\rhd_2} = \alpha_{\rhd_2,\rhd_3}$ for all $\rhd_1, \rhd_2, \rhd_3 \in \mathcal{R}$ such that $\text{swap}(\rhd_1, \rhd_2) = \text{swap}(\rhd_2, \rhd_3) = 1$ (Step 2). Finally, we infer that all these constants are equal and derive that there is $\alpha > 0$ such that $c(\succ, \rhd_1) - c(\succ, \rhd_2) = \alpha(c_{\text{SqK}}(\succ, \rhd_1) - c_{\text{SqK}}(\succ, \rhd_2))$ for all $\succ, \rhd_1, \rhd_2 \in \mathcal{R}$. Since $c(\rhd, \rhd) = c_{\text{SqK}}(\rhd, \rhd)$ for all $\rhd \in \mathcal{R}$, this means that $c$ is indeed proportional to $c_{\text{SqK}}$ (Step 3).

**Step 1:** Fix two rankings $\rhd_1, \rhd_2 \in \mathcal{R}$ with $\text{swap}(\rhd_1, \rhd_2) = 1$ and define $\alpha_{\rhd_1,\rhd_2} = c(\rhd_1, \rhd_2)$. We will show that $\alpha_{\rhd_1,\rhd_2} > 0$ and $c(\succ, \rhd_1) - c(\succ, \rhd_2) = \alpha_{\rhd_1,\rhd_2}(c_{\text{SqK}}(\succ, \rhd_1) - c_{\text{SqK}}(\succ, \rhd_2))$ for all $\succ \in \mathcal{R}$. Let us first prove that $\alpha_{\rhd_1,\rhd_2}$ is equal to $\alpha_{\rhd_2,\rhd_1} = c(\rhd_2, \rhd_1)$. For this, we consider the profile $R$ with $R(\rhd_1) = R(\rhd_2) = 1/2$. 2RP requires that $f(R) = \{\rhd_1, \rhd_2\}$. Thus, from the definition of ranking scoring functions and our assumption that $c(\succ, \succ) = 0$ for every $\succ \in \mathcal{R}$ we derive that $c(\rhd_1, \rhd_2) = c(\rhd_2, \rhd_1)$ and therefore

$$\alpha_{\rhd_1,\rhd_2} = \alpha_{\rhd_2,\rhd_1}. \tag{2}$$

To show that $\alpha_{\rhd_1,\rhd_2} > 0$, consider another profile $R'$ with $R'(\rhd_1) = 2/3$ and $R'(\rhd_2) = 1/3$. By 2RP, we get that $f(R') = \{\rhd_1\}$ and thus that $2/3 \cdot c(\rhd_1, \rhd_1) + 1/3 \cdot c(\rhd_2, \rhd_1) < 2/3 \cdot c(\rhd_1, \rhd_2) + 1/3 \cdot c(\rhd_2, \rhd_2)$. Because $c(\rhd_1, \rhd_2) = c(\rhd_2, \rhd_1)$ and $c(\rhd_1, \rhd_1) = c(\rhd_2, \rhd_2) = 0$, we derive that $\alpha_{\rhd_1,\rhd_2} > 0$.

Now, fix an arbitrary ranking $\succ$. Since $\text{swap}(\rhd_1, \rhd_2) = 1$, there is exactly one pair of alternatives on which $\rhd_1$ and $\rhd_2$ disagree. Let us denote them by $a$ and $b$, i.e., $a \rhd_1 b$ and $b \rhd_2 a$. We subsequently assume that $b \succ a$ and discuss the case that $a \succ b$ later. Let $d = \text{swap}(\succ, \rhd_1)$ and observe that $\text{swap}(\succ, \rhd_2) = d - 1$ (see Figure 6 for an illustration). Moreover, we define $R$ as the profile with $R(\succ) = 1/2d$ and $R(\rhd_1) = 1 - 1/2d$. Since $1/2d \cdot \text{swap}(\succ, \rhd_1) = 1/2$ and $(1 - 1/2d) \cdot \text{swap}(\succ, \rhd_1) = d - 1/2$, 2RP implies that $\rhd_1, \rhd_2 \in f(R)$. Thus, by the definition of ranking scoring functions, $C(R, \rhd_1) = C(R, \rhd_2)$, which further implies that $0 = 2d \cdot C(R, \rhd_1) - 2d \cdot C(R, \rhd_2)$. We can now compute that

$$\begin{aligned} 0 &= c(\succ, \rhd_1) + (2d-1)c(\rhd_1, \rhd_1) - c(\succ, \rhd_2) - (2d-1)c(\rhd_1, \rhd_2) \\ &= c(\succ, \rhd_1) - c(\succ, \rhd_2) - (2d-1)\alpha_{\rhd_1,\rhd_2}. \end{aligned}$$

Since $c_{\text{SqK}}(\succ, \rhd_1) - c_{\text{SqK}}(\succ, \rhd_2) = \text{swap}(\succ, \rhd_1)^2 - \text{swap}(\succ, \rhd_2)^2 = d^2 - (d-1)^2 = 2d - 1$, it follows that $c(\succ, \rhd_1) - c(\succ, \rhd_2) = \alpha_{\rhd_1,\rhd_2}(c_{\text{SqK}}(\succ, \rhd_1) - c_{\text{SqK}}(\succ, \rhd_2))$.

Lastly, let us consider the case of $a \succ b$. By analogous reasoning, we obtain that $c(\succ, \rhd_2) - c(\succ, \rhd_1) = \alpha_{\rhd_2,\rhd_1}(c_{\text{SqK}}(\succ, \rhd_2) - c_{\text{SqK}}(\succ, \rhd_1))$. By Equation (2) and sidewise multiplication by $-1$, we hence infer the thesis of this Step also in this case.

**Step 2:** Consider three rankings $\rhd_1, \rhd_2, \rhd_3 \in \mathcal{R}$ with $\text{swap}(\rhd_1, \rhd_2) = \text{swap}(\rhd_2, \rhd_3) = 1$ and let $\alpha_{\rhd_1,\rhd_2}$ and $\alpha_{\rhd_2,\rhd_3}$ denote the constants derived in the previous step. We will show in this step that $\alpha_{\rhd_1,\rhd_2} = \alpha_{\rhd_2,\rhd_3}$. If $m = 2$, then necessarily $\rhd_1 = \rhd_3$ and our claim directly follows from Equation (2).

Thus, assume $m \geqslant 3$, let $\blacktriangleleft_2$ denote the ranking that is completely reverse of $\rhd_2$, and let $d_{\max} = \binom{m}{2} = \text{swap}(\blacktriangleleft_2, \rhd_2)$. Furthermore, we define $R$ as the profile with $R(\blacktriangleleft_2) = 1/d_{\max}$ and $R(\rhd_2) = 1 - 1/d_{\max}$ (see Figure 6 for an illustration). Observe that $1/d_{\max} \cdot \text{swap}(\blacktriangleleft_2, \rhd_2) = 1$ and $(1 - 1/d_{\max}) \cdot \text{swap}(\blacktriangleleft_2, \rhd_2) = d_{\max} - 1$. Hence, 2RP implies that all rankings in swap distance 1 from $\rhd_2$ are selected by $f$. In particular, $\rhd_1, \rhd_3 \in f(R)$. By the definition of the ranking scoring function, this means that $C(R, \rhd_1) = C(R, \rhd_3)$, and therefore also $C(R, \rhd_1) - C(R, \rhd_2) = C(R, \rhd_3) - C(R, \rhd_2)$. Next, Step 1 implies that $C(R, \rhd_1) - C(R, \rhd_2) = \alpha_{\rhd_1,\rhd_2}(C_{\text{SqK}}(R, \rhd_1) - C_{\text{SqK}}(R, \rhd_2))$ and $C(R, \rhd_3) - C(R, \rhd_2) = \alpha_{\rhd_3,\rhd_2}(C_{\text{SqK}}(R, \rhd_3) - C_{\text{SqK}}(R, \rhd_2))$. Since we have $C_{\text{SqK}}(R, \rhd_1) = C_{\text{SqK}}(R, \rhd_1) = 1/d_{\max} \cdot (d_{\max} - 1)^2 + (1 - 1/d_{\max}) \cdot 1^2 < 1/d_{\max} \cdot d_{\max}^2 = C_{\text{SqK}}(R, \rhd_2)$, we now derive that

$$\alpha_{\rhd_1,\rhd_2} = \frac{C(R, \rhd_1) - C(R, \rhd_2)}{C_{\text{SqK}}(R, \rhd_1) - C_{\text{SqK}}(R, \rhd_2)} = \frac{C(R, \rhd_3) - C(R, \rhd_2)}{C_{\text{SqK}}(R, \rhd_3) - C_{\text{SqK}}(R, \rhd_2)} = \alpha_{\rhd_3,\rhd_2}.$$

Hence, by Equation (2), it follows that $\alpha_{\rhd_1,\rhd_2} = \alpha_{\rhd_2,\rhd_3}$.

**Step 3:** Finally, we will show that $c$ is proportional to $c_{\text{SqK}}$, so that $f$ is the Squared Kemeny rule. To this end, we first show that $\alpha_{\rhd,\rhd'} = \alpha_{\rhd^+,\rhd^*}$, for all $\rhd, \rhd', \rhd^+, \rhd^* \in \mathcal{R}$ with $\text{swap}(\rhd, \rhd') = \text{swap}(\rhd^+, \rhd^*) = 1$. To see this, observe that there exists a sequence of rankings $\rhd_1, \rhd_2, \dots, \rhd_k$ such that $\rhd_1 = \rhd$, $\rhd_2 = \rhd'$, $\rhd_{k-1} = \rhd^+$, $\rhd_k = \rhd^*$, and $\text{swap}(\rhd_i, \rhd_{i+1}) = 1$ for every $i \in \{1, \dots, k-1\}$. Then, Step 2 implies that $\alpha_{\rhd,\rhd'} = \alpha_{\rhd_1,\rhd_2} = \dots = \alpha_{\rhd_{k-1},\rhd_k} = \alpha_{\rhd^+,\rhd^*}$. We hence drop the index of these constants and simply refer to them by $\alpha$.

Next, take arbitrary rankings $\succ, \rhd, \rhd' \in \mathcal{R}$ and let $\rhd_1, \dots, \rhd_k$ be a sequence of rankings such that $\rhd_1 = \rhd$, $\rhd_k = \rhd'$, and $\text{swap}(\rhd_i, \rhd_{i+1}) = 1$ for every $i \in \{1, \dots, k-1\}$. Using the "telescoping sum" technique twice, we get that $c(\succ, \rhd) - c(\succ, \rhd') = \sum_{i=1}^{k-1}(c(\succ, \rhd_i) - c(\succ, \rhd_{i+1})) = \alpha \sum_{i=1}^{k-1}\left(c_{\text{SqK}}(\succ, \rhd_i) - c_{\text{SqK}}(\succ, \rhd_{i+1})\right) = \alpha(c_{\text{SqK}}(\succ, \rhd) - c_{\text{SqK}}(\succ, \rhd'))$.

From this we get that $c$ is proportional to $c_{\text{SqK}}$. Indeed, for every two rankings $\succ, \rhd \in \mathcal{R}$, we get that $c(\succ, \rhd) = c(\succ, \rhd) - c(\succ, \succ) = \alpha(c_{\text{SqK}}(\succ, \rhd) - c_{\text{SqK}}(\succ, \succ)) = \alpha(c_{\text{SqK}}(\succ, \rhd)$. Since $\alpha > 0$, this means that $f$ is the Squared Kemeny rule. □

### 3.3 Efficiency, Participation, and Strategyproofness

We will next show that the Squared Kemeny rule satisfies desirable efficiency and participation properties but violates strategyproofness. To define these axioms for SPFs, we first need to specify how we compare two output rankings $\rhd_1, \rhd_2$ based on an input ranking $\succ$. Following the literature [e.g., Athanasoglou, 2016, Bossert and Sprumont, 2014, Bossert and Storcken, 1992], we use the swap distance between the input ranking and the output rankings: given an input ranking $\succ$, $\rhd_1$ is *weakly preferred* to $\rhd_2$ (denoted by $\rhd_1 \succsim \rhd_2$) if $\text{swap}(\succ, \rhd_1) \leqslant \text{swap}(\succ, \rhd_2)$, and $\rhd_1$ is strictly preferred to $\rhd_2$ (denoted by $\rhd_1 \succ \rhd_2$) if $\text{swap}(\succ, \rhd_1) < \text{swap}(\succ, \rhd_2)$. It does not make a difference for these purposes whether we use the swap distance or the squared swap distance since they induce the same preferences. Next, we will define efficiency, participation, and strategyproofness.

*Efficiency.* An outcome is (Pareto) efficient if it is not possible to make one voter better off without making any other voter worse off. Formally, we say that a ranking $\rhd_1$ dominates another ranking $\rhd_2$ in a profile $R$ if $\rhd_1 \succsim \rhd_2$ for all $\succ \in \text{supp}(R)$ and $\rhd_1 \succ \rhd_2$ for some $\succ \in \text{supp}(R)$. Moreover, a ranking $\rhd$ is efficient for a profile $R$ if it is not dominated by any other ranking. Finally, an SPF $f$ is *efficient* if, for every profile $R$, every ranking $\rhd \in f(R)$ is efficient.[4]

*Participation.* The axiom of participation is typically formulated in electoral settings and intuitively requires that it is never better for a group of agents to abstain from an election than to participate. In our context, participation can be seen as a consistency notion: if $\rhd_1$ is a winning ranking in the profile $R$ and we add additional criteria to $R$ according to which $\rhd_1$ is better than $\rhd_2$, then $\rhd_2$ should not be winning in the extended profile. More formally, we say an SPF $f$ satisfies *participation* if there are no profiles $R^1, R^2$, a constant $\lambda \in (0,1) \cap \mathbb{Q}$, and rankings $\rhd_1 \in f(R^1)$, $\rhd_2 \in f(\lambda R^1 + (1-\lambda R^2))$ such that $\rhd_1 \succsim \rhd_2$ for all $\succ \in \operatorname{supp}(R^2)$ and $\rhd_1 \succ \rhd_2$ for some $\succ \in \operatorname{supp}(R^2)$.

*Strategyproofness.* As the last axiom of this section, we will consider strategyproofness. This axiom is also typically studied in electoral settings and requires that agents should never be better off by lying than by voting truthfully. We hence say an SPF $f$ is *strategyproof* if there are no profiles $R^1$, $R^2$, rankings $\succ_1 \in \operatorname{supp}(R^1)$, $\succ_2 \in \mathcal{R}$, and a constant $\varepsilon \in (0, R^1(\succ_1)]$ such that *(i)* $f(R^1) = \{\rhd_1\}$ and $f(R^2) = \{\rhd_2\}$ for some rankings $\rhd_1, \rhd_2$ with $\rhd_2 \succ_1 \rhd_1$, and *(ii)* $R^2(\succ_1) = R^1(\succ_1) - \varepsilon$, $R^2(\succ_2) = R^1(\succ_2) + \varepsilon$, and $R^2(\succ) = R^1(\succ)$ for all $\succ \in \mathcal{R} \setminus \{\succ_1, \succ_2\}$.

Proposition 3.4. *The Squared Kemeny rule satisfies efficiency and participation but violates strategyproofness.*

Proof. We first show that Squared Kemeny is efficient. For this, let $R$ be a profile and $\rhd_1$, $\rhd_2$ be two rankings such that $\rhd_1$ dominates $\rhd_2$ in $R$. It is easy to check that $C_{\mathrm{SqK}}(R, \rhd_1) < C_{\mathrm{SqK}}(R, \rhd_2)$, so $\rhd_2 \notin \mathrm{SqK}(R)$, which implies that the Squared Kemeny rule is indeed efficient.

For participation, consider two profiles $R^1$, $R^2$, and two rankings $\rhd_1$, $\rhd_2$ such that $\rhd_1 \in \mathrm{SqK}(R^1)$, $\rhd_1 \succsim \rhd_2$ for all $\succ \in \operatorname{supp}(R^2)$, and $\rhd_1 \succ \rhd_2$ for some $\succ \in \operatorname{supp}(R^2)$. Since $\rhd_1 \in \mathrm{SqK}(R^1)$, we have $C_{\mathrm{SqK}}(R^1, \rhd_1) \leqslant C_{\mathrm{SqK}}(R^1, \rhd_2)$. Moreover, from the conditions on $R^2$ we get that $C_{\mathrm{SqK}}(R^2, \rhd_1) < C_{\mathrm{SqK}}(R^1, \rhd_2)$. This implies for every $\lambda \in (0,1) \cap \mathbb{Q}$ that $\rhd_2 \notin \mathrm{SqK}(\lambda R^1 + (1-\lambda)R^2)$ because

$$\begin{aligned} C_{\mathrm{SqK}}(\lambda R^1 + (1-\lambda)R^2, \rhd_1) &= \lambda C_{\mathrm{SqK}}(R^1, \rhd_1) + (1-\lambda) C_{\mathrm{SqK}}(R^2, \rhd_1) \\ &> \lambda C_{\mathrm{SqK}}(R^1, \rhd_2) + (1-\lambda) C_{\mathrm{SqK}}(R^2, \rhd_2) = C_{\mathrm{SqK}}(\lambda R^1 + (1-\lambda)R^2, \rhd_2). \end{aligned}$$

Hence, the Squared Kemeny rule indeed satisfies participation.

Finally, we turn to strategyproofness, and consider the following two profiles $R^1$ and $R^2$.

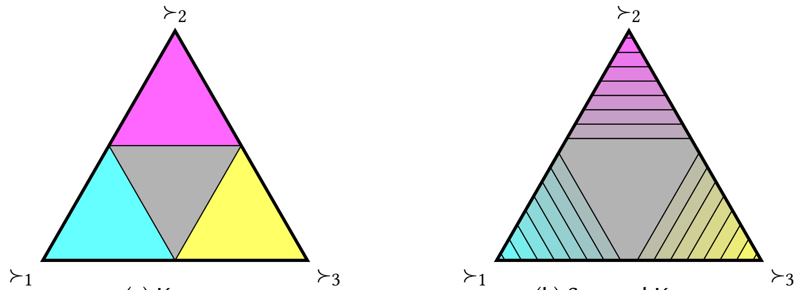
$$R^1: \begin{array}{ccc} 1/3 & 5/9 & 1/9 \\ \hline a & b & c \\ b & a & a \\ c & c & b \end{array} \qquad\qquad R^2: \begin{array}{cccc} 1/3 & 4/9 & 1/9 & 1/9 \\ \hline a & b & c & c \\ b & a & a & b \\ c & c & b & a \end{array}$$

It can be verified that Squared Kemeny uniquely chooses the ranking $\rhd_1 = a \rhd_1 b \rhd_1 c$ for $R^1$ and the ranking $\rhd_2 = b \rhd_2 a \rhd_2 c$ for $R^2$. Since the profile $R^2$ arises from the profile $R^1$ by moving probability $1/9$ from $\succ = b \succ a \succ c$ to $c \succ' b \succ' a$ and $\rhd_2 \succ \rhd_1$, this shows that the Squared Kemeny rule is manipulable. This example also shows that Squared Kemeny fails the weaker "betweenness" version of strategyproofness of Bossert and Sprumont [2014], which Kemeny does satisfy. □

## 4 PROPORTIONALITY GUARANTEES

In our axiomatic treatment, we have discussed the behavior of the Squared Kemeny rule on profiles with a lot of structure (single-crossing) and in particular those with only two rankings (2RP). Does Squared Kemeny retain its behavior as an average in general? This is what we will quantify in

[4] Bossert and Sprumont [2014] show that if $a \succ b$ for all $\succ \in \operatorname{supp}(R)$, then we also have $a \rhd b$ whenever $\rhd$ is efficient in $R$.




(a) Kemeny (b) Squared Kemeny

Fig. 7. The simplex of profiles in which the rankings $\succ_1 = abcdefgh$, $\succ_2 = fedcbahg$, and $\succ_3 = bahgfdec$ occur. Each point of the simplex is colored according to the swap distance of the (a) Kemeny and (b) Squared Kemeny ranking to the input rankings, with a point's cyan (resp., magenta, yellow) component being more intense if the output ranking is closer to $\succ_1$ (resp., $\succ_2$, $\succ_3$). Both rules output $\rhd^* = bafedchg$ (which is equidistant to the three input rankings) when no input ranking has weight greater than $\frac{1}{2}$. Otherwise, Kemeny outputs the majority ranking, while Squared Kemeny smoothly moves towards an input ranking as its weight increases.

this section. As an initial matter, we can check what Squared Kemeny does on profiles in which *three* rankings occur. For a fixed set of three rankings, we can use a simplex of weights to picture all profiles based on these rankings. For each point of the simplex (i.e., for each profile), we can compute the Kemeny and Squared Kemeny outcomes, and we can color the point to indicate how close the output rankings are to each of the input rankings. We show the result of this exercise in Figure 7. This confirms our expectation that Squared Kemeny takes rankings with smaller weight into account, while Kemeny frequently ignores them.

For general profiles with any number of rankings, we can ask about the maximum (over all possible profiles) swap distance between the output ranking and an input ranking, as a function of its weight $\alpha \in [0, 1]$. For an ideal proportional rule, there should be a roughly linear relationship between these. For Kemeny, the distance can be as large as $\binom{m}{2}$ (the maximum possible swap distance) when $\alpha < \frac{1}{2}$, and it is 0 when $\alpha > \frac{1}{2}$. For Squared Kemeny, we can compute its behavior for fixed $m$ using a linear program that searches for the worst-case profile, which yields the plot in Figure 3 shown in the introduction. That function is approximately linear, except for a "hump" for small $\alpha$, which indicates that Squared Kemeny can output the reverse of an input ranking which has weight as large as $\alpha \approx 17\%$. We do not have a satisfactory explanation for these humps (the profiles witnessing this worst-case behavior are very complicated), and we do not know how big the hump is as $m \to \infty$, but we do know it cannot exceed $\alpha = 25\%$ (by Theorem 4.2 below). In addition to computing the exact worst-case behavior for fixed $m$, we can also prove a theoretical upper bound that works for any $m$, which bounds the distance between the Squared Kemeny ranking and a weight-$\alpha$ input ranking. This bound is also shown in Figure 3.

Theorem 4.1. *Let $R$ be a profile and let $\succ^* \in \mathcal{R}$ be a ranking with weight $R(\succ^*) = \alpha$. Then*

$$\mathrm{swap}(\succ^*, \rhd) \leqslant \sqrt{\frac{1-\alpha}{\alpha}} \cdot \binom{m}{2} \quad \text{for every } \rhd \in \mathrm{SqK}(R).$$

Proof. Note that $C_{\mathrm{SqK}}(R, \succ^*) \leqslant (1-\alpha) \cdot \binom{m}{2}^2$ since $\binom{m}{2}$ is the maximum swap distance between two rankings, and an $\alpha$ fraction of the profile has swap distance 0 to $\succ^*$. Let $\rhd \in \mathrm{SqK}(R)$ be a ranking selected by Squared Kemeny and write $d = \mathrm{swap}(\succ^*, \rhd)$. Then we have $C_{\mathrm{SqK}}(R, \rhd) \geqslant \alpha \cdot d^2$. Because $\rhd$ optimizes the Squared Kemeny cost, we have $C_{\mathrm{SqK}}(R, \rhd) \leqslant C_{\mathrm{SqK}}(R, \succ^*)$ and thus $\alpha \cdot d^2 \leqslant (1-\alpha)\binom{m}{2}^2$. Solving for $d$, we get $d \leqslant \sqrt{(1-\alpha)/\alpha} \cdot \binom{m}{2}$, as required. $\square$

The above bound makes sense for profiles with few rankings that are not very similar to each other. But in contexts with many rankings, some of which are similar to each other, it would be better to guarantee to *groups* of rankings that the output ranking should not be too far away from them, on average. We will formalize this in a similar way to the work of Skowron and Górecki [2022], by considering arbitrary groups of rankings, without making any cohesiveness assumptions (that would say that the rankings in a group must be similar to each other). Note that such a setup limits the guarantees we can give: for a group of size $\alpha = 1$ (i.e. all the rankings together), we cannot guarantee that the output agrees with the group on more than half the pairwise comparisons on average (consider for example a profile where one ranking and its reverse each have weight $\frac{1}{2}$).

To state our result, given a profile $R$, we say that $S : \mathcal{R} \to [0, 1] \cap \mathbb{Q}$ is a *subprofile* of $R$ if $S(\succ) \leqslant R(\succ)$ for all rankings $\succ \in \mathcal{R}$. The *size* of $S$ is $\sum_{\succ \in \mathcal{R}} S(\succ)$. We can think of $S$ as a *group* of voters, and its size as the fraction of the entire electorate that they make up. We now provide a bound on the average satisfaction of any group.

THEOREM 4.2. *Let $R$ be a profile and let $S$ be a subprofile of $R$ with size $\alpha$. Then*

$$\frac{1}{\alpha} \sum_{\succ \in \mathcal{R}} S(\succ) \cdot \mathrm{swap}(\succ, \rhd) \leqslant \sqrt{\frac{1}{4\alpha}} \cdot \binom{m}{2} + o(m^{1.5}) \quad \textit{for every } \rhd \in \mathrm{SqK}(R).$$

PROOF. Let $d_{\max} = \binom{m}{2}$ be the maximum swap distance between two rankings. Fix any ranking $\succ \in \mathcal{R}$. For each $i \in \{0, 1, \dots, d_{\max}\}$, let $M_i$ be the number of rankings $\rhd \in \mathcal{R}$ with $\mathrm{swap}(\succ, \rhd) = i$. The values $M_0, M_1, \dots, M_{d_{\max}}$ are the *Mahonian numbers* and it is known [Ben-Naim, 2010] that

$$\sum_{i=0}^{d_{\max}} M_i \cdot i^2 = m! \left( \frac{1}{4} d_{\max}^2 + \frac{2m^3 + 3m^2 - 5m}{72} \right).$$

For a fixed profile $R$, the average Squared Kemeny cost $C_{\mathrm{SqK}}(R, \rhd)$ over all rankings $\rhd$ is

$$\frac{1}{m!} \sum_{\rhd \in \mathcal{R}} \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd)^2 = \frac{1}{m!} \sum_{\succ \in \mathcal{R}} \sum_{\rhd \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd)^2 = \frac{1}{m!} \cdot m! \left( \frac{1}{4} d_{\max}^2 + O(m^3) \right).$$

Therefore, there exists a ranking $\rhd \in \mathcal{R}$ such that $C_{\mathrm{SqK}}(R, \rhd) \leqslant d_{\max}^2/4 + O(m^3)$. From minimality of Squared Kemeny, we know that this is true for any $\rhd \in \mathrm{SqK}(R)$. Since $S$ is a subprofile of $R$,

$$\frac{1}{\alpha} \sum_{\succ \in \mathcal{R}} S(\succ) \cdot \mathrm{swap}(\succ, \rhd)^2 \leqslant \frac{1}{\alpha} \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd)^2 \leqslant \frac{1}{\alpha} d_{\max}^2/4 + O(m^3).$$

Finally, by Jensen's inequality we get that $(\sum_{\succ \in \mathcal{R}} S(\succ) \cdot \mathrm{swap}(\succ, \rhd)/\alpha)^2 \leqslant d_{\max}^2/(4\alpha) + O(m^3)$, since the square function is convex, which yields the thesis. □

In Figure 8, we show the upper bound obtained in Theorem 4.2 in terms of normalized swap distance (so $\binom{m}{2}$ is mapped to 1). For $m = 6$, we also show the actual worst-case performance of Squared Kemeny, which can be computed for fixed $m$ using linear programs for finding worst-case profiles that maximize the distance between the output and some size-$\alpha$ group. The figure also shows a lower bound, which is obtained by linear programs that find profiles where all $m!$ output rankings are bad simultaneously in the sense that some size-$\alpha$ group incurs at least the lower bound's distance.

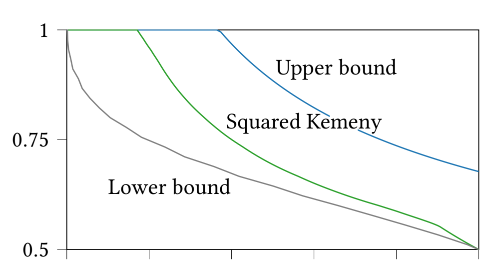



Fig. 8. Bound of Theorem 4.2 for groups of rankings.

## 5 COMPUTATION

The computational complexity of the Kemeny rule has been extensively studied. The problem of deciding if there is a ranking with at most a given cost is NP-complete [Bartholdi, III et al., 1989], even for a constant number of input rankings [Bachmeier et al., 2019, Biedl et al., 2009, Dwork et al., 2001]. Thus, it is reasonable to expect that the analogous problem for the Squared Kemeny rule is also NP-complete, and this is indeed the case, see Appendix B.1.

Theorem 5.1. *The problem of deciding, given a profile $R$ and a number $B$, whether there exists a ranking $\rhd$ with $C_{\mathrm{SqK}}(R, \rhd) \leqslant B$, is NP-complete, even for profiles with 4 rankings with equal weight.*

The proof is by reduction from the problem for the Kemeny rule, and uses the same technique used by Biedl et al. [2009] for showing that the egalitarian Kemeny rule (which selects the ranking where the maximum swap distance to any input ranking is minimized) is NP-complete to compute.

There exists an ILP formulation for computing the Kemeny rule, which is reasonably efficient in practice [Conitzer et al., 2006]. While this ILP formulation depends on the linear nature of the Kemeny objective, it is still possible to give an ILP formulation for the Squared Kemeny rule, using the same trick used by Caragiannis et al. [2019] for computing the maximum Nash welfare solution for fair allocation, and generalized by Bredereck et al. [2020] for various covering problems. The encoding is described in Appendix B.2. We found that it allows us to evaluate the Squared Kemeny rule reasonably efficiently up to $m = 80$ (see Figure 9).[5]

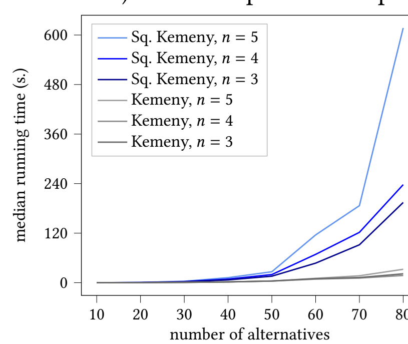



Fig. 9. The median running time of computing the Squared Kemeny using Gurobi for a given number of alternatives and $n = \{3, 4, 5\}$ rankings occurring in the profile with equal weights, drawn uniformly at random. The values are based on 50 samples.

The Kemeny rule also admits efficient approximation algorithms [Ailon et al., 2008, Coppersmith et al., 2010, Van Zuylen and Williamson, 2009] and even a PTAS [Kenyon-Mathieu and Schudy, 2007]. The Squared Kemeny rule admits a simple 4-approximation algorithm (output the input ranking that has the best score, see Appendix B.3). In addition, we can show that the optimal Kemeny ranking provides a 2-approximation to the Squared Kemeny rule. Combining this with the known PTAS for Kemeny, we obtain the following result, proved in Appendix B.4:

Theorem 5.2. *For every constant $\varepsilon > 0$, there exists a polynomial-time $(2 + \varepsilon)$-approximation to the Squared Kemeny rule.*

We believe, however, that such approximation algorithms have limited interest for our applications, since a ranking may have a good approximation factor to the optimum Squared Kemeny score while not satisfying the desirable proportionality properties of the Squared Kemeny rule. Indeed, the rankings returned by Kemeny and Squared Kemeny may be very far apart from each other (in the extreme case, they may be opposite to each other except for 1 shared pairwise comparison, see Appendix B.5), even though Kemeny provides a 2-approximation to Squared Kemeny. Still, approximation algorithms may have their use, for example, as subroutines in branch and bound algorithms. We leave the question of whether Squared Kemeny admits a PTAS to future work.

## 6 EMPIRICAL ANALYSIS

In this section, we compare the performance of the Squared Kemeny rule and the Kemeny rule based on several empirical experiments.

[5]We note that outliers (with up to 9 times longer running time than the median) do occur.

| Rank | GDP pc. (40%) | Air Quality (30%) | Sunniness (30%) | Kemeny | Sq. Kemeny | |
|---|---|---|---|---|---|---|
| 1. | San Francisco | New York City | Dubai | San Francisco | San Francisco | (–) |
| 2. | New York City | Dublin | Cairo | New York City | New York City | (–) |
| 3. | Zurich | Toronto | Johannesburg | Zurich | Sydney | (△2) |
| 4. | Dublin | Buenos Aires | San Francisco | Dublin | Toronto | (△2) |
| 5. | Sydney | London | Lahore | Sydney | Dubai | (△6) |
| 6. | Toronto | Sydney | Rome | Toronto | Rome | (△6) |
| 7. | London | Rio de Janeiro | Sydney | London | Johannesburg | (△6) |
| 8. | Paris | San Francisco | Mumbai | Tokyo | Buenos Aires | (△6) |
| 9. | Hong Kong | Zurich | New York City | Paris | Rio de Janeiro | (△6) |
| 10. | Tokyo | Tokyo | Mexico | Hong Kong | Zurich | (▽7) |
| 11. | Dubai | Rome | Buenos Aires | Dubai | Dublin | (▽7) |
| 12. | Rome | Moscow | Bangkok | Rome | London | (▽5) |
| 13. | Seoul | Paris | Rio de Janeiro | Johannesburg | Tokyo | (▽5) |
| 14. | Moscow | Hong Kong | Istanbul | Buenos Aires | Hong Kong | (▽4) |
| 15. | Shanghai | Mexico | Seoul | Rio de Janeiro | Mexico | (△3) |
| 16. | Johannesburg | Johannesburg | Toronto | Seoul | Seoul | (–) |
| 17. | Rio de Janeiro | Seoul | Tokyo | Moscow | Moscow | (–) |
| 18. | Istanbul | Bangkok | Shanghai | Mexico | Paris | (▽9) |
| 19. | Buenos Aires | Istanbul | Lagos | Bangkok | Bangkok | (–) |
| 20. | Mexico | Lagos | Hong Kong | Istanbul | Istanbul | (–) |
| 21. | Bangkok | Shanghai | Moscow | Shanghai | Shanghai | (–) |
| 22. | Mumbai | Cairo | Paris | Cairo | Cairo | (–) |
| 23. | Cairo | Dubai | Zurich | Mumbai | Mumbai | (–) |
| 24. | Lagos | Mumbai | London | Lagos | Lagos | (–) |
| 25. | Lahore | Lahore | Dublin | Lahore | Lahore | (–) |

Table 1. The three input rankings of cities and the results of their aggregation using the Kemeny and Squared Kemeny rule. In the rightmost column we also report how many positions a city has moved in the Squared Kemeny ranking in comparison to the Kemeny ranking.

## 6.1 Aggregate City Ranking

Our first experiment is a detailed example. We selected 25 cities around the world and ranked them, based on their Gross Domestic Product (GDP) per capita, air quality (measured by average PM 2.5 concentration), and sunniness (based on the average number of sunshine hours in a year). The details of the sources used can be found in Appendix C.1. Then, we assigned weight 40% to GDP per capita, 30% to air quality, and 30% to sunniness. We aggregated the rankings using the Kemeny and Squared Kemeny rules. While Squared Kemeny selected a unique ranking, Kemeny selected 4 tied outputs, in which a few cities differ by 1 to 3 positions. When increasing the weight on the GDP ranking by a small amount, only one of these outputs stays optimal, and we went with that output. The aggregation results are presented in Table 1.

Note that the top-5 of the Kemeny ranking is identical to the top-5 of the GDP per capita ranking. This includes Dublin and Zurich, which are both in the bottom-5 of the sunniness ranking. This is exactly the behavior we would like to avoid in the proportional aggregation of rankings: focusing mainly on one input ranking and disregarding another with still significant weight assigned to it.

In contrast, both Dublin and Zurich do not appear in the top-5 of the ranking selected by the Squared Kemeny rule. Instead, its top-5 includes Toronto (which is in the middle of the sunniness ranking and relatively high in both GDP per capita and air quality rankings) and Dubai (which, although near the bottom of the air quality ranking, is the first according to sunniness and 11-th based on GDP per capita). In this way, the top-5 of the Squared Kemeny rule output ranking arguably offers more uniform representation of the highest ranked cities across all input rankings.

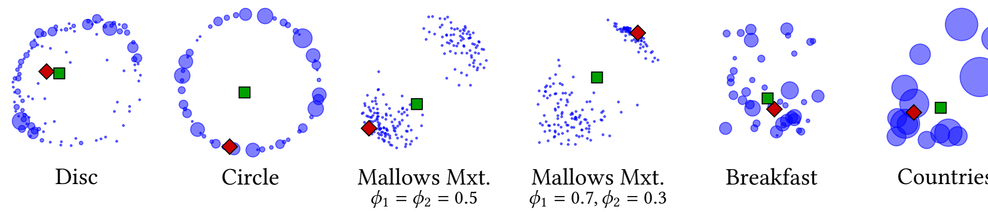



Fig. 10. Maps of preferences. For the Kemeny rule the position of its output is denoted with a red diamond, and for Squared Kemeny with a green square.

### 6.2 Drawing Embeddings of Rankings

Next, we visualize how the rankings output by the Kemeny and Squared Kemeny rules relate to the input rankings, using two methods of embedding rankings into 2-dimensional Euclidean space.

The first method is called *map of preferences*, introduced by Faliszewski et al. [2023], and starts by computing the swap distances between each pair of rankings present in a profile. Then, we apply a classical multidimensional scaling algorithm [Torgerson, 1952] to put each ranking as a dot on a plane in such a way that the Euclidean distances between the dots reflect the swap distances between the rankings as well as possible. By the size of a dot we signify the weight of a given ranking in the profile. In order to obtain the coordinates for the outputs of the Kemeny and Squared Kemeny rules as well, we simply include them while computing the distance matrix.

Figure 10 presents maps of preferences of six profiles which we have generated using different models. For each, we sampled 200 rankings, with possible repetitions, over 10 alternatives and constructed the profile by assigning each ranking a weight proportional to the number of times it was sampled. We have also verified that the behavior of Kemeny and Squared Kemeny visible on these examples is consistent across different profiles generated in the same way.

The first two pictures present profiles drawn from the *Euclidean model*, in which we sample 10 *alternative points* and 200 *voter points* uniformly at random from the unit disc (the first picture) or from its boundary, the unit circle (second picture). Then, for every voter point we record a ranking of all alternatives in order of increasing distance from the voter. For the disc, the Kemeny and Squared Kemeny rules select rankings very close to each other. However, for the circle, the difference is significant. While Squared Kemeny chooses a ranking that is in the center, Kemeny outputs a ranking that is similar to one of the input rankings on a circle. This is because for every ranking on the circle, we also have the reversed (or close to the reversed) ranking on the opposite side of the circle. Thus, all possible rankings have a similar average swap distance to the profile and the smallest occurs on the part of the circle from which, by chance, we sampled more voter points, which is then chosen by Kemeny. In contrast, since Squared Kemeny minimizes the average squared swap distance, it tries to equalize the distances to all rankings present in the profile.

The next two pictures show the maps of preferences for profiles generated from the mixture of two *Mallows models* [Mallows, 1957]. Given a central ranking $\succ$ and a noise parameter $\phi \in [0, 1]$, we sample each ranking $\succ'$ with probability proportional to $\phi^{\text{swap}(\succ,\succ')}$ under the Mallows model. Thus, for smaller $\phi$, the distribution is more concentrated around $\succ$. We generated 55% of the rankings using one Mallows model with the central ranking $\succ_1$ and noise $\phi_1$ and 45% using another Mallows model with the central ranking $\succ_2$ that is the complete reverse of $\succ_1$ and noise $\phi_2$. When $\phi_1 = \phi_2 = 0.5$ (the third picture), the Kemeny rule outputs $\succ_1$, the central ranking of the Mallows model responsible for 55% of the rankings while the Squared Kemeny rule selects a ranking in between $\succ_1$ and $\succ_2$. Interestingly, if $\phi_2$ is significantly smaller than $\phi_1$ (the fourth picture), the Kemeny rule outputs the central ranking of the smaller but more concentrated model, while Squared

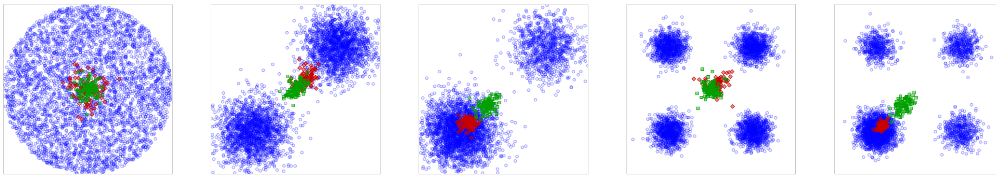


Fig. 11. Euclidean embeddings. For the Kemeny rule, the positions of its outputs are denoted with a red diamond, and for Squared Kemeny with a green square.

Kemeny is still between $\succ_1$ and $\succ_2$. This confirms our intuition that Squared Kemeny works like an average of rankings.

Finally, the last two pictures present the profiles drawn from models based on real-world data from Preflib [Mattei and Walsh, 2013]: the breakfast dataset [Green and Rao, 1972] that contains 42 preference orders over 15 breakfast items; and the 2016 countries ranking dataset [Boehmer and Schaar, 2023], where 107 countries are ranked according to 14 different criteria. For each dataset, we sampled with replacement 200 rankings and then restricted each ranking to 10 randomly chosen alternatives. For both profiles, Kemeny and Squared Kemeny choose similar rankings, with the former closer to the most concentrated part, and the latter closer to the center of the picture.

The second method of visualizing the positions of rankings is specific to profiles drawn from the Euclidean model. Given a profile specified by voter and alternative locations, and given an output ranking $\rhd$, we try to find a point in the same Euclidean space that would induce the ranking $\rhd$. In general, such a point may not exist, but using an ILP we can find a point that induces a ranking with minimal possible swap distance to $\rhd$.

In our experiments shown in Figure 11, we sample $m = 10$ candidate locations uniformly from the unit square, and $n = 40$ voter locations according to different distributions of interest. We then compute the outputs of the Kemeny and Squared Kemeny rules and embed them as a point in space. For each voter location distribution, we repeat this process 100 times and superimpose the results for the 100 profiles in the same figure, showing voters as blue dots, Kemeny rankings as red diamonds, and Squared Kemeny rankings as green squares.

The first voter distribution samples the voter locations uniformly from the unit disc, and we see that both rules select central rankings, with somewhat more variance in the Kemeny rankings. For the second and third picture, we sample voter locations from two Gaussians centered in the bottom left and top right corners. In the second picture, we sample 20 voters each from the two Gaussians and see that both rules select rankings roughly midway between the center points of the Gaussians. In the third picture, we sample 30 voters from the bottom left and 10 voters from the top right. Kemeny selects rankings located within the bigger cluster, while Squared Kemeny chooses locations that interpolate between the two (while still being closer to the larger cluster). The fourth and fifth picture repeat the same process with four Gaussians with voters uniformly distributed (in the fourth picture) or with 25 voters in the bottom left and 5 voters each in the other three clusters.

### 6.3 Worst-Case Average Distance

Our final experiment revisits the problem studied in Section 4, where we considered the average distance between the rankings of a group of voters and the output ranking. There, we bounded the distance for the worst-case profile, while here we compute it for randomly sampled profiles. Consider a profile $R$ and an output ranking $\rhd$. For each size $\alpha \in [0, 1]$, we look for the "unhappiest" group $S$ (i.e., subprofile of $R$) of size $\alpha$, in the sense that the average distance between $\rhd$ and the rankings of $S$ is large. Formally, we define $\mu_\alpha(R, \rhd) = \max_{S \subseteq R : |S| = \alpha} \frac{1}{\alpha} \sum_{\succ \in \mathcal{R}} S(\succ) \cdot \mathrm{swap}(\succ, \rhd)$ for this worst average distance, where $S \subseteq R$ denote that $S$ is a subprofile of $R$ with size $|S| = \alpha$.

For the experiment, we sampled 100 profiles with 8 alternatives and 50 rankings according to various distributions. The results of our experiment for the Euclidean disc model is presented in Figure 12. (See Appendix C.2 for other distributions.) For each profile and each size $\alpha \in \{1/50, 2/50, \dots, 1\}$, we put a red dot indicating the value of $\mu_\alpha(R, \rhd)$ for Kemeny, and a green dot for Squared Kemeny. The lines show the average value for each $\alpha$.

Figure 12 also shows a lower bound. For each $\alpha$, this is computed by finding the ranking $\rhd$ that optimizes $\mu_\alpha(R, \rhd)$, and placing a gray dot at that value. Note that different rankings may be optimum for different $\alpha$, and so this lower bound is "unfair" to rules like Kemeny or Squared Kemeny, which must choose a single ranking which gets evaluated for all $\alpha$ simultaneously.

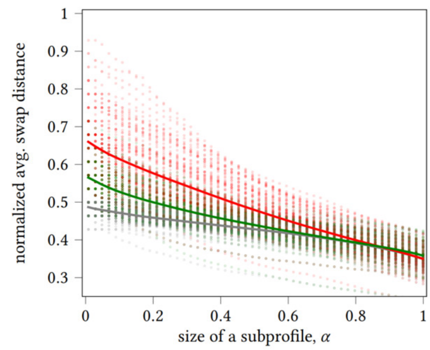



Fig. 12. The maximal average distances between the subprofile of size $\alpha$ and the output of the Kemeny (red) and Squared Kemeny (green) rules, plus lower bound (gray).

We see that when $\alpha$ is close to 1, Kemeny leads to smaller distances than Squared Kemeny. This is to be expected, as Kemeny minimizes the overall average distance. However, for $\alpha$ smaller than 0.8, it is Squared Kemeny that returns the lower values on average. Observe also that the difference between the Kemeny and Squared Kemeny for $\alpha$ close to 1 is negligible, while the differences for $\alpha$ close to 0 are significant or even substantial depending on the considered model.

## 7 CONCLUSIONS AND FUTURE WORK

We have studied the Squared Kemeny rule and argued that it behaves more appropriately in contexts where we want to aggregate rankings proportionally, compared to its better-known cousin the Kemeny rule. In particular, we have shown a full characterization of the Squared Kemeny rule based on a proportionality axiom, proved general proportionality guarantees for this rule, and demonstrated in an experimental study that it behaves similar to a mean. Based on these results, we conclude that the Squared Kemeny rule has the potential of providing a consensus ranking in situations where a majoritarian rule such as Kemeny is undesirable.

There are many interesting directions for future work exploring the topic of proportional rank aggregation. In particular, one could study new SPFs with the aim to find more proportional rules. For instance, one could consider rules based on the Spearman footrule distance instead of the swap distance [Diaconis and Graham, 1977, Viappiani, 2015], analogues of Proportional Approval Voting [Aziz et al., 2017], or the family of "$p$-Kemeny rules" that minimize the $p$-th power of the swap distance. One could also derive other proportionality axioms that are not defined in terms of swap distance. For example, following Skowron et al. [2017] who study proportional rankings based on approval votes, one could phrase proportionality as requiring that every top-initial segment of the output ranking, viewed as a set, should be a proportional committee. How to adapt this to ranking input is not clear, though, since it is an open question whether axioms for proportional multi-winner rules (such as Proportionality for Solid Coalitions, PSC, Aziz and Lee, 2020) are compatible with *committee monotonicity*, which is necessary to adapt a multi-winner rule to output a ranking.

Finally, we note that the methods that we have introduced may prove useful in other contexts. For example, we considered bounds on the maximum dissatisfaction of a voter, as a function of the voter's weight. Plotting and bounding these functions could provide insights in all kinds of collective decision-making problems. Further, our work can be seen as proportional decision-making on binary issues ("should $a$ be ranked above $b$?") under constraints (in our case, transitivity). This general topic has just started to be explored by researchers [Chandak et al., 2024, Lackner and Maly, 2023, Masařík et al., 2023, Skowron and Górecki, 2022].

## Acknowledgements

This work was funded in part by the French government under management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute).

## REFERENCES

Nir Ailon, Moses Charikar, and Alantha Newman. 2008. Aggregating inconsistent information: ranking and clustering. *Journal of the ACM (JACM)* 55, 5 (2008), 1–27. https://doi.org/10.1007/BF01200760 [→ p. 14, 38]

Stergios Athanasoglou. 2016. Strategyproof and efficient preference aggregation with Kemeny-based criteria. *Games and Economic Behavior* 95 (2016), 156–167. https://doi.org/10.1016/j.geb.2015.12.002 [→ p. 10]

Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare* 48, 2 (2017), 461–485. https://doi.org/10.1007/s00355-016-1019-3 [→ p. 2, 18]

Haris Aziz and Barton E. Lee. 2020. The expanding approvals rule: improving proportional representation and monotonicity. *Social Choice and Welfare* 54 (2020), 1–45. https://doi.org/10.1007/s00355-019-01208-3 [→ p. 18]

Georg Bachmeier, Felix Brandt, Christian Geist, Paul Harrenstein, Keyvan Kardel, Dominik Peters, and Hans Georg Seedig. 2019. $k$-Majority digraphs and the hardness of voting with a constant number of voters. *J. Comput. System Sci.* 105 (2019), 130–157. https://doi.org/10.1016/j.jcss.2019.04.005 [→ p. 14, 36]

John Bartholdi, III, Craig A. Tovey, and Michael A. Trick. 1989. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare* 6 (1989), 157–165. https://doi.org/10.1007/BF00303169 [→ p. 14]

Eli Ben-Naim. 2010. Mixing of diffusing particles. *Physical Review E* 82, 6 (2010), 061103. https://doi.org/10.1103/PhysRevLett.102.190602 [→ p. 13]

Therese Biedl, Franz J. Brandenburg, and Xiaotie Deng. 2009. On the complexity of crossings in permutations. *Discrete Mathematics* 309, 7 (2009), 1813–1823. https://doi.org/10.1016/j.disc.2007.12.088 [→ p. 14, 36]

Niclas Boehmer and Nathan Schaar. 2023. Collecting, classifying, analyzing, and using real-world elections. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 1706–1715. https://www.ifaamas.org/Proceedings/aamas2023/pdfs/p1706.pdf [→ p. 17]

Walter Bossert and Yves Sprumont. 2014. Strategy-proof preference aggregation: Possibilities and characterizations. *Games and Economic Behavior* 85 (2014), 109–126. https://doi.org/10.1016/j.geb.2014.01.015 [→ p. 10, 11]

Walter Bossert and Ton Storcken. 1992. Strategy-proofness of social welfare functions: The use of the Kemeny distance between preference orderings. *Social Choice and Welfare* 9 (1992), 345–360. https://doi.org/10.1007/BF00182575 [→ p. 10]

Robert Bredereck, Piotr Faliszewski, Rolf Niedermeier, Piotr Skowron, and Nimrod Talmon. 2020. Mixed integer programming with convex/concave constraints: Fixed-parameter tractability and applications to multicovering and voting. *Theoretical Computer Science* 814 (2020), 86–105. https://doi.org/10.1016/j.tcs.2020.01.017 [→ p. 14, 37]

Ioannis Caragiannis, David Kurokawa, Hervé Moulin, Ariel D. Procaccia, Nisarg Shah, and Junxing Wang. 2019. The unreasonable fairness of maximum Nash welfare. *ACM Transactions on Economics and Computation (TEAC)* 7, 3 (2019), 1–32. https://doi.org/10.1002/cpa.3160130102 [→ p. 14, 37]

Diptarka Chakraborty, Syamantak Das, Arindam Khan, and Aditya Subramanian. 2022. Fair rank aggregation. *Advances in Neural Information Processing Systems (NeurIPS)* 35 (2022), 23965–23978. https://proceedings.neurips.cc/paper_files/paper/2022/file/974309ef51ebd89034adc64a57e304f2-Paper-Conference.pdf [→ p. 2]

Nikhil Chandak, Shashwat Goel, and Dominik Peters. 2024. Proportional aggregation of preferences for sequential decision making. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*. https://arxiv.org/abs/2306.14858 [→ p. 18]

Vincent Conitzer, Andrew Davenport, and Jayant Kalagnanam. 2006. Improved bounds for computing Kemeny rankings. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI)*. 620–626. https://www.cs.cmu.edu/~conitzer/kemenyAAAI06.pdf [→ p. 14]

Vincent Conitzer, Matthew Rognlie, and Lirong Xia. 2009. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*. 109–115. http://www.cs.cmu.edu/~conitzer/preferenceIJCAI09.pdf [→ p. 8]

Don Coppersmith, Lisa K. Fleischer, and Atri Rurda. 2010. Ordering by weighted number of wins gives a good ranking for weighted tournaments. *ACM Transactions on Algorithms (TALG)* 6, 3 (2010), 1–13. https://doi.org/10.1007/BF01200760 [→ p. 14]

Persi Diaconis and Ronald L. Graham. 1977. Spearman's footrule as a measure of disarray. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 39, 2 (1977), 262–268. https://doi.org/10.1111/j.2517-6161.1977.tb01624.x [→ p. 18]

Cynthia Dwork, Ravi Kumar, Moni Naor, and Dandapani Sivakumar. 2001. Rank aggregation methods for the web. In *Proceedings of the 10th International Conference on World Wide Web (WWW)*. 613–622. https://doi.org/10.1145/371920.372165 [→ p. 14, 36]

Edith Elkind, Martin Lackner, and Dominik Peters. 2022. Preference restrictions in computational social choice: A survey. arXiv:2205.09092 [cs.GT] [→ p. 7]

Piotr Faliszewski, Andrzej Kaczmarczyk, Krzysztof Sornat, Stanisław Szufa, and Tomasz Wąs. 2023. Diversity, agreement, and polarization in elections. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*. 2684–2692. https://doi.org/10.24963/ijcai.2023/299 [→ p. 16]

Peter C. Fishburn. 1978. Axioms for approval voting: Direct proof. *Journal of Economic Theory* 19, 1 (1978), 180–185. https://doi.org/10.1016/0022-0531(78)90062-5 [→ p. 3]

Paul E. Green and Vithala R. Rao. 1972. *Applied multidimensional scaling: A comparison of approaches and algorithms.* [→ p. 17]

John G. Kemeny. 1959. Mathematics without numbers. *Daedalus* 88, 4 (1959), 577–591. https://www.jstor.org/stable/20026529 [→ p. 0, 1]

John G. Kemeny and J. Laurie Snell. 1960. Preference rankings: An axiomatic approach. In *Mathematical Models in the Social Sciences*. Ginn, 9–23. [→ p. 1]

Claire Kenyon-Mathieu and Warren Schudy. 2007. How to rank with few errors: A PTAS for weighted feedback arc set on tournaments. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC)*. 95–103. https://doi.org/10.1145/1250790.1250806 [→ p. 14, 38]

Caitlin Kuhlman and Elke Rundensteiner. 2020. Rank aggregation algorithms for fair consensus. *Proceedings of the VLDB Endowment* 13 (2020), 2706–2719. https://doi.org/10.14778/3407790.3407855 [→ p. 2]

Martin Lackner and Jan Maly. 2023. Proportional decisions in perpetual voting. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*. 5722–5729. https://doi.org/10.1609/aaai.v37i5.25710 [→ p. 18]

Martin Lackner and Piotr Skowron. 2021. Consistent approval-based multi-winner rules. *Journal of Economic Theory* 192 (2021), 105173. https://doi.org/10.1016/j.jet.2020.105173 [→ p. 3, 8]

Martin Lackner and Piotr Skowron. 2023. *Multi-Winner Voting with Approval Preferences*. Springer. https://doi.org/10.1007/978-3-031-09016-5 [→ p. 2, 3]

Patrick Lederer. 2023. Bivariate scoring rules: Unifying the characterizations of positional scoring rules and Kemeny's rule. (2023). https://pub.dss.in.tum.de/brandt-research/consistent_ranking.pdf [→ p. 8]

Colin L. Mallows. 1957. Non-null ranking models. *Biometrika* 44 (1957), 114–130. https://doi.org/10.2307/2333244 [→ p. 16]

Tomáš Masařík, Grzegorz Pierczyński, and Piotr Skowron. 2023. A generalised theory of proportionality in collective decision making. arXiv:2307.06077 [cs.GT] [→ p. 18]

Nicholas Mattei and Toby Walsh. 2013. Preflib: A library for preferences. In *Proceedings of the 4rd International Conference on Algorithmic Decision Theory (ADT)*. Springer, 259–270. https://doi.org/10.1111/j.1467-9280.2007.01950.x [→ p. 17]

Roger B. Myerson. 1995. Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare* 12, 1 (1995), 59–74. https://doi.org/10.1007/BF00182193 [→ p. 3, 8]

Selin Oğuz. 2023. Mapped: Air pollution levels around the world in 2022. *VC Elements* (2023). https://elements.visualcapitalist.com/mapped-air-pollution-levels-around-the-world-2022/ [→ p. 40]

Dominik Peters, Grzegorz Pierczyński, and Piotr Skowron. 2021. Proportional participatory budgeting with additive utilities. *Advances in Neural Information Processing Systems* 34 (2021), 12726–12737. https://arxiv.org/abs/2008.13276 [→ p. 2]

Paul Rothstein. 1991. Representative voter theorems. *Public Choice* 72, 2-3 (1991), 193–212. https://doi.org/10.1007/BF00123744 [→ p. 6]

Piotr Skowron, Piotr Faliszewski, and Arkadii Slinko. 2019. Axiomatic characterization of committee scoring rules. *Journal of Economic Theory* 180 (2019), 244–273. https://doi.org/10.1016/j.jet.2018.12.011 [→ p. 8]

Piotr Skowron and Adrian Górecki. 2022. Proportional public decisions. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*. 5191–5198. https://doi.org/10.1609/aaai.v36i5.20454 [→ p. 13, 18]

Piotr Skowron, Martin Lackner, Markus Brill, Dominik Peters, and Edith Elkind. 2017. Proportional rankings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. 409–415. https://www.ijcai.org/Proceedings/2017/0058.pdf [→ p. 18]

Warren S. Torgerson. 1952. Multidimensional scaling: I. Theory and method. *Psychometrika* 17, 4 (1952), 401–419. https://doi.org/10.1007/BF02288916 [→ p. 16]

Anke Van Zuylen and David P. Williamson. 2009. Deterministic pivoting algorithms for constrained ranking and clustering problems. *Mathematics of Operations Research* 34, 3 (2009), 594–620. https://doi.org/10.1007/978-3-540-77918-6_21 [→ p. 14]

Paolo Viappiani. 2015. Characterization of scoring rules with distances: application to the clustering of rankings. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*. 104–110. http://ijcai.org/papers15/Papers/IJCAI15-022.pdf [→ p. 18]

Dong Wei, Md Mouinul Islam, Baruch Schieber, and Senjuti Basu Roy. 2022. Rank aggregation with proportionate fairness. In *Proceedings of the 2022 International Conference on Management of Data (SIGMOD)*. 262–275. https://doi.org/10.1145/3514221.3517865 [→ p. 2]

WHO. 2024. WHO Ambient Air Quality Database (update 2024). Version 6.1. https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database [→ p. 40]

H. Peyton Young. 1975. Social choice scoring functions. *SIAM J. Appl. Math.* 28, 4 (1975), 824–838. https://doi.org/10.1016/0022-0531(74)90073-8 [→ p. 3, 8, 24]

H. Peyton Young. 1988. Condorcet's theory of voting. *American Political Science Review* 82, 4 (1988), 1231–1244. https://doi.org/10.1016/0022-0531(77)90011-4 [→ p. 1]

H. Peyton Young and Arthur Levenglick. 1978. A consistent extension of Condorcet's election principle. *SIAM J. Appl. Math.* 35, 2 (1978), 285–300. https://doi.org/10.1007/BFb0121198 [→ p. 1, 3, 8, 22]

## A PROOF OF THEOREM 3.2

In this appendix, we will provide a full proof of Theorem 3.2. To this end, we first note that the direction from left to right is easy, especially since we have already shown that the Squared Kemeny rule satisfies 2RP. We thus focus on the converse and suppose for this that $f$ is an SPF that satisfies neutrality, continuity, reinforcement, and 2RP. To prove that $f$ is the Squared Kemeny rule, we will use an equivalent formulation of this rule by exchanging the minimum with a maximum in its definition: $\mathrm{SqK}(R) = \arg\max_{\rhd\in\mathcal{R}} -\sum_{\succ\in\mathcal{R}} R(\succ)\,\mathrm{swap}(\succ,\rhd)^2$. Then, our goal is to show that $f$ is the SPF that chooses the rankings that maximize $-\sum_{\succ\in\mathcal{R}} R(\succ)\,\mathrm{swap}(\succ,\rhd)^2$. For this, we will use a hyperplane argument as, e.g., showcased by Young and Levenglick [1978].

As a first step, we hence change the domain of $f$ from ranking profiles to a numerical space. To this end, let $b : \{1,\dots,|\mathcal{R}|\} \to \mathcal{R}$ denote an enumeration of all possible input rankings. Moreover, we define $\mathbb{T} = \{v \in \mathbb{Q}^{m!} \colon \sum_{i=1}^{m!} v_i = 1 \wedge v_i \geqslant 0 \text{ for all } i \in \{1,\dots,m!\}\}$. Using the enumeration $b$, we can represent every profile $R$ as a vector $v \in \mathbb{T}$ by defining $v_i = R(b(i))$ for all $i \in \{1,\dots,m!\}$. For simplicity, we will denote the vector associated with a profile $R$ by $v(R)$. Moreover, for every permutation $\tau : A \to A$, we define the permutation of a vector $v$ by $\tau(v)_i = v_j$ for all $i, j$ such that $\tau(b(j)) = b(i)$. That is, if the ranking $b(j)$ has weight $v_j$ in the profile $R$, then the ranking $\tau(b(j)) = b(i)$ has weight $\tau(v)_i = v_j$ in the permuted profile $\tau(R)$. Hence, $\tau(v(R)) = v(\tau(R))$.

By the definition of SPFs and profiles, it is straightforward that there is a function $g : \mathbb{T} \to 2^{\mathcal{R}} \setminus \{\emptyset\}$ such that $f(R) = g(v(R))$ for all profiles $R$. Furthermore, $g$ inherits the desirable properties of $f$:

- $g$ satisfies neutrality: it holds for all $v \in \mathbb{T}$ and all permutations $\tau : A \to A$ that $g(\tau(v)) = \{\tau(\rhd) \colon \rhd \in g(v)\}$.
- $g$ satisfies reinforcement: it holds for all $v, v' \in \mathbb{T}$ with $g(v) \cap g(v') \neq \emptyset$ and all $\lambda \in (0,1) \cap \mathbb{Q}$ that $g(\lambda v + (1-\lambda)v') = g(v) \cap g(v')$.
- $g$ satisfies continuity: it holds for all $v, v' \in \mathbb{T}$ that there is a constant $\lambda \in (0,1) \cap \mathbb{Q}$ such that $g(\lambda v + (1-\lambda)v') \subseteq g(v)$.

We next extend $g$ to the domain $\mathbb{Q}^{m!}$ while preserving its desirable properties.

Lemma A.1. *There is a neutral, reinforcing, and continuous function $\hat{g} : \mathbb{Q}^{m!} \to 2^{\mathcal{R}} \setminus \{\emptyset\}$ such that $f(R) = \hat{g}(v(R))$ for all profiles $R \in \mathcal{R}^*$.*

Proof. As we observed before this lemma, there is a neutral, reinforcing, and continuous function $g : \mathbb{T} \to 2^{\mathcal{R}} \setminus \{\emptyset\}$ such that $f(R) = g(v(R))$ for all $R \in \mathcal{R}^*$. We will prove this lemma by extending $g$ to $\mathbb{Q}^{m!}$. To this end, we will first extend the domain of $g$ to $\mathbb{T}^+ = \{v \in \mathbb{Q}^{m!} \colon \sum_{i=1}^{m!} v_i > 0 \wedge v_i \geqslant 0 \text{ for all } i \in \{1,\dots,m!\}\}$ and then to $\mathbb{Q}^{m!}$.

**Step 1: Extension to $\mathbb{T}^+$**

For extending $g$ to the domain $\mathbb{T}^+$, we let $\bar{g}(v) = g(\lambda v)$ for all $v \in \mathbb{T}$, where $\lambda = \frac{1}{\sum_{i=1}^{m!} v_i}$ is the unique scalar such that $\sum_{i=1}^{m!} \lambda v_i = 1$. First, we note that $\bar{g}$ is defined for all $v \in \mathbb{T}^+$ as the only difference between $\mathbb{T}$ and $\mathbb{T}^+$ is the assumption that $\sum_{i=1}^{m!} v_i = 1$ for all $v \in \mathbb{T}$, whereas $\sum_{i=1}^{m!} v_i > 0$ for all $v \in \mathbb{T}^+$. Moreover, we note that $f(R) = g(v(R)) = \bar{g}(v(R))$ for all profiles $R$ as $\bar{g}(v) = g(v)$ for all profiles $R$. Next, we will show that $\bar{g}$ is neutral, continuous, and reinforcing. To this end, let $v^1 \in \mathbb{T}^+$ denote an arbitrary vector and let $\lambda_1 > 0$ denote the scalar such that $\bar{g}(v^1) = g(\lambda_1 v^1)$.

For neutrality, we note that $\tau(\lambda_1 v^1) = \lambda_1 \tau(v^1)$ for every permutation $\tau : A \to A$, so it follows that $\bar{g}(\tau(v^1)) = g(\lambda\tau(v^1)) = g(\tau(\lambda_1 v^1)) = \{\tau(\rhd) \colon \rhd \in g(\lambda_1 v^1)\} = \{\tau(\rhd) \colon \rhd \in \bar{g}(v^1)\}$. Here, the third equality follows from the neutrality of $g$. This argument shows that $\bar{g}$ is neutral.

Next, we turn to reinforcement and hence let $v^2 \in \mathbb{T}^+$ denote a second vector with $\bar{g}(v^1) \cap \bar{g}(v^2) \neq \emptyset$ and $\lambda_2$ the scalar such that $\bar{g}(v^2) = g(\lambda_2 v^2)$. We need to show that $\bar{g}(\kappa v^1 + (1-\kappa)v^2) = \bar{g}(v^1) \cap \bar{g}(v^2)$ for all $\kappa \in (0,1) \cap \mathbb{Q}$. To this end, we fix such a $\kappa$ and first note that there is by definition a scalar $\lambda_3 > 0$ such that $\bar{g}(\kappa v^1 + (1-\kappa)v^2) = g(\lambda_3(\kappa v^1 + (1-\kappa)v^2))$. Next, we observe that $\bar{g}(v^1) \cap \bar{g}(v^2) = g(\lambda_1 v^1) \cap \bar{g}(\lambda_2 v^2)$, so we can infer from the reinforcement of $g$ that $\bar{g}(v^1) \cap \bar{g}(v^2) = g(\kappa'\lambda_1 v^1 + (1-\kappa')\lambda_2 v^2)$ for every $\kappa' \in (0,1) \cap \mathbb{Q}$. To show that $\bar{g}$ is reinforcing, it hence suffices to find a $\kappa' \in (0,1) \cap \mathbb{Q}$ such that $\lambda_3(\kappa v^1 + (1-\kappa)v^2) = \kappa'\lambda_1 v^1 + (1-\kappa')\lambda_2 v^2$. To this end, we note that $\lambda_3(\kappa v^1 + (1-\kappa)v^2) = (\frac{\lambda_3}{\lambda_1}\cdot\kappa)\cdot\lambda_1 v^1 + (\frac{\lambda_3}{\lambda_2}\cdot(1-\kappa))\cdot\lambda_2 v^2$. Now, since $\sum_{i=1}^{m!} \lambda_3(\kappa v_i^1 + (1-\kappa)v_i^2) = \sum_{i=1}^{m!} \lambda_1 v_i^1 = \sum_{i=1}^{m!} \lambda_1 v_i^2 = 1$, we can infer that $\frac{\lambda_3}{\lambda_1}\cdot\kappa + \frac{\lambda_3}{\lambda_2}\cdot(1-\kappa) = 1$. Moreover, because $\frac{\lambda_3}{\lambda_1} = \frac{\sum_{i=1}^{m!} v_i^1}{\sum_{i=1}^{m!} v_i^1 + v_i^2} \in (0,1) \cap \mathbb{Q}$ and $\kappa \in (0,1) \cap \mathbb{Q}$, we can infer that both $\frac{\lambda_3}{\lambda_1}\cdot\kappa \in (0,1) \cap \mathbb{Q}$ and $\frac{\lambda_3}{\lambda_2}\cdot(1-\kappa) \in (0,1) \cap \mathbb{Q}$. Hence, we now define $\kappa' = \frac{\lambda_3}{\lambda_1}\cdot\kappa$ (which implies that $1-\kappa' = \frac{\lambda_3}{\lambda_2}\cdot(1-\kappa)$). It then follows that $\lambda_3(\kappa v^1 + (1-\kappa)v^2) = \kappa'\lambda_1 v^1 + (1-\kappa')\lambda_2 v^2$, thus proving that $\bar{g}$ is reinforcing.

Finally, we turn to the continuity of $\bar{g}$ and again consider a vector $v^2 \in \mathbb{T}^+$ with its scalar $\lambda^2$. Our goal is to show that there is $\kappa \in (0,1) \cap \mathbb{Q}$ such that $\bar{g}(\kappa v^1 + (1-\kappa)v^2) \subseteq \bar{g}(v^1)$. This is equivalent to finding a $\hat{\kappa} \in \mathbb{Q}$ such that $\hat{\kappa} > 0$ and $\bar{g}(\hat{\kappa}v^1 + v^2) \subseteq \bar{g}(v^1)$. To see this equivalence, we define $\kappa = \frac{\hat{\kappa}}{\hat{\kappa}+1}$, $\hat{\lambda}_3 = \frac{1}{\hat{\kappa}\sum_{i=1}^{m!} v_i^1 + \sum_{i=1}^{m!} v_i^2}$, and $\lambda_3 = \frac{1}{\kappa\sum_{i=1}^{m!} v_i^1 + (1-\kappa)\sum_{i=1}^{m!} v_i^2}$. It holds that $\hat{\lambda}_3 = \frac{1}{\hat{\kappa}+1}\lambda_3$, so $\bar{g}(\hat{\kappa}v^1 + v^2) = g(\hat{\lambda}_3(\hat{\kappa}v^1 + v^2)) = g(\lambda_3(\kappa v^1 + (1-\kappa)v^2)) = \bar{g}(\kappa v^1 + (1-\kappa)v^2)$. Finally, to infer a suitable $\hat{\kappa}$, we note that $g$ itself is continuous. Hence, there is a $\kappa' \in (0,1) \cap \mathbb{Q}$ such that $g(\kappa'\lambda_1 v^1 + (1-\kappa')\lambda_2 v^2) \subseteq g(\lambda v^1) = \bar{g}(v^1)$. We thus define $\hat{\kappa}$ by $\hat{\kappa} = \frac{\kappa'\lambda_1}{(1-\kappa')\lambda_2}$ because then $\bar{g}(\hat{\kappa}v^1 + v^2) = g(\kappa'\lambda_1 v^1 + (1-\kappa')\lambda_2 v^2) \subseteq \bar{g}(v^1)$. This implies that $\bar{g}$ is continuous.

**Step 2: Extension to $\mathbb{Q}^{m!}$**

Next, we will extend $\bar{g}$ to the domain $\mathbb{Q}^{m!}$. To this end, we define $v^*$ as the vector with $v_i^* = \frac{1}{m!}$ for all $i \in \{1, \dots, m!\}$. Since $\tau(v^*) = v^*$ for every permutation $\tau : A \to A$, it follows from the neutrality of $\bar{g}$ and $g$ that $\bar{g}(v^*) = g(v^*) = \mathcal{R}$. Now, to extend $\bar{g}$ to $\mathbb{Q}^{m!}$, we define $\hat{g}(v) = \bar{g}(v + \lambda v^*)$ for all $v \in \mathbb{Q}^{m!}$, where $\lambda \in \mathbb{Q}$ is a positive constant such that $v + \lambda v^* \in \mathbb{T}^+$.

As the first point, we will show that $\hat{g}$ is well-defined despite the fact that we do not fully specify $\lambda$. To this end, let $v \in \mathbb{Q}^{m!}$ and consider two distinct positive constants $\lambda_1, \lambda_2$ such that $v + \lambda_1 v^* \in \mathbb{T}$ and $v + \lambda_2 v^* \in \mathbb{T}$. We need to prove that $\bar{g}(v + \lambda_1 v^*) = \bar{g}(v + \lambda_2 v^*)$. Note for this that $\bar{g}$ is by definition homogenous, i.e., it holds for every rational constant $\kappa > 0$ that $\bar{g}(v') = \bar{g}(\kappa v')$ because $\bar{g}$ will rescale its input vector such that $\alpha\sum_{i=1}^{m!} v_i' = 1$ and then apply $g$. Hence, we have that $\bar{g}(v + \lambda_i v^*) = \bar{g}(\frac{1}{2}v + \frac{1}{2}\lambda_i v^*)$ for $i \in \{1,2\}$. Now, if $\lambda_1 < \lambda_2$, it follows that $\bar{g}(v + \lambda_1 v^*) = \bar{g}(v + \lambda_1 v^*) \cap \bar{g}((\lambda_2 - \lambda_1)v^*) = \bar{g}(\frac{1}{2}v + \frac{1}{2}\lambda_1 v^*) \cap \bar{g}(\frac{1}{2}(\lambda_2 - \lambda_1)v^*) = \bar{g}(\frac{1}{2}v + \frac{1}{2}\lambda_2 v^*) = \bar{g}(v + \lambda_2 v^*)$. The first equality here uses that $\bar{g}((\lambda_2 - \lambda_1)v^*) = \mathcal{R}$, the second one that $\bar{g}$ is homogenous, and the third one that $\bar{g}$ is reinforcing. Since an analogous argument works if $\lambda_1 > \lambda_2$, it now follows that $\hat{g}$ is well-defined. Moreover, we note that $f(R) = \bar{g}(v(R) + 0v^*) = \hat{g}(v(R))$ for all profiles $R$.

It remains to show that $\hat{g}$ is neutral, reinforcing, and continuous. For this, let $v^1$ denote a vector in $\mathbb{Q}^{m!}$ and let $\lambda_1$ denote a positive scalar in $\mathbb{Q}$ such that $v^1 + \lambda_1 v^* \in \mathbb{T}^+$. For neutrality, let $\tau$ additionally denote a permutation on $A$. Due to the neutrality of $\bar{g}$, it holds that $\hat{g}(\tau(v^1)) = \bar{g}(\tau(v^1) + \lambda_1 v^*) = \bar{g}(\tau(v^1 + \lambda_1 v^*)) = \{\tau(\rhd) : \rhd \in \bar{g}(v^1 + \lambda_1 v^*)\} = \{\tau(\rhd) : \rhd \in \hat{g}(v^1)\}$. This shows that $\bar{g}$ is neutral.

For reinforcement, let $v^2$ denote a vector in $\mathbb{Q}^{m!}$ and $\lambda_2$ a scalar such that $v^2 + \lambda_2 v^* \in \mathbb{T}^+$ and $\hat{g}(v^1) \cap \hat{g}(v^2) = \bar{g}(v^1 + \lambda_1 v^*) \cap \bar{g}(v^2 + \lambda_2 v^*) \neq \emptyset$. Since $\bar{g}$ is reinforcing, it follows for every $\kappa \in (0,1) \cap \mathbb{Q}$ that $\hat{g}(\kappa v^1 + (1 - \kappa v^2)) = \bar{g}(\kappa(v^1 + \lambda_1 v^*) + (1-\kappa)(v^2 + \lambda_2 v^*)) = \bar{g}(v^1 + \lambda_1 v^*) \cap \bar{g}(v^2 + \lambda_2 v^*) = \hat{g}(v^1) \cap \hat{g}(v^2)$. This proves that $\hat{g}$ is also reinforcing.

Finally, for continuity, we let $v^2 \in \mathbb{Q}^{m!}$ again denote a second vector and $\lambda_2$ a corresponding scalar. We need to show that there is $\kappa \in (0,1) \cap \mathbb{Q}$ such that $\hat{g}(\kappa v^1 + (1-\kappa)v^2) \subseteq \hat{g}(v^1)$. To this end, we note that there is $\kappa' \in (0,1) \cap \mathbb{Q}$ such that $\bar{g}(\kappa'(v^1 + \lambda_1 v^*) + (1-\kappa')(v^2 + \lambda_2 v^*)) \subseteq \bar{g}(v^1 + \lambda_1 v^*) = \hat{g}(v^1)$. It now follows from the definition of $\hat{g}$ that $\hat{g}(\kappa' v^1 + (1-\kappa')v^2) = \bar{g}(\kappa' v^1 + (1-\kappa')v^2 + (\kappa'\lambda_1 + (1-\kappa')\lambda_2)v^*) \subseteq \hat{g}(v^1)$. This shows that $\hat{g}$ is also continuous and hence completes the proof of this lemma. □

We note that our SPF $f$ uniquely entails the function $\hat{g}$ and that $\hat{g}$ fully describes $f$. Hence, we aim to describe the function $\hat{g}$ based on a scoring function $s(\succ, \rhd)$. To this end, we first note that $\hat{g}$ is *homogenous* (i.e., $\hat{g}(v) = \hat{g}(\lambda v)$ for every $v \in \mathbb{Q}^{m!}$ and $\lambda \in \mathbb{Q}$ with $\lambda > 0$). To see this, we recall that $\bar{g}$ is by definition homogeneous and let $v \in \mathbb{Q}^{m!}$ denote an arbitrary vector and $\kappa \in \mathbb{Q}$ a positive scalar. Moreover, let $\lambda$ denote a scalar such that $v + \lambda v^* \in \mathbb{T}^+$. Then, the homogeneity of $\hat{g}$ follows as $\hat{g}(v) = \bar{g}(v + \lambda v^*) = \bar{g}(\kappa(v + \lambda v^*)) = \hat{g}(\kappa v)$.

Next, we further modify the representation of $f$ by considering the sets $R_{\rhd_i} = \{v \in \mathbb{Q}^{m!} : \rhd_i \in \hat{g}(v)\}$ for all $\rhd_i \in \mathcal{R}$. All sets $R_{\rhd_i}$ are symmetric to each other (i.e., if $v \in R_{\rhd_i}$, then $\tau(v) \in R_{\tau(\rhd_i)}$) and $\mathbb{Q}$-convex (i.e, if $v, v' \in R_{\rhd_i}$, then $\lambda v + (1-\lambda)v' \in R_{\rhd_i}$ for all $\lambda \in (0,1) \cap \mathbb{Q}$) because $\hat{g}$ is neutral and reinforcing. Moreover, since the domain of $\hat{g}$ is $\mathbb{Q}^{m!}$, it follows that $\bigcup_{\rhd_i \in \mathcal{R}} R_{\rhd_i} = \mathbb{Q}^{m!}$. Further, we denote with $\bar{R}_{\rhd_i}$ the closure of $R_{\rhd_i}$ with respect to $\mathbb{R}^{m!}$. In particular, the sets $\bar{R}_{\rhd_i}$ are convex and symmetric to each other and $\bigcup_{\rhd_i \in \mathcal{R}} \bar{R}_{\rhd_i} = \mathbb{R}^{m!}$ (see Young [1975]). As the last point, we note that $\hat{g}(v) = \{\rhd_i \in \mathcal{R} : v \in R_{\rhd_i}\} \subseteq \{\rhd_i \in \mathcal{R} : v \in \bar{R}_{\rhd_i}\}$ for all $v \in \mathbb{Q}^{m!}$.

We next aim to find a suitable representation for the sets $\bar{R}_{\rhd_i}$ and use for this the separating hyperplane theorem for convex sets. In the next lemmas, we write $uv = \sum_{i=1}^{k} u_i v_i$ for the standard scalar product between two vectors $u, v \in \mathbb{R}^k$.

Lemma A.2. *For all distinct rankings $\rhd_i, \rhd_j \in \mathcal{R}$, there is a non-zero vector $u^{\rhd_i, \rhd_j} \in \mathbb{R}^{m!}$ such that $vu^{\rhd_i, \rhd_j} \geqslant 0$ if $v \in \bar{R}_{\rhd_i}$ and $vu^{\rhd_i, \rhd_j} \leqslant 0$ if $v \in \bar{R}_{\rhd_j}$.*

Proof. Consider two distinct rankings $\rhd_i, \rhd_j \in \mathcal{R}$ and their respective sets $\bar{R}_{\rhd_i}$ and $\bar{R}_{\rhd_j}$. First, we recall that $\bigcup_{\succ_k \in \mathcal{R}} \bar{R}_{\rhd_k} = \mathbb{R}^{m!}$ and that all these sets are symmetric to each other. Since there are only a finitely many such sets, this implies that all $\bar{R}_{\rhd_k}$ are fully dimensional. Consequently, $\operatorname{int} \bar{R}_{\rhd_i} \neq \emptyset$ and $\operatorname{int} \bar{R}_{\rhd_j} \neq \emptyset$.

We will next show that $\operatorname{int} \bar{R}_{\rhd_i} \cap \operatorname{int} \bar{R}_{\rhd_j} = \emptyset$. Assume for contradiction that this is not the case. Then, there is a vector $v \in \operatorname{int} \bar{R}_{\rhd_i} \cap \operatorname{int} \bar{R}_{\rhd_j} \cap \mathbb{Q}^{m!}$. In particular, these conditions entail that $\rhd_i, \rhd_j \in \hat{g}(v)$. Next, consider the profile $R$ such that $R(\rhd_i) = \frac{m(m-1)}{m(m-1)+1}$ and $R(\rhd_j) = \frac{1)}{m(m-1)+1}$, and let $v' = v(R)$ denote the corresponding vector. By 2RP, we have that $f(R) = \hat{g}(v') = \{\rhd_i\}$ because

$$\operatorname{swap}(\rhd_i, \rhd_j) \cdot \frac{1}{m(m-1)+1} \leqslant \frac{\binom{m}{2}}{m(m-1)+1} < \frac{1}{2}.$$

The homogeneity of $\hat{g}$ then shows that $\hat{g}(\lambda v') = \{\rhd_i\}$ for every $\lambda \in \mathbb{Q}$ with $\lambda > 0$. Moreover, reinforcement shows that $\hat{g}(v + \lambda v') = \hat{g}(\frac{1}{2}v + \frac{1}{2}\lambda v') = \hat{g}(v) \cap \hat{g}(\lambda v') = \{\rhd_i\}$ for all $\lambda \in \mathbb{Q}$ with $\lambda > 0$. However, $v \in \operatorname{int} \bar{R}_{\rhd_i} \cap \operatorname{int} \bar{R}_{\rhd_j} \cap \mathbb{Q}^{m!}$ implies that there is $\lambda \in \mathbb{Q}$ with $\lambda > 0$ such that $\rhd_i, \rhd_j \in \hat{g}(v + \lambda v')$, contradicting our previous insight. Hence, our initial assumption must have been wrong and the sets $\operatorname{int} \bar{R}_{\rhd_i}$ and $\operatorname{int} \bar{R}_{\rhd_j}$ are indeed disjoint.

Now, by the separating hyperplane theorem for convex sets, there is a non-zero vector $u^{\rhd_i, \rhd_j}$ such that $vu^{\rhd_i, \rhd_j} > 0$ if $v \in \operatorname{int} \bar{R}_{\rhd_i}$ and $vu^{\rhd_i, \rhd_j} < 0$ if $v \in \operatorname{int} \bar{R}_{\rhd_j}$. In particular, note that the constant on the right side of our inequalities must be 0 as $\bar{R}_{\rhd_i}$ and $\bar{R}_{\rhd_j}$ are cones. This implies that $vu^{\rhd_i, \rhd_j} \geqslant 0$ for all $v \in \bar{R}_{\rhd_i}$ and $vu^{\rhd_i, \rhd_j} \leqslant 0$ for all $v \in \bar{R}_{\rhd_j}$. ◻

We say that a non-zero vector $u$ *separates* a set $\bar{R}_{\rhd_i}$ from another set $\bar{R}_{\rhd_j}$ if $vu \geqslant 0$ for all $v \in \bar{R}_{\rhd_i}$ and $vu \leqslant 0$ for all $v \in \bar{R}_{\rhd_j}$. Our interest in these vectors comes from the next lemma which states that the vectors that separate $\bar{R}_{\rhd_i}$ from any other set $\bar{R}_{\rhd_j}$ fully describe the set $\bar{R}_{\rhd_i}$.

Lemma A.3. *Let $\rhd_i \in \mathcal{R}$ denote an arbitrary ranking and let $u^{\rhd_i, \rhd_j}$ denote non-zero vectors that separates $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$ for every $\rhd_j \in \mathcal{R} \setminus \{\rhd_i\}$. It holds that*

$$\bar{R}_{\rhd_i} = \{v \in \mathbb{R}^{m!} : \forall \rhd_j \in \mathcal{R} \setminus \{\rhd_i\} : vu^{\rhd_i, \rhd_j} \geqslant 0\}.$$

Proof. Fix a ranking $\rhd_i$ and the vectors $u^{\rhd_i,\rhd_j}$ for all $\rhd_j \in \mathcal{R} \setminus \{\rhd_i\}$. For a simpler notation, we define $S_{\rhd_i} = \{v \in \mathbb{R}^{m!} \colon \forall \rhd_j \in \mathcal{R} \setminus \{\rhd_i\} \colon v u^{\rhd_i,\rhd_j} \geqslant 0\}$, and we will show that $S_{\rhd_i} \subseteq \bar{R}_{\rhd_i}$ and $\bar{R}_{\rhd_i} \subseteq S_{\rhd_i}$. The second subset relation is straightforward by the definition of the vectors $u^{\rhd_i,\rhd_j}$: if $v \in \bar{R}_{\rhd_i}$, then $v u^{\rhd_i,\rhd_j} \geqslant 0$ for all $\rhd_j \neq \rhd_i$ and therefore $v \in S_{\rhd_i}$.

For the converse direction, we first note that $\operatorname{int} S_{\rhd_i}$ is non-empty as $\bar{R}_{\rhd_i} \subseteq S_{\rhd_i}$ and $\operatorname{int} \bar{R}_{\rhd_i} \neq \emptyset$. Now, let $v \in \operatorname{int} S_{\rhd_i}$, which means that $v u^{\rhd_i,\rhd_j} > 0$ for all $\rhd_j \in \mathcal{R} \setminus \{\rhd_i\}$. By the definition of the vectors $u^{\rhd_i,\rhd_j}$, it thus follows that $v \notin \bar{R}_{\rhd_j}$ for all $\rhd_j \in \mathcal{R} \setminus \{\rhd_i\}$. Since $\bigcup_{\rhd_i \in \mathcal{R}} \bar{R}_{\rhd_i} = \mathbb{R}^{m!}$, we derive that $v \in \bar{R}_{\rhd_i}$, so $\operatorname{int} S_{\rhd_i} \subseteq \bar{R}_{\rhd_i}$. Because $\bar{R}_{\rhd_i}$ is a closed set, we finally conclude that $S_{\rhd_i} \subseteq \bar{R}_{\rhd_i}$, which completes the proof of this lemma. □

Lemma A.3 states that the vectors $u^{\rhd_i,\rhd_j}$ fully describe the sets $\bar{R}_{\rhd_i}$ which, in turn, describe our function $\hat{g}$ because $\hat{g}(v) \subseteq \{\rhd_i \in \mathcal{R} \colon v \in \bar{R}_{\rhd_i}\}$ for all $v \in \mathbb{Q}^{m!}$ as $R_{\rhd_i} \subseteq \bar{R}_{\rhd_i}$. We show next that the subset relation between $\hat{g}(v)$ and $\{\rhd_i \in \mathcal{R} \colon v \in \bar{R}_{\rhd_i}\}$ is an equality for vectors in $\mathbb{Q}^{m!}$.

Lemma A.4. *It holds for all $v \in \mathbb{Q}^{m!}$ that $v \in \bar{R}_{\rhd_i}$ if and only if $\rhd_i \in \hat{g}(v)$.*

Proof. Consider an arbitrary vector $v \in \mathbb{Q}^{m!}$. First, it immediately follows from the definition of the sets $\bar{R}_{\rhd_i}$ that $v \in \bar{R}_{\rhd_i}$ if $\rhd_i \in \hat{g}(v)$. For the other direction, suppose for contradiction that there is $v \in \mathbb{Q}^{m!}$ and $\rhd_i \in \mathcal{R}$ such that $\rhd_i \notin \hat{g}(v)$ but $v \in \bar{R}_{\rhd_i}$. In this case, we consider the vectors $u^{\rhd_i,\rhd_j}$ that separate $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$ for all $\rhd_j \in \mathcal{R}$. Due to the definitions of these vectors, we derive that $v u^{\rhd_i,\rhd_j} = 0$ for all $\rhd_j \in \hat{g}(v)$ because $v \in \bar{R}_{\rhd_j}$ if $\rhd_j \in \hat{g}(v)$. Next, we note that there is a profile $R$ such that $\hat{g}(v(R)) = f(R) = \{\rhd_i\}$ because of 2RP. By the continuity of $\hat{g}$, there is for every vector $v' \in \mathbb{Q}^{m!}$ a $\lambda \in (0,1) \cap \mathbb{Q}$ such that $\hat{g}(\lambda v(R) + (1-\lambda)v') = \{\rhd_i\}$. This shows that $v(R) \in \operatorname{int} \bar{R}_{\rhd_i}$. By Lemma A.3, it thus follows that $v(R)u^{\rhd_i,\rhd_j} > 0$ for all $\rhd_j \in \mathcal{R} \setminus \{\rhd_i\}$. Now, using again continuity, there is $\lambda \in (0,1) \cap \mathbb{Q}$ such that $\hat{g}(\lambda v + (1-\lambda)v(R)) \subseteq \hat{g}(v)$. However, if we consider the vectors $u^{\rhd_i,\rhd_j}$, we infer for all $\rhd_j \in \hat{g}(v)$ that $v(\lambda v + (1-\lambda)v(R))u^{\rhd_i,\rhd_j} = (1-\lambda)v(R)u^{\rhd_i,\rhd_j} > 0$, which contradicts that $v \in \bar{R}_{\rhd_j}$ for $\rhd_j \in \hat{g}(v)$. Hence, our assumption that $v \in \bar{R}_{\rhd_i}$ and $\rhd_i \notin g(v)$ is wrong. □

Lemmas A.3 and A.4 show that it suffices to determine the vectors $u^{\rhd_i,\rhd_j}$ to characterize $\hat{g}$ and thus $f$. We start the analysis of these vectors for rankings that only differ in a single swap. Recall for the next lemma that $b$ is the function that enumerates all possible rankings. We moreover note that the subsequent lemma is very similar to Step 1 in the proof of Theorem 3.3.

Lemma A.5. *Let $\rhd_i, \rhd_j \in \mathcal{R}$ denote two rankings such that* $\operatorname{swap}(\rhd_i, \rhd_j) = 1$. *The vector $u$ defined by $u_k = -\operatorname{swap}(\rhd_i, b(k))^2 + \operatorname{swap}(\rhd_j, b(k))^2$ for all $k \in \{1, \dots, m!\}$ separates $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$.*

Proof. Consider two arbitrary rankings $\rhd_i, \rhd_j \in \mathcal{R}$ such that $\operatorname{swap}(\rhd_i, \rhd_j) = 1$ and let $a, b$ denote the alternatives with $a \rhd_i b$ and $b \rhd_j a$. Moreover, let $u^{\rhd_i,\rhd_j}$ denote the non-zero vector that separates $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$ given by Lemma A.2. Based on $u^{\rhd_i,\rhd_j}$, we will show that the vector $u$ defined in this lemma also separates $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$.

As a first step, we consider the profiles $R$ and $R'$ with $R(\rhd_i) = R(\rhd_j) = \frac{1}{2}$, and $R'(\rhd_1) = \frac{2}{3}$ and $R'(\rhd_2) = \frac{1}{3}$. 2RP implies for the profile $R$ that $f(R) = \hat{g}(v(R)) = \{\rhd_i, \rhd_j\}$, so we conclude that $v(R) \in \bar{R}_{\rhd_i} \cap \bar{R}_{\rhd_j}$. This means that $v(R)u^{\rhd_i,\rhd_j} = 0$ and hence $u_i^{\rhd_i,\rhd_j} = -u_j^{\rhd_i,\rhd_j}$ (we assume here that $b(i) = \rhd_i$, and $b(j) = \rhd_j$ for simplicity). By contrast, 2RP requires for $R'$ that $f(R') = \hat{g}(v(R')) = \{\rhd_i\}$. Thus, Lemma A.4 entails that $v(R)$ is only contained in the set $\bar{R}_{\rhd_i}$. Moreover, based on continuity, it is easy to show that $v(R)$ is even in $\operatorname{int} \bar{R}_{\rhd_i}$. Consequently, Lemma A.3 implies that $v(R')u^{\rhd_i,\rhd_j} > 0$. Since $v(R')u^{\rhd_i,\rhd_j} = \frac{2}{3}u_i^{\rhd_i,\rhd_j} + \frac{1}{3}u_j^{\rhd_i,\rhd_j}$ and $u_i^{\rhd_i,\rhd_j} = -u_j^{\rhd_i,\succ_j}$, we can now conclude that $u_i^{\rhd_i,\rhd_j} > 0$ and $u_j^{\rhd_i,\rhd_j} < 0$. Because every rescaling of $u^{\rhd_i,\rhd_j}$ also separates $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$, we assume from now on that

$$u_i^{\rhd_i,\rhd_j} = 1 = -\operatorname{swap}(\rhd_i,\rhd_i)^2 + \operatorname{swap}(\rhd_i,\rhd_j) \qquad \text{and}$$
$$u_j^{\rhd_i,\rhd_j} = -1 = -\operatorname{swap}(\rhd_j,\rhd_i)^2 + \operatorname{swap}(\rhd_j,\rhd_j).$$

Next, we consider an arbitrary index $k$ such that $b \succ a$ for the ranking $\succ\, = b(k)$. Moreover, we define $d = \operatorname{swap}(\succ, \rhd_i)$ and consider the profile $R$ with $R(\succ) = \frac{1}{2d}$ and $R(\rhd_i) = \frac{2d-1}{2d}$. For this profile, it holds that $d \cdot (1 - R(\succ)) = d - \frac{1}{2}$ and $d \cdot (1 - R(\rhd_i)) = \frac{1}{2}$. Hence, 2RP requires that

$$f(R) = \{\rhd \in \mathcal{R} : (\operatorname{swap}(\succ,\rhd) = d \quad \text{and} \quad \operatorname{swap}(\rhd_i,\rhd) = 0)$$
$$\text{or } (\operatorname{swap}(\succ,\rhd) = d-1 \quad \text{and} \quad \operatorname{swap}(\rhd_i,\rhd) = 1)\}.$$

Thus, $\rhd_i \in f(R) = \hat{g}(v(R))$ because $\operatorname{swap}(\rhd_i,\succ) = d$ and $\operatorname{swap}(\rhd_i,\rhd_i) = 0$, and $\rhd_j \in f(R)$ because $\operatorname{swap}(\rhd_j,\succ) = d-1$ and $\operatorname{swap}(\rhd_i,\rhd_j) = 1$. This means that $v(R) \in \bar{R}_{\rhd_i} \cap \bar{R}_{\rhd_j}$ and the definition of the vector $u^{\rhd_i,\rhd_j}$ therefore implies that $v(R)u^{\rhd_i,\rhd_j} = 0$. We thus infer that

$$u_k^{\rhd_i,\rhd_j} = -(2d-1)u_i^{\rhd_i,\rhd_j} = -2(d-1) = -d^2 + (d-1)^2 = \operatorname{swap}(b(k),\rhd_i)^2 - \operatorname{swap}(b(k),\rhd_j)^2. \quad (3)$$

Finally, we can infer the entries $u_\ell^{\rhd_i,\rhd_j}$ for every input ranking $\succ\, = b(\ell)$ with $a \succ b$ based on a symmetric argument by exchanging the roles of $\rhd_i$ and $\rhd_j$. In more detail, we first compute $d = \operatorname{swap}(\succ,\rhd_j)$, then consider the profile $R$ with $R(\succ) = \frac{1}{2d}$ and $R(\rhd_j) = \frac{2d-1}{2d}$, and lastly determine the output for this profile based on 2RP. This approach yields that

$$u_\ell^{\rhd_i,\rhd_j} = -\operatorname{swap}(b(\ell),\rhd_i)^2 + \operatorname{swap}(b(\ell),\rhd_j)^2.$$

This completes the proof as we have shown that the vector $u$ with $u_k = -\operatorname{swap}(b(k),\rhd_i)^2 + \operatorname{swap}(b(k),\rhd_j))^2$ indeed separates $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_j}$. □

For the sake of readability, we assume from now on that

$$u_k^{\rhd_i,\rhd_j} = -\operatorname{swap}(b(k),\rhd_i)^2 + \operatorname{swap}(b(k),\rhd_j)^2$$

for all output rankings $\rhd_i, \rhd_j$ with $\operatorname{swap}(\rhd_i,\rhd_j) = 1$ and all $k \in \{1,\dots,m!\}$. Put differently, this means that the vectors that separate adjacent rankings are described by the score function of the Squared Kemeny rule.

We next aim to extend this insight to pairs of rankings $\rhd_i, \rhd_j$ with larger swap distance. To this end, we proceed as follows: first, we investigate some general consequences of 2RP for the vector $u^{\rhd_i,\rhd_j}$. Next, we start to investigate single-crossing swap sequences. Formally, a *swap sequence* $\rhd_1,\dots,\rhd_k$ (of length $k$) is a sequence of rankings such that $\operatorname{swap}(\rhd_i,\rhd_{i+1}) = 1$ for all $i \in \{1,\dots,k-1\}$. Moreover, this sequence is *single-crossing* if every pair of alternatives is swapped at most once. The motivation for these single-crossing swap sequences is Lemma A.7: we can find a single-crossing swap sequence from $\rhd_i$ to $\rhd_j$ of length $\operatorname{swap}(\rhd_i,\rhd_j)+1$ for any pair of $\rhd_i, \rhd_j$. Finally, we show in Lemmas A.8 to A.10 that $f$ omits for every single-crossing swap sequence a profile such that $f(R)$ contains precisely the rankings in our sequence. Based on this insight, we then show that $u^{\rhd_i,\rhd_j}$ can be represented as $\sum_{k=0}^{t-1} u^{\rhd_{i_k},\rhd_{i_{k+1}}}$ for a single-crossing swap sequence $\rhd_{i_1},\dots,\rhd_{i_t}$ from $\rhd_i$ to $\rhd_j$.

We start this analysis by proving some consequences of 2RP for arbitrary pairs of rankings.

Lemma A.6. *Consider two arbitrary rankings $\rhd_i, \rhd_j \in \mathcal{R}$ such that $\operatorname{swap}(\rhd_i,\rhd_j) \geqslant 1$. For all $b(k) \in \mathcal{R}$, it holds that*

(1) $u_k^{\rhd_i,\rhd_j} = 0$ *if* $\operatorname{swap}(b(k),\rhd_i) = \operatorname{swap}(b(k),\rhd_j)$,
(2) $u_k^{\rhd_i,\rhd_j} \geqslant 0$ *if* $\operatorname{swap}(b(k),\rhd_i) < \operatorname{swap}(b(k),\rhd_j)$,
(3) $u_k^{\rhd_i,\rhd_j} > 0$ *if* $b(k) = \rhd_i$.

Proof. Consider two arbitrary rankings $\rhd_i, \rhd_j \in \mathcal{R}$ with $\text{swap}(\rhd_i, \rhd_j) \geqslant 1$ and let $\succ\, = b(k)$ denote an arbitrary input ranking. Moreover, we define $\bar{\succ}$ as the ranking that is completely inverse to $\succ$. We prove each of our three claims separately.

**Claim 1)**: Assume that $\text{swap}(\succ, \rhd_i) = \text{swap}(\succ, \rhd_j)$. This implies also that

$$\begin{aligned}\text{swap}(\bar{\succ}, \rhd_i) &= \tbinom{m}{2} - \text{swap}(\succ, \rhd_i)\\ &= \tbinom{m}{2} - \text{swap}(\succ, \rhd_j)\\ &= \text{swap}(\bar{\succ}, \rhd_j).\end{aligned}$$

Moreover, since $\text{swap}(\succ, \rhd_i) = \text{swap}(\succ, \rhd_j)$, we can conclude that $\succ\, \neq \rhd_i$ and $\succ\, \neq \rhd_j$. Now, consider the profiles $R$ and $R'$ defined by

$$\begin{aligned} R(\bar{\succ}) &= \frac{\text{swap}(\succ, \rhd_i)}{\binom{m}{2}} & R(\succ) &= \frac{\binom{m}{2} - \text{swap}(\succ, \rhd_i)}{\binom{m}{2}}\\ R'(\bar{\succ}) &= \frac{\text{swap}(\succ, \rhd_i)}{\binom{m}{2}} - \varepsilon & R'(\succ) &= \frac{\binom{m}{2} - \text{swap}(\succ, \rhd_i)}{\binom{m}{2}} + \varepsilon;\end{aligned}$$

$\varepsilon > 0$ denotes here a rational constant such that $\text{round}(\binom{m}{2}(1 - R(\bar{\succ}))) = \{\binom{m}{2} - \text{swap}(\succ, \rhd_i))\}$ and $\text{round}(\binom{m}{2}(1 - R(\succ))) = \{\text{swap}(\succ, \rhd_i))\}$. 2RP thus implies that

$$f(R) = f(R') = \{\rhd \in \mathcal{R} \colon\ \text{swap}(\succ, \rhd) = \text{swap}(\succ, \rhd_i) \text{ and } \text{swap}(\bar{\succ}, \rhd) = \text{swap}(\bar{\succ}, \rhd_i)\}.$$

Consequently, $\rhd_i$ and $\rhd_j$ are both chosen for $R$ and $R'$. Hence, $v(R), v(R') \in \bar{R}_{\rhd_i} \cap \bar{R}_{\rhd_j}$ and $v(R)u^{\rhd_i,\rhd_j} = v(R')u^{\rhd_i,\rhd_j} = 0$ by the definitions of the sets $\bar{R}_{\rhd}$ and the vector $u^{\rhd_i,\rhd_j}$. Finally, let $\lambda = \frac{R'(\bar{\succ})}{R(\bar{\succ})}$. It is now easy to verify that

$$0 = (v(R') - \lambda v(R))u^{\rhd_i,\rhd_j} = (R'(\succ) - \lambda R(\succ))u_k^{\rhd_i,\rhd_j}.$$

Since $R'(\succ) > R(\succ)$ and $0 < \lambda < 1$, this implies that $u_k^{\rhd_i,\rhd_j} = 0$, which completes the proof of this claim.

**Claim 2)**: Next, assume that $\text{swap}(\succ, \rhd_i) < \text{swap}(\succ, \rhd_j)$. In this case, we consider the following two profiles $R$ and $R'$:

$$\begin{aligned} R(\bar{\succ}) &= \frac{\text{swap}(\succ, \rhd_i)}{\binom{m}{2}} & R(\succ) &= \frac{\binom{m}{2} - \text{swap}(\succ, \rhd_i)}{\binom{m}{2}}\\ R'(\bar{\succ}) &= \frac{\text{swap}(\succ, \rhd_j)}{\binom{m}{2}} & R'(\succ) &= \frac{\binom{m}{2} - \text{swap}(\succ, \rhd_j)}{\binom{m}{2}}.\end{aligned}$$

Using 2RP, it can be verified that $\rhd_i \in f(R)$ and $\rhd_j \notin f(R)$, and $\rhd_i \in f(R')$ and $\rhd_j \notin f(R')$. Consequently, $v(R) \in \bar{R}_{\rhd_i}$ and $v(R') \in \bar{R}_{\rhd_j}$, which implies that $v(R)u^{\rhd_i,\rhd_j} \geqslant 0$ and $v(R')u^{\rhd_i,\rhd_j} \leqslant 0$. Finally, we define $\lambda = \frac{R(\bar{\succ})}{R'(\bar{\succ})}$ and note that $0 \leqslant \lambda < 1$. Since $-\lambda v(R')u^{\rhd_i,\rhd_j} \geqslant 0$, we derive that

$$0 \leqslant (v(R) - \lambda v(R'))u^{\rhd_i,\rhd_j} = (R(\succ) - \lambda R'(\succ))u_k^{\rhd_i,\rhd_j}.$$

Since $R(\succ) > R'(\succ)$ (as $\text{swap}(\succ, \rhd_i) < \text{swap}(\succ, \rhd_j)$) and $0 \leqslant \lambda < 1$, this inequality shows that $u_k^{\rhd_i,\rhd_j} \geqslant 0$.

**Claim 3)**: For our last claim, we assume that $\succ\, = \rhd_i$, which means that $\text{swap}(\succ, \rhd_i) = 0$ and $\text{swap}(\succ, \rhd_j) > 0$. By Claim 2), we know that $u_k^{\rhd_i,\rhd_j} \geqslant 0$ and we only need to show that this inequality is strict. For doing so, consider the profile $R$ with $R(\succ) = 1 - \varepsilon$ and $R(\bar{\succ}) = \varepsilon$, where $\varepsilon > 0$ is so small that $\text{round}(\binom{m}{2}(1 - R(\succ))) = \{0\}$. As a consequence, 2RP implies that $f(R) = \{\rhd_i\}$. By Lemma A.4, this means that $v(R) \in \bar{R}_{\rhd_i}$ and $v(R) \notin \bar{R}_{\rhd'}$ for all $\rhd' \in \mathcal{R} \setminus \{\rhd_i\}$. Lemma A.3, in turn, implies that $v(R)u^{\rhd_i,\rhd_j} \geqslant 0$ and that there is for every $\rhd_k \in \mathcal{R} \setminus \{\rhd_i\}$ another ranking $\phi(\rhd_k)$ such that $v(R)u^{\rhd_k,\phi(\rhd_k)} < 0$.

We will next show that $u^{\rhd_i,\rhd_j} > 0$ and assume thus for contradiction that $u^{\rhd_i,\rhd_j} = 0$. Moreover, let $v'$ denote a vector such that $v'u^{\rhd_i,\rhd_j} < 0$; such a vector exists as $u^{\rhd_i,\rhd_j}$ is non-zero. We then define $v^* = v(R) + \delta v'$, where $\delta > 0$ is so small that $v(R)u^{\rhd_1,\rhd_2} < 0$ implies that $v^*u^{\rhd_1,\rhd_2} < 0$ for all $\rhd_1, \rhd_2 \in \mathcal{R}$. In particular, this means that $v^*u^{\rhd_k,\phi(\rhd_k)} < 0$ for all $\rhd_k \in \mathcal{R} \setminus \{\rhd_i\}$, so $v^* \notin \bar{R}_{\rhd_k}$ for these rankings. Moreover, $v^*u^{\rhd_i,\rhd_j} = \delta v'u^{\rhd_i,\rhd_j} < 0$, so $v' \notin \bar{R}_{\rhd_i}$. However, this contradicts that $\bigcup_{\rhd_k \in \mathcal{R}} \bar{R}_{\rhd_k} = \mathbb{R}^{m!}$, so it must hold that $v(R)u^{\rhd_i,\rhd_j} > 0$.

As the last point, we observe that $v(R)u^{\rhd_i,\rhd_j} > 0$ is only possible if $u_k^{\rhd_i,\rhd_j} > 0$ or $u_{k'}^{\rhd_i,\rhd_j} > 0$ ($k'$ denotes here the index such that $\bar{\succ} = b(k')$). Now, assume for contradiction that $u_k^{\rhd_i,\rhd_j} \leqslant 0$, so $u_{k'}^{\rhd_i,\rhd_j} > 0$. By Claim 2), we know that $u_k^{\rhd_i,\rhd_j} \geqslant 0$ and thus $u_k^{\rhd_i,\rhd_j} = 0$. Finally, consider the profile $R'$ in the proof of Claim 2) with $\rhd_j \in f(R')$. In particular, this means that $v(R') \in \bar{R}_{\rhd_j}$. However, if $u_k^{\rhd_i,\rhd_j} = 0$ and $u_{k'}^{\rhd_i,\rhd_j} > 0$, then $v(R')u^{\rhd_i,\rhd_j} > 0$ which contradicts that $v(R') \in \bar{R}_{\rhd_j}$. Hence, the assumption that $u_k^{\rhd_i,\rhd_j} \leqslant 0$ is wrong and our third claim follows. □

As explained before, we now turn our focus to swap sequences. For the sake of completeness, we will next show that for every pair of rankings $\rhd_i, \rhd_j$, there is a swap sequence from $\rhd_i$ to $\rhd_j$ of length $\text{swap}(\rhd_i, \rhd_j) + 1$.

Lemma A.7. *Let $\rhd_i, \rhd_j \in \mathcal{R}$ denote two rankings such that* $\text{swap}(\rhd_i, \rhd_j) = k$ *for some $k \in \mathbb{N}$. There is a swap sequence $\hat{\rhd}_1, \dots, \hat{\rhd}_{k+1}$ of length $k+1$ such that $\hat{\rhd}_1 = \rhd_i$ and $\hat{\rhd}_{k+1} = \rhd_j$.*

Proof. We prove the claim by induction over the swap distance $\text{swap}(\rhd_i, \rhd_j) = k$ for our two considered rankings $\rhd_i, \rhd_j$. For the induction basis, assume that $k = 1$. This means that there is a single pair of alternatives $a, b$ such that $a \rhd_i b$ and $b \rhd_j a$. This is only possible if we can transform $\rhd_i$ to $\rhd_j$ by simply swapping $a$ and $b$ and thus, $\rhd_i, \rhd_j$ forms our swap sequence of length $k + 1$.

Now, suppose that we can construct a swap sequence of length $k$ between any two rankings with $\rhd_i, \rhd_j$ with $\text{swap}(\rhd_i, \rhd_j) = k - 1$. We will prove that the same holds for rankings $\rhd_i, \rhd_j$ with $\text{swap}(\rhd_i, \rhd_j) = k$. To this end, let $D = \{(a, b) \in \rhd_i \setminus \rhd_j\}$ and note that $|D| = k$ by definition. For proving the lemma, let $x_k$ denote the $k$-th best alternative in $\rhd_i$ and $y_k$ the $k$-th best alternative in $\rhd_j$ for every $k \in \{1, \dots, m\}$. Since $\rhd_i \neq \rhd_j$, there is an integer $k$ such that $x_k \neq y_k$ and we let $k^*$ denote the largest such integer. Put differently, this means that $x_k = y_k$ for all $k > k^*$. Consequently, there is an integer $\ell < k^*$ such that $x_\ell = y_{k^*}$. We claim that $(x_\ell, x_{\ell+1}) \in D$. If this was not the case, then $x_\ell \rhd_j x_{\ell+1}$, so there is an index $k' > k^*$ such that $x_{\ell+1} = y_{k'}$. However, this contradicts that $x_{k'} = y_{k'}$, so $(x_\ell, x_{\ell+1}) \in D$ is true.

Next, let $\rhd_i'$ denote the ranking derived from $\rhd_i$ by swapping $x_\ell$ and $x_{\ell+1}$. It holds that $\text{swap}(\rhd_i', \rhd_j) = k - 1$, so there is a swap sequence $\hat{\rhd}_1, \dots, \hat{\rhd}_k$ of length $k$ from $\rhd_i'$ to $\rhd_j$ by the induction hypothesis. Finally, $\rhd_i, \hat{\rhd}_1, \dots, \hat{\rhd}_k$ is then a swap sequence of length $k + 1$ connecting $\rhd_i$ and $\rhd_j$. □

We note that the swap sequences constructed in Lemma A.7 are minimal: for any pair of rankings $\rhd_i, \rhd_j$ with $\text{swap}(\rhd_i, \rhd_j) = k$, there cannot be a swap sequence of length less than $k+1$ that connects these two rankings. In particular, this means that the constructed sequences are single-crossing.

We next aim to show that $f$ admits for every single-crossing swap sequence $\rhd_{i_0}, \dots, \rhd_{i_t}$ a profile $R$ such that $f(R) = \{\rhd_{i_0}, \dots, \rhd_{i_t}\}$. To this end, we analyze the linear independence of the vectors $u^{\rhd_{i_k},\rhd_{i_{k+1}}}$ for $k \in \{0, \dots, t-1\}$. Since all these vectors separate rankings that differ only in a single swap, Lemma A.5 applies and shows that they can be described by the scores assigned by the Squared Kemeny rule. Hence, we will first analyze the Squared Kemeny rule in more detail. In particular, we will show that the Squared Kemeny rule admits for every single-crossing swap sequence $\rhd_0, \dots, \rhd_t$ a profile $R$ such that $\text{SqK}(R) = \{\rhd_0, \dots, \rhd_t\}$.

Lemma A.8. *Let $\rhd_0, \dots, \rhd_t$ denote a single-crossing swap sequence. There is a profile $R$ such that* $\mathrm{SqK}(R) = \{\rhd_0, \dots, \rhd_t\}$.

Proof. We will first show the lemma for a single-crossing swap sequence $\rhd_0, \dots, \rhd_t$ of length $t+1 = \binom{m}{2} + 1$. To this end, we introduce some auxiliary notation: for every ranking $\succ$, we define $R^{\succ}$ as the profile with $R^{\succ}(\succ) = 1$ and $R^{-\succ}$ as the profile with $R^{-\succ}(\succ') = \frac{1}{m!-1}$ for all $\succ' \in \mathcal{R} \setminus \{\succ\}$. Moreover, we recall that $C_{\mathrm{SqK}}(R, \rhd) = \sum_{\succ \in \mathcal{R}} R(\succ)\, \mathrm{swap}(\succ, \rhd)^2$ denotes the Squared Kemeny cost of a ranking $\rhd$ in the profile $R$. It immediately follows that $C_{\mathrm{SqK}}(R^{\succ}, \rhd) = \mathrm{swap}(\succ, \rhd)^2$ for all $\succ, \rhd \in \mathcal{R}$. Next, for determining $C_{\mathrm{SqK}}(R^{-\succ}, \rhd)$, we first consider the profile $R$ with $R(\succ) = \frac{1}{m!}$ for all $\succ \in \mathcal{R}$. Due to the symmetry of this profile, it holds that $C_{\mathrm{SqK}}(R, \rhd) = C_{\mathrm{SqK}}(R, \rhd')$ for all $\rhd \in \mathcal{R}$ and we thus define $c = m! \cdot C_{\mathrm{SqK}}(R, \rhd) = \sum_{\succ \in \mathcal{R}} \mathrm{swap}(\succ, \rhd)^2$. We then compute that $(m!-1) C_{\mathrm{SqK}}(R^{-\succ}, \rhd) = \sum_{\succ' \in \mathcal{R}} \mathrm{swap}(\succ', \rhd)^2 - \mathrm{swap}(\succ, \rhd)^2 = c - \mathrm{swap}(\succ, \rhd)^2$ for all $\succ, \rhd \in \mathcal{R}$.

Based on these insights, we now define the profile $R^*$ as convex combination of profiles $R^{\succ}$ and $R^{-\succ}$ for $\succ \in \{\rhd_0, \dots, \rhd_t\}$. In more detail, if $t$ is odd, then

$$R^* = \frac{1}{Z}\Big((m!-1)\frac{t-1}{2}(R^{-\rhd_0} + R^{-\rhd_t}) + \sum_{k=1}^{t-1} R^{\rhd_k}\Big),$$

and if $t$ is even, then

$$R^* = \frac{1}{Z}\Big((m!-1)\frac{t}{2}(R^{-\rhd_0} + R^{-\rhd_t}) + R^{\rhd_{t/2}} + \sum_{k=1}^{t-1} R^{\rhd_k}\Big).$$

For both cases, $Z$ denotes a normalization constant such that $\sum_{\succ \in \mathcal{R}} R^*(\succ) = 1$. Based on our previous insights, we can compute for all rankings $\rhd$ that

$$Z \cdot C_{\mathrm{SqK}}(R^*, \rhd) = (t-1)c - \frac{t-1}{2}(\mathrm{swap}(\rhd_0, \rhd)^2 + \mathrm{swap}(\rhd_t, \rhd)^2) + \sum_{k=1}^{t-1} \mathrm{swap}(\rhd_k, \rhd)^2$$

if $t$ is odd, and

$$Z \cdot C_{\mathrm{SqK}}(R^*, \rhd) = tc - \frac{t}{2}(\mathrm{swap}(\rhd_0, \rhd)^2 + \mathrm{swap}(\rhd_t, \rhd)^2) + \mathrm{swap}(\rhd_{t/2}, \rhd)^2 + \sum_{k=1}^{t-1} \mathrm{swap}(\rhd_k, \rhd)^2$$

if $t$ is even.

We will next show that $\mathrm{SqK}(R^*) = \{\rhd_0, \dots, \rhd_t\}$. As a first point, we will show that $f(R^*) \subseteq \{\rhd_0, \dots, \rhd_t\}$. For this, let $\rhd$ denote a ranking that is not on our swap sequence. In particular, this means that $d = \mathrm{swap}(\rhd_0, \rhd) > 0$ and $\mathrm{swap}(\rhd_t, \rhd) > 0$. Moreover, since our swap sequence $\rhd_0, \dots, \rhd_t$ has maximal length, we know that $\rhd_0$ and $\rhd_t$ are completely inverse. This implies that $\mathrm{swap}(\rhd_t, \rhd) = \binom{m}{2} - d$, so $d < \binom{m}{2}$. Now, let $\rhd_d$ denote the ranking in our swap sequence such that $\mathrm{swap}(\rhd_0, \rhd_d) = d$ and $\mathrm{swap}(\rhd_t, \rhd_d) = \binom{m}{2} - d$; we will show that $Z \cdot C_{\mathrm{SqK}}(R^*, \rhd_d) < Z \cdot C_{\mathrm{SqK}}(R^*, \rhd)$. We hence note that $\mathrm{swap}(\rhd_0, \rhd_d) = \mathrm{swap}(\rhd_0, \rhd)$ and $\mathrm{swap}(\rhd_t, \rhd_d) = \mathrm{swap}(\rhd_t, \rhd)$, so the cost caused by $\rhd_0$ and $\rhd_t$ does not matter for comparing these rankings. Moreover, it is obvious that $\mathrm{swap}(\rhd_d, \rhd_d) = 0 < \mathrm{swap}(\rhd_d, \rhd)$. Further, consider an arbitrary ranking $\rhd_k$ in our swap sequence with $k \in \{1, \dots, d-1\}$. It holds that $\mathrm{swap}(\rhd_k, \rhd) + \mathrm{swap}(\rhd_k, \rhd_0) \geqslant \mathrm{swap}(\rhd_0, \rhd)$, so we can infer that $\mathrm{swap}(\rhd_k, \rhd) \geqslant d - k = \mathrm{swap}(\rhd_k, \rhd_d)$. Since a symmetric argument holds for all $\rhd_k$ with $k \in \{d+1, \dots, t-1\}$, we can now conclude that $\sum_{k=1}^{t-1} \mathrm{swap}(\rhd_k, \rhd)^2 > \sum_{k=1}^{t-1} \mathrm{swap}(\rhd_k, \rhd_d)^2$, which implies that $\rhd \notin \mathrm{SqK}(R)$.

Next, we need to show that $C_{\mathrm{SqK}}(R, \rhd) = C_{\mathrm{SqK}}(R, \rhd')$ for all $\rhd, \rhd' \in \{\rhd_0, \dots, \rhd_t\}$. For this, we define $\ell = \frac{t}{2} - 1$ if $t$ is even and $\ell = \frac{t-3}{2}$ if $t$ is odd. Moreover, we consider the auxiliary profiles $R^i$ (for $i \in \{0, \dots, \ell\}$) defined by

$$R^i = \frac{1}{Z^i}\big((m!-1)(R^{-\rhd_i} + R^{-\rhd_{t-i}}) + R^{\rhd_{i+1}} + R^{\rhd_{t-(i+1)}}\big),$$

where $Z^i$ is again a normalization constant. Now, we recall that $(m!-1)R^{-\rhd} + R^{\rhd} = m!R$ for all $\rhd$, where $R$ is the profile in which every ranking has weight $\frac{1}{m!}$. Hence, we infer that

$$\begin{aligned}\sum_{i=0}^{\ell}(\ell+1-i)\cdot Z^i\cdot R^i &= \sum_{i=0}^{\ell}(\ell+1-i)\cdot\big((m!-1)(R^{-\rhd_i}+R^{-\rhd_{t-i}}) + R^{\rhd_{i+1}} + R^{\rhd_{t-(i+1)}}\big)\\ &= \sum_{i=1}^{\ell}(\ell+1-i)(m!-1)(R^{-i})(R^{-\rhd_i}+R^{-\rhd_{t-i}}) + (\ell+2-i)(R^{\rhd_i}+R^{\rhd_{t-i}})\\ &\quad + (\ell+1)(m!-1)(R^{-\rhd_0}+R^{-\rhd_t})\\ &= \sum_{i=1}^{\ell} R^{\rhd_i} + R^{\rhd_{t-i}} + 2(\ell+1-i)m!R + (\ell+1)(m!-1)(R^{-\rhd_0}+R^{-\rhd_t})\\ &= ZR^* + \ell' m!R.\end{aligned}$$

Here, we define $\ell' = 2\sum_{i=1}^{\ell}(\ell+1-i)$ for simplicity. Since $C_{\text{SqK}}(R,\rhd) = C_{\text{SqK}}(R,\rhd')$ for all rankings $\rhd,\rhd'$, we infer from this equation that $C_{\text{SqK}}(R^*,\rhd) = C_{\text{SqK}}(R^*,\rhd')$ if and only if $C_{\text{SqK}}(\sum_{i=0}^{\ell}(\ell+1-i)\cdot Z^i\cdot R^i,\rhd) = C_{\text{SqK}}(\sum_{i=0}^{\ell}(\ell+1-i)\cdot Z^i\cdot R^i,\rhd')$ for all rankings $\rhd,\rhd'$. In turn, the latter equality holds if $C_{\text{SqK}}(R^i,\rhd) = C_{\text{SqK}}(R^i,\rhd')$ for all $i\in\{0,\dots,\ell\}$. We thus consider now an arbitrary ranking $\rhd_d\in\{\rhd_0,\dots,\rhd_t\}$ and such a profile $R^i$. First, we note that $\text{swap}(\rhd_j,\rhd_d)^2 = (d-j)^2$ for every ranking $\rhd_j$ in our swap sequence. Hence, we compute that

$$\begin{aligned}&(m-1)!C_{\text{SqK}}(R^{-\rhd_i},\rhd_d) = c - \text{swap}(\rhd_i,\rhd_d)^2 = c-(d-i)^2\\ &(m-1)!C_{\text{SqK}}(R^{-\rhd_{t-i}},\rhd_d) = c - \text{swap}(\rhd_{t-i},\rhd_d)^2 = c-(d-(t-i))^2\\ &C_{\text{SqK}}(R^{\rhd_{i+1}},\rhd_d) = \text{swap}(\rhd_{i+1},\rhd_d) = (d-(i+1))^2\\ &C_{\text{SqK}}(R^{\rhd_{t-(i+1)}},\rhd_d) = \text{swap}(\rhd_{t-(i+1)},\rhd_d) = (d-(t-(i+1)))^2\end{aligned}$$

Our central observation is now that

$$\begin{aligned}&-(d-i)^2-(d-(t-i))^2+(d-(i+1))^2+(d-(t-(i+1)))^2\\ &=-(d-i)^2-(d-(t-i))^2+(d-i-1)^2+(d-(t-i)+1)^2\\ &=-(d-i)^2-(d-(t-i))^2+(d-i)^2-2(d-i)+1+(d-(t-i))^2+2(d-(t-i))+1\\ &=2+4i-2t.\end{aligned}$$

Hence, $C_{\text{SqK}}(R^i,\rhd^d) = \frac{1}{Z^i}(2c+2+4i-2t)$ for all $\rhd_d\in\{\rhd_0,\dots,\rhd_t\}$ and all $i\in\{0,\dots,\ell\}$, so we can conclude that $\text{SqK}(R^*) = \{\rhd_0,\dots,\rhd_t\}$.

As the last point, we need to extend our argument to swap sequence of length $t'+1 < t+1 = \binom{m}{2}+1$. Hence, consider a single-crossing swap sequence $\rhd_0,\dots,\rhd_{t'}$ for $t'<t$. First, if $t'=0$, then there is clearly a profile $R$ such that $\text{SqK}(R) = \{\rhd_0\}$ due to 2RP. We hence suppose that $t'\geqslant 1$. We can extend this sequence to a single-crossing swap sequence $\rhd'_0,\dots,\rhd'_t$ of length $t+1$ using Lemma A.7. This means that $\rhd'_i = \rhd_i$ for all $i\in\{0,\dots,t'\}$ and the remaining rankings form a path to $\rhd'_t$ (which is completely inverse to $\rhd_0$). Now, let $\bar{\rhd}$ denote the completely inverse ranking of every $\rhd$ (i.e., $\rhd'_t = \bar{\rhd}_0$). It is easy to see that

$$\bar{\rhd}'_{t-i},\bar{\rhd}'_{t-i+1},\dots,\bar{\rhd}'_{t-1},\bar{\rhd}'_t = \rhd'_0,\rhd'_1,\dots,\rhd'_i$$

is a single-crossing swap sequence for every $i$. Moreover, our construction yields for every $i$ a profile $\hat{R}^i$ such that the Squared Kemeny rule chooses exactly the given swap sequence. Now, consider the profile $\hat{R}^0$ (for which $\mathrm{SqK}(\hat{R}^0) = \{\rhd'_0, \dots, \rhd'_t\}$) and the profile $\hat{R}^{t-t'}$ (for which $\mathrm{SqK}(\hat{R}^{t-t'}) = \{\bar{\rhd}'_{t-t'}, \dots \bar{\rhd}'_t, \rhd'_1, \dots, \rhd'_{t'}\}$). It can be verified that $\mathrm{SqK}(\hat{R}^0) \cap \mathrm{SqK}(\hat{R}^{t-t'}) = \{\rhd'_0, \dots, \rhd'_{t'}\}$, so $\mathrm{SqK}(\frac{1}{2}\hat{R}^0 + \frac{1}{2}\hat{R}^{t-t'}) = \{\rhd'_0, \dots, \rhd'_{t'}\}$ as the Squared Kemeny rule is reinforcing. □

Based on Lemma A.8, we next show that the vectors $u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_{t-1}, \rhd_t}$ are linearly independent for every swap sequence $\rhd_0, \dots, \rhd_t$.

Lemma A.9. *Consider a single-crossing swap sequence $\rhd_0, \dots, \rhd_t$. The vectors $u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_{t-1}, \rhd_t}$ are linearly independent.*

Proof. Let $\rhd_0, \dots, \rhd_t$ denote a single-crossing swap sequence and let $u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_{t-1}, \rhd_t}$ denote the vectors that separate $\bar{R}_{\rhd_i}$ from $\bar{R}_{\rhd_{i+1}}$. Since any two consecutive rankings in our sequence only differ in a swap, we know that $u_k^{\rhd_i, \rhd_{i+1}} = -\mathrm{swap}(\rhd_i, b(k))^2 + \mathrm{swap}(\rhd_{i+1}, b(k))^2$ for all $i \in \{0, \dots, t-1\}$ and $k \in \{1, \dots, m!\}$ (see Lemma A.5). In particular, this means that $v(R)u^{\rhd_i, \rhd_{i+1}} = 0$ for every profile $R$ in which $\rhd_i$ and $\rhd_{i+1}$ have the same Squared Kemeny score. Now, by Lemma A.8, there are profiles $R^j$ such that $\mathrm{SqK}(R^j) = \{\rhd_0, \dots, \rhd_j\}$ for all $j \in \{1, \dots, t\}$. By the definition of the Squared Kemeny rule, this means that $v(R^j)u^{\rhd_i, \rhd_{i+1}} = 0$ for all $i < j$ and $v(R^j)u^{\rhd_j, \rhd_{j+1}} > 0$.

We will now use these profiles to inductively show that the considered vectors are linearly independent. For the induction basis, let $j = 1$ and note that the set $\{u^{\rhd_0, \rhd_1}\}$ is trivially linearly independent. Now, assume that the set $\{u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_{j-1}, \rhd_j}\}$ is linearly independent for some $j \leqslant t - 1$. For the set $\{u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_j, \rhd_{j+1}}\}$, the linear independence follows by considering the vector $v(R^j)$ because $v(R^j)u^{\rhd_i, \rhd_{i+1}} = 0$ for all $i < j$ and $v(R^j)u^{\rhd_j, \rhd_{j+1}} > 0$. This is only possible if $u^{\rhd_j, \rhd_{j+1}}$ is linearly independent of the remaining vectors in our set. Moreover, since the set $\{u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_{j-1}, \rhd_j}\}$ is linearly independent by the induction hypothesis, the full set $\{u^{\rhd_0, \rhd_1}, \dots, u^{\rhd_j, \rhd_{j+1}}\}$ is linearly independent and the lemma follows. □

As the next step, we show that Lemma A.8 also holds for our SPF $f$: for every single-crossing swap sequence $\rhd_0, \dots, \rhd_t$, there is a profile $R$ such that $f(R) = \{\rhd_0, \dots, \rhd_t\}$. As it will turn out, the same profiles as for the Squared Kemeny rule show this claim.

Lemma A.10. *Let $\rhd_0, \dots, \rhd_t$ denote a single-crossing swap sequence. There is a profile $R$ such that $f(R) = \{\rhd_0, \dots, \rhd_t\}$.*

Proof. We will prove the lemma only for single-crossing swap sequences $\rhd_0, \dots, \rhd_t$ with $t = \binom{m}{2}$; since $f$ satisfies 2RP and reinforcement, we can shorten the sequence as demonstrated in Lemma A.8. Hence, consider such a sequence, and let $R^*$ denote the corresponding profile defined in Lemma A.8.

We will first show that $f(R^*) \subseteq \{\rhd_0, \dots, \rhd_t\}$. To this end, consider an arbitrary ranking $\rhd$ not in our sequence. Moreover, we define $d = \mathrm{swap}(\rhd, \rhd_0)$ and note that $0 < d < \binom{m}{2}$ as $d \notin \{\rhd_0, \rhd_t\}$. Now, let $\rhd_d$ denote the $d$-th ranking in our sequence. We will show that $vu^{\rhd_d, \rhd} > 0$ for the vector $v = v(R^*)$. This shows that $v \notin \bar{R}_\rhd$ which, in turn, implies that $\rhd \notin \hat{g}(v) = f(R^*)$ due to Lemma A.4.

To this end, we first note that $\mathrm{swap}(\rhd_0, \rhd_d) = d = \mathrm{swap}(\rhd_0, \rhd)$ and $\mathrm{swap}(\rhd_t, \rhd_d) = t - d = \mathrm{swap}(\rhd_t, \rhd)$. Hence, it holds by Claim 1) of Lemma A.6 that $u_k^{\rhd_d, \rhd} = u_{k'}^{\rhd_d, \rhd} = 0$ for the indices $k, k'$ with $b(k) = \rhd_0$ and $b(k') = \rhd_t$. Next, let $v'$ denote the vector with $v'_i = \frac{1}{m!-1}$ for all $i \in \{1, \dots, m!\}$ and note that $v' \in \bar{R}_{\rhd'}$ for all $\rhd' \in \mathcal{R}$ due to the symmetry of this vector. Hence, $v'u^{\rhd_d, \rhd} = 0$. We can now compute that $v(R^{-\rhd_0})u^{\rhd_d, \rhd} = v'u^{\rhd_d, \rhd} - \frac{1}{m!-1}u_k^{\rhd_d, \rhd} = 0$ and $v(R^{-\rhd_t})u^{\rhd_d, \rhd} = v'u^{\rhd_d, \rhd} - \frac{1}{m!-1}u_{k'}^{\rhd_d, \rhd} = 0$.

Moreover, an analogous analysis as in the proof of Lemma A.8 shows that $\mathrm{swap}(\rhd_i, \rhd_d) \leqslant \mathrm{swap}(\rhd_i, \rhd)$ for every ranking $\rhd_i$ in our swap sequence. By Claims 1) and 2) in Lemma A.6, this means that $u_k^{\rhd_d,\rhd} \geqslant 0$ for the index $k$ with $b(k) = \rhd_i$. Finally, $R^*(\rhd_d) > 0$ and we know by Claim 3) in Lemma A.6 that $u_k^{\rhd_d,\rhd} > 0$ for the corresponding index. Since $v(R^{\rhd_i})u^{\rhd_d,\rhd} = u_k^{\rhd_d,\rhd_k}$ for all $\rhd_i$ in our swap sequence and the corresponding index $k = b(\rhd_i)$, we can now infer that $vu^{\rhd_d,\rhd} > 0$ as $R^*$ is a convex combination of $R^{-\rhd_0}$, $R^{-\rhd_t}$, and all $R^{\rhd_i}$ for $i \in \{1, \dots, t-1\}$. This proves that $f(R^{\rhd_0,\dots,\rhd_t}) \subseteq \{\rhd_0, \dots, \rhd_t\}$.

Next, we need to show that $f(R^*) = \{\rhd_0, \dots, \rhd_t\}$. Assume for a contradiction that this is not the case, i.e., there is a ranking $\rhd_i$ in our sequence that is not chosen. By Lemma A.4, this means that $v \notin \bar{R}_{\rhd_i}$. Moreover, by combining Lemmas A.3 and A.4, we know that, for every $\rhd_j \notin f(R^*)$, there is another ranking $\phi(\rhd_j)$ such that $vu^{\rhd_j,\phi(\rhd_j)} < 0$. Next, we note that $\mathrm{SqK}(R^*) = \{\rhd_0, \dots, \rhd_t\}$ because of Lemma A.8. Combined with Lemma A.5, this implies that $vu^{\rhd_j,\rhd_{j+1}} = 0$ for all $j \in \{0, \dots, t-1\}$. Finally, all these vectors $u^{\rhd_j,\rhd_{j+1}}$ are linearly independent of each other (see Lemma A.9). Hence, the matrix $M$ that contains these vectors as rows has full (row) rank, which equivalently means that its image has full dimension. As a consequence, there is a vector $v' \in \mathbb{R}^{m!}$ such that

$$v'u^{\rhd_j,\rhd_{j+1}} < 0 \qquad \text{for all } j \in \{0, \dots, i-1\}, \text{ and}$$
$$v'u^{\rhd_j,\rhd_{j+1}} > 0 \qquad \text{for all } j \in \{i, \dots, t-1\}.$$

Finally, we consider the vector $v + \varepsilon v'$, where $\varepsilon > 0$ is so small that $(v + \varepsilon v')u^{\rhd_j,\phi(\rhd_j)} < 0$ still holds for all $\rhd_j \notin f(R^*)$. This shows that $v + \varepsilon v' \notin \bar{R}_{\rhd_j}$ for any $\rhd_j \notin f(R^*)$. Next, it holds that

$$(v + \varepsilon v')u^{\rhd_j,\rhd_{j+1}} = \varepsilon v'u^{\rhd_j,\rhd_{j+1}} < 0 \qquad \text{for all } j \in \{0, \dots, i-1\}, \text{ and}$$
$$(v + \varepsilon v')u^{\rhd_j,\rhd_{j+1}} = \varepsilon v'u^{\rhd_j,\rhd_{j+1}} > 0 \qquad \text{for all } j \in \{i, \dots, t-1\}.$$

Hence, we also have that $v + \varepsilon v' \notin \bar{R}_{\rhd_j}$ for all $\rhd_j \in \{\rhd_0, \dots, \rhd_t\} \setminus \{\rhd_i\}$. However, since $\rhd_i \notin f(R^*)$ by assumption, this means that $v + \varepsilon v' \notin \bar{R}_{\rhd_j}$ for every $\rhd_j \in \mathcal{R}$. This contradicts that $\bigcup_{\rhd_j \in \mathcal{R}} \bar{R}_{\rhd_j} = \mathbb{R}^{m!}$, so our initial assumption that $f(R^*) \subsetneq \{\rhd_0, \dots, \rhd_t\}$ must have been wrong and $f(R^*) = \{\rhd_1, \dots, \rhd_t\}$. □

We are now ready to fully generalize Lemma A.5 to all vectors $u^{\rhd_i,\rhd_j}$. We note that the subsequent lemma is the equivalent of Step 3 in the proof of Theorem 3.3.

Lemma A.11. *Consider a single-crossing swap sequence $\rhd_0, \dots, \rhd_t$ for some $t \geqslant 1$. There is $\lambda > 0$ such that $u_k^{\rhd_0,\rhd_t} = \lambda(-\operatorname{swap}(\rhd_0, b(k))^2 + \operatorname{swap}(\rhd_k, b(k))^2)$ for all $k \in \{1, \dots, m!\}$.*

Proof. First, we note that the lemma follows immediately from Lemma A.5 if $t = 1$, so we focus subsequently on the case that $t \geqslant 2$. We denote by $\rhd_0, \dots, \rhd_t$ a given single-crossing swap sequence and prove the lemma in multiple steps. In particular, we first show the vector $u^{\rhd_0,\rhd_t}$ is linearly dependent on the vectors $u^{\rhd_0,\rhd_1}, \dots, u^{\rhd_{t-1},\rhd_t}$, which means that there are scalars $\lambda_i$, not all of which are 0, such that $u^{\rhd_0,\rhd_t} = \sum_{i=0}^{t-1} \lambda_i u^{\rhd_i,\rhd_{i+1}}$. The lemma now follows by showing that all scalars $\lambda_i$ are non-negative and equal since not all of them are 0. We thus prove in the second step all $\lambda_i$ are non-negative. In the third step, we then prove the lemma for the case that $t = 2$ and finally generalize the lemma to arbitrary $t$ in the last step.

**Step 1:** As first step, we show that $u^{\rhd_0,\rhd_t}$ is linearly dependent on $u^{\rhd_0,\rhd_1}, \dots, u^{\rhd_{t-1},\rhd_t}$. Assume for contradiction that this is not the case, which means that the set $\{u^{\rhd_0,\rhd_1}, \dots, u^{\rhd_{t-1},\rhd_t}, u^{\rhd_0,\rhd_t}\}$ is linearly independent. We consider now the matrix $M$ that contains all these vectors as rows. By basic linear algebra, this matrix has full (row) rank, so its image has full dimension. This implies that there is a vector $v'$ such that $v'u^{\rhd_0,\rhd_t} < 0$ and $vu^{\rhd_i,\rhd_{i+1}} > 0$ for all $i \in \{0, \dots, t-1\}$. By the definition of these vectors, this means that $v' \notin \bar{R}_{\rhd_i}$ for any $\rhd_i \in \{\rhd_0, \dots, \rhd_t\}$.

Moreover, by Lemma A.10, there is a profile $R$ such that $f(R) = \{\rhd_0, \dots, \rhd_t\}$. Next, by Lemma A.4, it follows for $v = v(R)$ that $v \in \bar{R}_{\rhd_i}$ if and only if $\rhd_i \in \{\rhd_0, \dots, \rhd_t\}$. By Lemma A.3, this means that there is a mapping $\phi$ from $\mathcal{R} \setminus f(R)$ to $\mathcal{R}$ such that $vu^{\rhd,\phi(\rhd)} < 0$ for all $\rhd \in \mathcal{R} \setminus f(R)$. Moreover, it holds that $v \in \bar{R}_{\rhd_i}$ for all $\rhd_i \in f(R)$. By the definition of the vectors $u^{\rhd_i,\rhd_j}$, it hence follows that $vu^{\rhd_i,\rhd_j} = 0$ for all $\rhd_i, \rhd_j \in \{\rhd_0, \dots, \rhd_t\}$.

Finally, we can find as sufficiently small $\varepsilon > 0$ such that $(v + \varepsilon v')u^{\rhd,\phi(\rhd)} < 0$ still holds for every $\rhd \in \mathcal{R} \setminus f(R)$. It is also straightforward to verify that $(v + \varepsilon v')u^{\rhd_i,\rhd_j} = \varepsilon v' u^{\rhd_i,\rhd_j}$ for all $\rhd_i, \rhd_j \in \{\rhd_0, \dots, \rhd_t\}$. By the definition of $v'$, this means that $(v + \varepsilon v') \notin \bar{R}_{\rhd_i}$ for any $\rhd_i \in \mathcal{R}$, which contradicts that $\bigcup_{\rhd_i \in \mathcal{R}} \bar{R}_{\rhd_i} = \mathbb{R}^{m!}$. Hence, the initial assumption is wrong and $u^{\rhd_0,\rhd_t}$ is linearly depending on $u^{\rhd_0,\rhd_1}, \dots, u^{\rhd_{t-1},\rhd_t}$.

**Step 2:** We next show that all $\lambda_i$ are non-negative. Assume for contradiction that this is not true, i.e., there is an index $i$ with $\lambda_i < 0$. Now, recall that the vectors $u^{\rhd_i,\rhd_{i+1}}$ for $i \in \{0, \dots, t-1\}$ are linearly independent, so there is a vector $v'$ such that

$$v' u^{\rhd_j,\rhd_{j+1}} = \begin{cases} \varepsilon & \text{for all } j \in \{0, \dots, t-1\} \setminus \{i\}, \\ 1 & \text{for } j = i. \end{cases}$$

In particular, we can choose $\varepsilon > 0$ so small that

$$-\lambda_i v' u^{\rhd_i,\rhd_{i+1}} > \textstyle\sum_{j=0, j\neq i}^{t-1} \lambda_j v' u^{\rhd_j,\rhd_{j+1}}.$$

This means that $vu^{\rhd_0,\rhd_t} < 0$. However, it follows now that $v' \notin \bar{R}_{\rhd_i}$ for all $\rhd_i \in \{\rhd_0, \dots, \rhd_t\}$. Analogous to the last step, we can combine $v'$ again with the vector $v \in \mathbb{Q}^{m!}$ with $\hat{g}(v) = \{\rhd_0, \dots, \rhd_t\}$ to derive a vector $v + \varepsilon' v'$ such that $v + \varepsilon v' \notin \bar{R}_{\rhd_i}$ for all $\rhd_i \in \mathcal{R}$. This gives the same contradiction as in the last step, so $\lambda_i \geqslant 0$ for all $i \in \{0, \dots, t-1\}$.

**Step 3:** In our third step, we prove the lemma for the case that $t = 2$ and hence let $\rhd_0, \rhd_1, \rhd_2$ denote the considered sequence. By the first two steps, we know that $u^{\rhd_0,\rhd_2} = \lambda_0 u^{\rhd_0,\rhd_1} + \lambda_1 u^{\rhd_1,\succ_2}$ for some values $\lambda_0, \lambda_1$ such that both are non-negative and at least one is strictly positive. We hence only need to show that $\lambda_0 = \lambda_1$. To this end, we note that $\rhd_0$ differs from $\rhd_2$ either in two disjoint swaps, or we shift an alternative $x$ by two positions.

We continue with a case distinction with respect to these two options and first assume that $\rhd_0$ differs from $\rhd_2$ in two disjoint swaps. In this case, let $\succ_1$ and $\succ_2$ denote the rankings where we have swapped precisely one of these two pairs. In more detail, we assume subsequently that

$$\begin{aligned} \rhd_0 &= \dots, a, b, \dots, c, d, \dots \\ \rhd_2 &= \dots, b, a, \dots, d, c, \dots \\ \succ_1 &= \dots, b, a, \dots, c, d \dots \\ \succ_2 &= \dots, a, b, \dots, d, c, \dots \end{aligned}$$

for some alternatives $a, b, c, d$. It is easy to check that $\text{swap}(\rhd_0, \succ_1) = \text{swap}(\rhd_0, \succ_2) = 1$, $\text{swap}(\rhd_2, \succ_1) = \text{swap}(\rhd_2, \succ_2) = 1$, and either $\text{swap}(\rhd_1, \succ_1) = 0$ and $\text{swap}(\rhd_1, \succ_2) = 2$, or $\text{swap}(\rhd_1, \succ_1) = 2$ and $\text{swap}(\rhd_1, \succ_2) = 0$. Next, it follows from 2RP that $f(R) = \{\rhd_0, \rhd_2\}$ for the profile $R$ with $R(\succ_1) = \frac{1}{2}$ and $R(\succ_2) = \frac{1}{2}$. Moreover, using Lemma A.5, we derive that

$$\begin{aligned} v(R)u^{\rhd_0,\rhd_1} &= \frac{1}{2}(-\text{swap}(\rhd_0, \succ_1)^2 + \text{swap}(\rhd_1, \succ_1)^2 - \text{swap}(\rhd_0, \succ_2)^2 + \text{swap}(\rhd_1, \succ_2)^2) \\ &= \frac{1}{2}(-2 + 4) \\ &= 1 \end{aligned}$$

and

$$\begin{aligned}v(R)u^{\rhd_1,\rhd_2} &= \frac{1}{2}(-\operatorname{swap}(\rhd_1,\succ_1)^2+\operatorname{swap}(\rhd_2,\succ_1)^2-\operatorname{swap}(\rhd_1,\succ_2)^2+\operatorname{swap}(\rhd_2,\succ_2)^2)\\ &= \frac{1}{2}(-4+2)\\ &= -1.\end{aligned}$$

Finally, since $\rhd_0, \rhd_2 \in f(R) = \hat{g}(v)$, we have that $v(R)u^{\rhd_0,\rhd_2} = \lambda_0 v u^{\rhd_0,\rhd_1} + \lambda_1 v u^{\rhd_1,\rhd_2} = \lambda_0 - \lambda_1 = 0$. This is only true if $\lambda_0 = \lambda_1$ which proves our third step in this case.

Next, suppose that we derive $\rhd_2$ from $\rhd_0$ by shifting an alternative $a$ by two positions. We assume that $\rhd_2$ is derived from $\rhd_0$ by shifting an alternative $a$ down as the case of shifting an alternative up is symmetric. Hence, let

$$\begin{aligned}\rhd_0 &= \dots, a, b, c, \dots\\ \rhd_1 &= \dots, b, a, c, \dots\\ \rhd_2 &= \dots, b, c, a, \dots.\end{aligned}$$

Now, let $\succ = \dots, c, a, b, \dots$ and consider the profile $R$ such that $R(\rhd_0) = R(\rhd_2) = R(\succ) = \frac{1}{3}$. First, we note that $v(R) \notin \bar{R}_{\rhd}$ for any $\rhd$ that differs from $\rhd_0$ in any other pair as $(a, b)$, $(b, c)$, and $(a, c)$. Indeed, if such a pair exists in $\rhd$, there is also a pair $(x', y')$ such that $x' \rhd y'$, $y' \rhd_0 x'$, and $x'$ and $y'$ are adjacent in $\rhd$. Using Lemma A.5, we can then infer that $v(R)u^{\rhd,\rhd'} < 0$ for the ranking $\rhd'$ in which we swapped $x'$ and $y'$, which implies that $v(R) \notin \bar{R}_{\rhd}$. This is true because the rankings $\rhd_0$, $\rhd_2$, and $\succ$ agree on the order of all alternatives except $a, b, c$ and thus $\operatorname{swap}(\rhd', \succ') < \operatorname{swap}(\rhd, \succ')$ for all $\succ' \in \{\rhd_0, \rhd_2, \succ\}$.

Furthermore, it holds for the ranking $\rhd_1$ that $v(R)u^{\rhd_0,\rhd_1} > 0$ as we can simply compare the Squared Kemeny costs of both rankings ($\rhd_0$ has a cost of $\frac{8}{3}$ in $R$, $\rhd_{i_1}$ of $\frac{11}{3}$). Analogously, we can infer that $v(R)u^{\rhd_2,\rhd_2'} > 0$ for the ranking $\rhd_2' = \dots, c, b, a, \dots$, and that $v(R)u^{\succ,\succ'} > 0$ for the ranking $\succ' = \dots, a, c, b, \dots$. Due to Lemmas A.3 and A.4 and neutrality, we can now infer that $f(R) = g(v(R)) = \{\rhd_0, \rhd_2, \succ\}$. This means that $v(R)u^{\rhd_0,\rhd_2} = 0$. Furthermore, we can compute that

$$\begin{aligned}v(R)u^{\rhd_0,\rhd_1} =& \frac{1}{3}(-\operatorname{swap}(\rhd_0,\rhd_0)^2-\operatorname{swap}(\rhd_0,\succ)^2-\operatorname{swap}(\rhd_0,\rhd_2)^2)\\ &+\frac{1}{3}(\operatorname{swap}(\rhd_1,\rhd_0)^2+\operatorname{swap}(\rhd_1,\succ)^2+\operatorname{swap}(\rhd_1,\rhd_2)^2)\\ =& \frac{1}{3}(-0-4-4+1+1+9) = 1,\end{aligned}$$

and an analogous calculation shows that $v(R)u^{\rhd_1,\rhd_2} = -1$. Since

$$v(R)u^{\rhd_0,\rhd_2} = \lambda_0 v(R)u^{\rhd_0,\rhd_1} + \lambda_1 v(R)u^{\rhd_1,\rhd_2},$$

we can now infer that $\lambda_0 = \lambda_1$. This completes the case that $t = 2$.

**Step 4:** Finally, we consider the case that $t > 2$. First, we assume for contradiction that there is $i \in \{0, \dots, t-2\}$ such that $\lambda_i < \lambda_{i+1}$. Clearly, this means that there is $k$ such that $k(\lambda_{i+1} - \lambda_i) > \sum_{j\in\{0,\dots,t-1\}\setminus\{i+1\}} \lambda_i$. Since the vectors $u^{\rhd_0,\rhd_1}, \dots, u^{\rhd_{t-1},\rhd_t}$ are linearly independent, we can find a vector $v'$ such that $v'u^{\rhd_i,\rhd_{i+1}} = k+1$, $v'u^{\rhd_{i+1},\rhd_{i+2}} = -k$, and $v'u^{\rhd_j,\rhd_{j+1}} = 1$ for all other vectors. By the linear dependence of $u^{\rhd_0,\rhd_t}$, we derive that

$$\begin{aligned}v'u^{\rhd_0,\rhd_t} &= \textstyle\sum_{j\in\{0,\dots,t-1\}} \lambda_j v' u^{\rhd_j,\rhd_{j+1}}\\ &= -k\lambda_{i+1} + k\lambda_i + \textstyle\sum_{j\in\{0,\dots,t-1\}\setminus\{i+1\}} \lambda_j\\ &< 0.\end{aligned}$$

This means that $v \notin \bar{R}_{\rhd_0}$. Furthermore, it holds for all $\rhd_j$ with $j \neq i+2$ that $v' u^{\rhd_{j-1}, \rhd_j} > 0$, which implies that $v' \notin \bar{R}_{\rhd_j}$ either. Finally, for $\rhd_{i+2}$, we use the fact that $u^{\rhd_i, \rhd_{i+2}} = \alpha(u^{\rhd_i, \rhd_{i+1}} + u^{\rhd_{i+1}, \rhd_{i+2}})$ for some $\alpha > 0$ (see Step 3) to derive that

$$v' u^{\rhd_i, \rhd_{i+2}} = \alpha(v' u^{\rhd_i, \rhd_{i+1}} + v' u^{\rhd_{i+1}, \rhd_{i+2}}) = \alpha(k + 1 - k) > 0,$$

so $v' \notin \bar{R}_{\rhd_{i+2}}$ either. Just as in Steps 2 and 3, we can now infer a contradiction by combining $v'$ with the vector $v \in \mathbb{Q}^{m!}$ which guarantees that $\hat{g}(v) = \{\rhd_0, \dots, \rhd_t\}$. In particular, there is a sufficiently small $\varepsilon > 0$ such that $v + \varepsilon v' \notin \bar{R}_{\rhd_j}$ for all $\rhd_j \in \mathcal{R}$. This contradicts that $\bigcup_{\rhd_j \in \mathcal{R}} \bar{R}_{\rhd_j} = \mathbb{R}^{m!}$, so we infer that hold that $\lambda_0 \geqslant \lambda_1 \geqslant \dots \geqslant \lambda_{t-1}$. Finally, we note that the case $\lambda_i > \lambda_{i+1}$ follows symmetrically by simply changing the "direction" of our construction: we now choose $v'$ such that

$$\begin{aligned} v' u^{\rhd_i, \rhd_{i+1}} &= k, \\ v' u^{\rhd_{i+1}, \rhd_{i+2}} &= -(k+1), \\ v' u^{\rhd_j, \rhd_{j+1}} &= -1 \quad \text{for all other } j. \end{aligned}$$

An analogous analysis as in the last case leads to a contradiction and the lemma follows. □

Finally, we are ready to show our main result.

THEOREM 3.2. *An SPF satisfies neutrality, reinforcement, continuity, and 2RP if and only if it is the Squared Kemeny rule.*

PROOF. Consider an SPF $f$ that satisfies all our requirements. By Lemma A.1, there is a function $\hat{g} : \mathbb{Q}^{m!} \to 2^{\mathcal{R}} \setminus \{\emptyset\}$ such that $f(R) = g(v(R))$ for all profiles $R \in \mathcal{R}^*$. Now, let $R_{\rhd_i} = \{v \in \mathbb{Q}^{m!} : \rhd_i \in \hat{g}(v)\}$ for every $\rhd_i \in \mathcal{R}$ and let $\bar{R}_{\rhd_i}$ denote the closure of $R_{\rhd_i}$ with respect to $\mathbb{R}^{m!}$. By Lemma A.2, we know that there are non-zero vectors $u^{\rhd_i, \rhd_j}$ for all $\rhd_i, \rhd_j \in \mathcal{R}$ such that $vu^{\rhd_i, \rhd_j} \geqslant 0$ if $v \in \bar{R}_{\rhd_i}$ and $vu^{\rhd_i, \rhd_j} \leqslant 0$ if $v \in \bar{R}_{\rhd_j}$. Moreover, by Lemma A.5 and Lemma A.11, we also know these vectors can be represented as follows: there is $\lambda > 0$ such that

$$u_k^{\rhd_i, \rhd_j} = \lambda(-\operatorname{swap}(\rhd_i, b(k))^2 + \operatorname{swap}(\rhd_j, b(k))^2)$$

for all $k \in \{1, \dots, m!\}$. In turn, Lemma A.3 shows that $\bar{R}_{\rhd_i} = \{v \in \mathbb{R}^{m!} : \forall \rhd_j \in \mathcal{R} \setminus \{\rhd_i\} : vu^{\rhd_i, \rhd_j} \geqslant 0\}$.

Now, let $s(v, \rhd_i) = \sum_{k=1}^{m!} -v_k \operatorname{swap}(\rhd_i, b(k))^2$ for every vector $v \in \mathbb{R}^{m!}$ and every $\rhd_i \in \mathcal{R}$. Since our normal vectors are invariant under scaling, it is easy to see that

$$\bar{R}_{\rhd_i} = \{v \in \mathbb{R}^{m!} : \forall \rhd_j \in \mathcal{R} \setminus \{\rhd_i\} : s(v, \rhd_i) \geqslant s(v, \rhd_j)\}.$$

Finally, Lemma A.4 shows for all $v \in \mathbb{Q}^{m!}$ that

$$\hat{g}(v) = \{\rhd_i \in \mathcal{R} : v \in \bar{R}_{\rhd_i}\} = \{\rhd_i \in \mathcal{R} : \forall \rhd_j \in \mathcal{R} \setminus \{\rhd_i\} : s(v, \rhd_i) \geqslant s(v, \rhd_j)\}.$$

Hence, $f(R) = g(v(R)) = \arg\max_{\rhd \in \mathcal{R}} s(v(R), \rhd) = \arg\min_{\rhd \in \mathcal{R}} C_{\text{SqK}}(R, \rhd)$ for all profiles $R \in \mathcal{R}^*$. This proves that $f$ is the Squared Kemeny rule. □

## B ADDITIONAL MATERIAL FOR SECTION 5

### B.1 Proof of Theorem 5.1

Before we prove the NP-completeness of computing the Squared Kemeny rule, we should clarify how the input profiles are to be encoded: they should be given as a list of rankings that occur in the profile, together with their weights.

THEOREM 5.1. *The problem of deciding, given a profile $R$ and a number $B$, whether there exists a ranking $\rhd$ with $C_{\text{SqK}}(R, \rhd) \leqslant B$, is NP-complete, even for profiles with 4 rankings with equal weight.*

Proof. Recall a standard inequality between the 1- and 2-norms:

$$\|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{n} \cdot \|x\|_2 \quad \text{for all } x \in \mathbb{R}^n,$$

where $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$ and $\|x\|_1 = \sum_{i=1}^n |x_i|$. We will use this inequality with $n = 4$:

$$\|x\|_1 \leqslant 2 \cdot \|x\|_2 \tag{4}$$

Write $C_{\mathrm{K}}(R, \rhd) = \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd)$ for the Kemeny score of the ranking $\rhd$ in profile $R$. It is well-known that the following problem of computing the Kemeny rule on profiles with 4 voters is NP-complete [Bachmeier et al., 2019, Biedl et al., 2009, Dwork et al., 2001].

| | |
|---|---|
| Input: | Profile $R$ in which exactly 4 rankings occur with equal weight, target score $k$ |
| Question: | Does there exist $\rhd$ with $C_{\mathrm{K}}(R, \rhd) \leqslant k$? |

By reducing from that problem, we will show that the following problem is NP-complete:

| | |
|---|---|
| Input: | Profile $R$ in which exactly 4 rankings occur with equal weight, target score $s$ |
| Question: | Does there exist $\rhd$ with $C_{\mathrm{SqK}}(R, \rhd) \leqslant s$? |

Consider an instance $(R, k)$ of the Kemeny problem, with $R$ defined on alternative set $A$ with $|A| = m$ and with rankings $\succ_1, \succ_2, \succ_3, \succ_4$ occurring in $R$ with weights $1/4$ each. We construct an instance of the Squared Kemeny problem. Our alternative set is going to be $A' = \bigcup_{a \in A} \{a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}\}$, consisting of 4 copies of each alternative $a \in A$. Thus $|A| = 4m$. The profile $R'$ on $A'$ consists of the following 4 rankings, each with weight $1/4$:[6]

$$\begin{aligned}
\succ'_a &= \succ_1^{(1)} \cdot \succ_2^{(2)} \cdot \succ_3^{(3)} \cdot \succ_4^{(4)} \\
\succ'_b &= \succ_2^{(1)} \cdot \succ_3^{(2)} \cdot \succ_4^{(3)} \cdot \succ_1^{(4)} \\
\succ'_c &= \succ_3^{(1)} \cdot \succ_4^{(2)} \cdot \succ_1^{(3)} \cdot \succ_2^{(4)} \\
\succ'_d &= \succ_4^{(1)} \cdot \succ_1^{(2)} \cdot \succ_2^{(3)} \cdot \succ_3^{(4)}
\end{aligned}$$

This notation is to be understood as follows: $\succ_i^{(j)}$ refers to the ranking induced by $\succ_i$ applied to the alternatives $\{a^{(j)} : a \in A\}$, i.e., the $j$-th copy of $A$. Rankings separated by a dot are concatenated. Thus, each ranking in $R'$ has in its top $m$ positions the alternatives from the first copy of $A$, in the next $m$ positions the alternatives from the second copy of $A$, and so on. However, the rankings in $R'$ differ in how they order each copy; for example $\succ_a$ ranks the first copy in the same way that $\succ_1$ ranks $A$. Our target score is going to be $s = \frac{1}{4}k^2$.

We show that there exist a ranking $\rhd$ with $C_{\mathrm{K}}(R, \rhd) \leqslant k$ if and only if there exists a ranking $\rhd'$ with $C_{\mathrm{SqK}}(R', \rhd') \leqslant s$.

$\Longrightarrow$ : Let $\rhd$ be such that $C_{\mathrm{K}}(R, \rhd) \leqslant k$. Let $\rhd' = \rhd^{(1)} \cdot \rhd^{(2)} \cdot \rhd^{(3)} \cdot \rhd^{(4)}$, i.e., the concatenation of 4 copies of $\rhd$. Note that we have

$$\mathrm{swap}(\succ'_a, \rhd') = \mathrm{swap}(\succ_1, \rhd) + \mathrm{swap}(\succ_2, \rhd) + \mathrm{swap}(\succ_3, \rhd) + \mathrm{swap}(\succ_4, \rhd) = C_{\mathrm{K}}(R, \rhd) \leqslant k.$$

Similarly, $\mathrm{swap}(\succ'_b, \rhd') = \mathrm{swap}(\succ'_c, \rhd') = \mathrm{swap}(\succ'_d, \rhd') \leqslant k$. Hence

$$C_{\mathrm{SqK}}(R', \rhd') \leqslant \tfrac{1}{4}(k^2 + k^2 + k^2 + k^2) = \tfrac{1}{4}k^2 = s.$$

$\Longleftarrow$ : Let $\rhd'$ be such that $C_{\mathrm{SqK}}(R', \rhd') \leqslant s = \frac{1}{4}k^2$.

Consider an optimum Kemeny ranking $\rhd$ in $R$, and let $t = C_{\mathrm{K}}(R, \rhd)$ be its Kemeny score. We want to show that $t \leqslant k$.

[6] The same profile construction is used in the reduction in Theorem 6 of Biedl et al. [2009] (showing that egalitarian Kemeny is NP-complete).

It is easy to see that $\rhd^* = \rhd^{(1)} \cdot \rhd^{(2)} \cdot \rhd^{(3)} \cdot \rhd^{(4)}$ is an optimum Kemeny ranking in $R'$, and it has Kemeny score $C_{\mathrm{K}}(R', \rhd') = t$ because $\mathrm{swap}(\succ_x, \rhd^*) = t$ for $x = a, b, c, d$. Now, we have

$$t = C_{\mathrm{K}}(R', \rhd^*) \overset{\text{opt}}{\leqslant} C_{\mathrm{K}}(R', \rhd') \overset{(4)}{\leqslant} 2 \cdot \sqrt{C_{\mathrm{SqK}}(R', \rhd')} \leqslant 2 \cdot \sqrt{s} = 2 \cdot \sqrt{\tfrac{1}{4}k^2} = k,$$

and thus $t \leqslant k$, as required. ◻

## B.2 ILP Formulation

Here, we present an Integer Linear Programming formulation for computing the Squared Kemeny rule. Let $R$ be a profile with $n$ different rankings. Then the ILP formulation will contain $\binom{m}{2}$ binary variables, $2n$ continuous variables, and $O(m^3 + n)$ constraints. To transform swap distances into squared swap distances, we use the trick described by Caragiannis et al. [2019] for computing the maximum Nash welfare solution for fair allocation, and generalized by Bredereck et al. [2020] to other ILPs with convex or concave objective functions.

In the formulation, for every pair $a, b \in A$ of distinct alternatives, we have a binary variable $x_{a,b}$ encoding whether $a \rhd b$ in the output ranking $\rhd$. The first type of constraint encodes completeness of the binary relation $\rhd$, while the second encodes transitivity of $\rhd$. For each ranking $\succ$ appearing in the profile $R$, the formulation includes a (continuous) variable $\mathrm{dist}_\succ$ which is constrained to equal $\mathrm{swap}(\succ, \rhd)$.[7] There is also a continuous variable $\mathrm{sqdist}_\succ$ which is constrained to be at least $\mathrm{swap}(\succ, \rhd)^2$. It will equal that value in the optimum solution.

$$\begin{array}{lll}
\text{minimize} & \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{sqdist}_\succ & \\
\text{subject to} & x_{a,b} + x_{b,a} = 1 & \text{for all } a, b \in A, a \neq b \\
& x_{a,b} + x_{b,c} + x_{c,a} \leqslant 2 & \text{for all } a, b, c \in A, \text{ all distinct} \\
& \mathrm{dist}_\succ = \sum_{a,b \in A \,:\, a \succ b} x_{b,a} & \text{for all } \succ \in \mathcal{R} \\
& \mathrm{sqdist}_\succ \geqslant k^2 + ((k+1)^2 - k^2) \cdot (\mathrm{dist}_\succ - k) & \text{for all } \succ \in \mathcal{R} \text{ and all } k \in [\binom{m}{2}] \\
& x_{a,b} \in \{0, 1\} & \text{for all } a, b \in A, a \neq b
\end{array}$$

## B.3 4-Approximation to Squared Kemeny

Theorem B.1. *There is a polynomial-time 4-approximation algorithm for the Squared Kemeny rule.*

Proof. Let $R$ be a profile and let $\rhd^* \in \mathrm{SqK}(R)$ be some Squared Kemeny output ranking. Consider the expected Squared Kemeny cost of a random ranking $\rhd$ drawn according to the weights in $R$ (i.e., viewing $R$ as a probability distribution). We have

$$\begin{aligned}
\mathbb{E}_{\rhd \sim R}[C_{\mathrm{SqK}}(R, \rhd)] &= \mathbb{E}_{\rhd \sim R}[\mathbb{E}_{\succ \sim R}[\mathrm{swap}(\succ, \rhd)^2]] && \text{(definition)} \\
&= \mathbb{E}_{\succ \sim R}[\mathbb{E}_{\rhd \sim R}[\mathrm{swap}(\succ, \rhd)^2]] && \text{(linearity of expectation)} \\
&\leqslant \mathbb{E}_{\succ \sim R}[\mathbb{E}_{\rhd \sim R}[(\mathrm{swap}(\succ, \rhd^*) + \mathrm{swap}(\rhd^*, \rhd))^2]] && \text{(triangle inequality)} \\
&\leqslant \mathbb{E}_{\succ \sim R}[\mathbb{E}_{\rhd \sim R}[2\,\mathrm{swap}(\succ, \rhd^*)^2 + 2\,\mathrm{swap}(\rhd^*, \rhd)^2]] && ((a+b)^2 \leqslant 2a^2 + 2b^2) \\
&= \mathbb{E}_{\succ \sim R}[\mathbb{E}_{\rhd \sim R}[2\,\mathrm{swap}(\succ, \rhd^*)^2]] + \mathbb{E}_{\succ \sim R}[\mathbb{E}_{\rhd \sim R}[2\,\mathrm{swap}(\rhd^*, \rhd)^2]] && \text{(linearity)} \\
&= \mathbb{E}_{\succ \sim R}[2\,\mathrm{swap}(\succ, \rhd^*)^2] + \mathbb{E}_{\rhd \sim R}[2\,\mathrm{swap}(\rhd^*, \rhd)^2] && (\mathbb{E}[\text{const.}] = \text{const.}) \\
&= 4 \cdot C_{\mathrm{SqK}}(R, \rhd^*). && \text{(definition)}
\end{aligned}$$

Thus, it follows that there exists $\rhd \in \mathrm{supp}(R)$ with

$$C_{\mathrm{SqK}}(R, \rhd) \leqslant 4 \cdot C_{\mathrm{SqK}}(R, \rhd^*).$$

[7] This variable can be eliminated from the formulation by replacing its value in the constraints placed on $\mathrm{sqdist}_\succ$.

Thus, the algorithm that goes through all rankings $\rhd$ in $\mathrm{supp}(R)$ and outputs one with minimum $C_{\mathrm{SqK}}(R, \rhd)$ is a 4-approximation of the Squared Kemeny rule. □

The same proof strategy has been used to obtain a 2-approximation to the Kemeny rule [Ailon et al., 2008].

### B.4 Proof of Theorem 5.2

Theorem 5.2. *For every constant $\varepsilon > 0$, there exists a polynomial-time $(2 + \varepsilon)$-approximation to the Squared Kemeny rule.*

Proof. Fix an arbitrary profile $R$ and ranking $\rhd \in \mathcal{R}$. Also, let $\rhd_{\mathrm{SqK}} \in \mathrm{SqK}(R)$ be a ranking selected by the Squared Kemeny rule. Furthermore, let $\alpha \geqslant 1$ be such that $\sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd) = \alpha \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}})$. Observe that we can always find such an $\alpha$: if the (weighted) average swap distance to $\rhd$ is smaller than to $\rhd_{\mathrm{SqK}}$, we can set $\alpha = 1$, and otherwise we can set it as the ratio between the distances.

It is sufficient to prove that

$$C_{\mathrm{SqK}}(R, \rhd) \leqslant 2\alpha^2 \cdot C_{\mathrm{SqK}}(R, \rhd_{\mathrm{SqK}}).$$

Indeed, if $\rhd$ is a $(\sqrt{1 + \varepsilon/2})$-approximation of Kemeny (and as e.g., $(1 + \varepsilon^2) < \sqrt{1 + \varepsilon/2}$ for small enough values of $\varepsilon$, we can find such a ranking in polynomial time [Kenyon-Mathieu and Schudy, 2007]), then surely $\alpha \leqslant (\sqrt{1 + \varepsilon/2})$, thus $\rhd$ is a $(2 + \varepsilon)$-approximation of Squared Kemeny as well. Observe that this inequality is equivalent to $C_{\mathrm{SqK}}(R, \rhd) - \alpha^2 \cdot C_{\mathrm{SqK}}(R, \rhd_{\mathrm{SqK}}) \leqslant \alpha^2 \cdot C_{\mathrm{SqK}}(R, \rhd_{\mathrm{SqK}})$. Which, using the $x^2 - y^2 = (x + y)(x - y)$ formula, we can write as

$$\begin{aligned}\textstyle\sum_{\succ \in \mathcal{R}} R(\succ)(\mathrm{swap}(\succ, \rhd) + \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))(\mathrm{swap}(\succ, \rhd) - \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}})) \leqslant \\ \textstyle\sum_{\succ \in \mathcal{R}} R(\succ) \cdot (\alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))^2. \quad (5)\end{aligned}$$

We will prove Inequality (5) using the following lemma.

Lemma B.2. *Given real numbers $x_1, x_2, \dots, x_n \geqslant 0$, $y_1, y_2, \dots, y_n \geqslant 0$ and $z_1, z_2, \dots, z_n$ such that $x_1 + x_2 + \cdots + x_n = 1$, $x_1 z_1 + x_2 z_2 + \cdots + x_n z_n \leqslant 0$ and $z_i \leqslant y_j$ for every $i, j \in [n]$, it holds that*

$$\sum_{i=1}^{n} x_i y_i z_i \leqslant \frac{1}{4} \left( \sum_{i=1}^{n} x_i y_i \right)^2.$$

Proof. Without loss of generality, let us assume that $y_1 \geqslant y_2 \geqslant \ldots \geqslant y_n \geqslant 0$. Then, the condition that $z_i \leqslant y_j$, for every $i, j \in [n]$, means simply that $z_i \leqslant y_n$, for every $i \in [n]$.

We will first show that for fixed values of $x_1, \dots, x_n, y_1, \dots, y_n$ the sum $\sum_{i=1}^{n} x_i y_i z_i$ is maximized if $z_i = y_n$, for every $i \in [n-1]$, and $z_n = (1 - x_n) y_n$. To this end, take $z_1, \dots, z_n$ defined like that and arbitrary $z'_1, \dots, z'_n$ such that $x_1 z'_1 + \cdots + x_n z'_n = 1$ and $z'_i \leqslant y_n$, for every $i \in [n]$. We will show that $\sum_{i=1}^{n} x_i y_i z_i \geqslant \sum_{i=1}^{n} x_i y_i z'_i$. Since for every $i \in [n-1]$ it holds that $y_i \geqslant y_n$ and $z'_i \leqslant y_n = z_i$ we get that

$$\sum_{i=1}^{n} x_i y_i z_i - \sum_{i=1}^{n} x_i y_i z'_i = \sum_{i=1}^{n} x_i y_i (z_i - z'_i) \geqslant y_n \sum_{i=1}^{n} x_i (z_i - z'_i).$$

Furthermore, we have that $\sum_{i=1}^{n} x_i z_i = 0$ and $\sum_{i=1}^{n} x'_i z_i \leqslant 0$, hence $\sum_{i=1}^{n} x_i (z_i - z'_i) \geqslant 0$. Thus, indeed $\sum_{i=1}^{n} x_i y_i z_i \geqslant \sum_{i=1}^{n} x_i y_i z'_i$.

In the remainder of the proof of this lemma, let us show that with such values of $z_1, \dots z_n$ maximizing the sum $\sum_{i=1}^{n} x_i y_i z_i$, the upper bound from the thesis still holds. Observe that for such values of $z_1, \dots, z_n$ we have

$$\sum_{i=1}^{n} x_i y_i z_i = \sum_{i=1}^{n-1} x_i y_i y_n - (1-x_n)y_n^2 = y_n\left(\sum_{i=1}^{n-1}(x_i y_i) - (1-x_n)y_n\right) = y_n\left(\sum_{i=1}^{n}(x_i y_i) - y_n\right).$$

Now, assume that the sum $\sum_{i=1}^{n}(x_i y_i)$ is a constant equal to $S$. What, given the value of $S$, would be the value of $y_n$ that would maximize the $\sum_{i=1}^{n} x_i y_i z_i$? The one that would maximize the term $y_n(S - y_n)$, which is a quadratic function with roots in 0 and $S$. Hence, since it is concave, the maximum is obtained halfway between the roots at $y_n = S/2$. Thus, we obtain that $\sum_{i=1}^{n} x_i y_i z_i \leqslant 1/2 \cdot \sum_{i=1}^{n}(x_i y_i) \cdot 1/2 \cdot \sum_{i=1}^{n}(x_i y_i)$, form which the thesis of the lemma follows. □

Now, let us use Lemma B.2 in order to prove Inequality (5), by showing that $x_\succ = R(\succ)$, $y_\succ = (\mathrm{swap}(\succ, \rhd) + \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))$, and $z_\succ = (\mathrm{swap}(\succ, \rhd) - \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))$ would satisfy the conditions of the lemma. To this end, observe that indeed $\sum_{\succ \in \mathcal{R}} x_\succ = \sum_{\succ \in \mathcal{R}} R(\succ) = 1$. Moreover, by the definition of $\alpha$,

$$\sum_{\succ \in \mathcal{R}} x_\succ z_\succ = \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd) - \alpha \sum_{\succ \in \mathcal{R}} R(\succ) \cdot \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}) \leqslant 0.$$

Thus, it suffices to show that for every $\succ, \succ' \in \mathcal{R}$ it holds that $z_\succ \leqslant y_{\succ'}$, i.e., $\mathrm{swap}(\succ, \rhd) - \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}) \leqslant \mathrm{swap}(\succ', \rhd) + \alpha\, \mathrm{swap}(\succ', \rhd_{\mathrm{SqK}})$. For this, observe that using triangle inequality for swap distance two times, we get that

$$\mathrm{swap}(\succ, \rhd) \leqslant \mathrm{swap}(\succ', \rhd) + \mathrm{swap}(\succ, \succ') \leqslant \mathrm{swap}(\succ', \rhd) + \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}) + \mathrm{swap}(\succ', \rhd_{\mathrm{SqK}}).$$

Since, as we assumed, $\alpha \geqslant 1$, we get the desired inequality. Therefore, we can indeed use Lemma B.2 to show that

$$\sum_{\succ \in \mathcal{R}} R(\succ)(\mathrm{swap}(\succ, \rhd) + \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))(\mathrm{swap}(\succ, \rhd) - \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}})) \leqslant \frac{1}{4}\left(\sum_{\succ \in \mathcal{R}} R(\succ)(\mathrm{swap}(\succ, \rhd) + \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))\right)^2. \quad (6)$$

Observe that from the definition of $\alpha$, we can bound the right-hand side of Inequality (6) by

$$\frac{1}{4}\left(\sum_{\succ \in \mathcal{R}} R(\succ)(\mathrm{swap}(\succ, \rhd) + \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))\right)^2 \leqslant \alpha^2 \left(\sum_{\succ \in \mathcal{R}} R(\succ)\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}})\right)^2.$$

Next, observe that $\sum_{\succ \in \mathcal{R}} R(\succ)\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}})$ is a weighted average, hence from Jensen's inequality, we get that

$$\frac{1}{4}\left(\sum_{\succ \in \mathcal{R}} R(\succ)(\mathrm{swap}(\succ, \rhd) + \alpha\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}}))\right)^2 \leqslant \alpha^2 \sum_{\succ \in \mathcal{R}} R(\succ)\, \mathrm{swap}(\succ, \rhd_{\mathrm{SqK}})^2.$$

Combining this with Inequality (6), yields Inequality (5). □

### B.5 Distance Between Kemeny and Squared Kemeny

There are profiles where the outputs of the Kemeny and the Squared Kemeny rules are almost reverse to each other, namely have distance $\binom{m}{2} - 1$. Write $A = \{a_1, \dots, a_m\}$, $m \geqslant 4$, let $\succ_1$ be the ranking $a_1 \succ \cdots \succ a_m$, and let $\succ_2$ be the ranking $a_2 \succ a_1 \succ a_3 \succ \cdots \succ a_m$. Note that $\mathrm{swap}(\succ_1, \succ_2) = 1$.

Let $\varepsilon > 0$. Consider the profile $R$ with

$$R(\succ_1) = 2 + \varepsilon,\ R(\succ_2) = 0, \text{ and } R(\succ) = 1 \text{ for all } \succ \in \mathcal{R} \setminus \{\succ_1, \succ_2\}.$$

Note that these weights sum up to more than 1, but we can normalize the weights without changing the argument.

Now, let us discuss the outputs of the Kemeny and Squared Kemeny rules for such profile $R$. To this end, consider an arbitrary ranking $\rhd \in \mathcal{R}$ and denote the sum of distances from $\rhd$ to every other ranking in $\mathcal{R}$ by $D_1 = \sum_{\succ \in \mathcal{R}} \text{swap}(\succ, \rhd)$. Observe that the Kemeny rule cost of ranking $\rhd$ is equal to

$$C_{\text{Kemeny}}(R, \rhd) = D_1 + (1+\varepsilon) \cdot \text{swap}(\succ_1, \rhd) - \text{swap}(\succ_2, \rhd).$$

Since $\succ_1$ and $\succ_2$ differ only on the ordering of the pair of alternatives, $a_1, a_2$, we get that

$$\text{swap}(\succ_1, \rhd) - \text{swap}(\succ_2, \rhd) = \begin{cases} 1, & \text{if } a_2 \rhd a_1, \text{ and} \\ -1, & \text{otherwise.} \end{cases} \tag{7}$$

Thus, we obtain

$$C_{\text{Kemeny}}(R, \rhd) = \begin{cases} D_1 + \varepsilon \cdot \text{swap}(\succ_1, \rhd) + 1, & \text{if } a_2 \rhd a_1, \text{ and} \\ D_1 + \varepsilon \cdot \text{swap}(\succ_1, \rhd) - 1, & \text{otherwise,} \end{cases}$$

which is minimized for $\rhd = \succ_1$. Thus, $\text{Kemeny}(R) = \{\succ_1\}$ for every $\varepsilon > 0$.

Now, let us consider the output of the Squared Kemeny rule. To this end, let us denote $D_2 = \sum_{\succ \in \mathcal{R}} \text{swap}(\succ, \rhd)^2$ and observe that

$$C_{\text{SqK}}(R, \rhd) = D_2 + (1+\varepsilon) \cdot \text{swap}(\succ_1, \rhd)^2 - \text{swap}(\succ_2, \rhd)^2.$$

Let us denote $d = \min(\text{swap}(\succ_1, \rhd), \text{swap}(\succ_2, \rhd))$ and observe that $\text{swap}(\succ_1, \rhd) + \text{swap}(\succ_2, \rhd) = 2d + 1$. Hence, by Equation (7), we get

$$\text{swap}(\succ_1, \rhd)^2 - \text{swap}(\succ_2, \rhd)^2 = (2d+1)(\text{swap}(\succ_1, \rhd) - \text{swap}(\succ_2, \rhd)) = \begin{cases} 2d+1, & \text{if } a_2 \rhd a_1, \text{ and} \\ -2d-1, & \text{otherwise.} \end{cases}$$

Therefore, we obtain that

$$C_{\text{SqK}}(R, \rhd) = \begin{cases} D_2 + \varepsilon \cdot \text{swap}(\succ_1, \rhd)^2 + 2d + 1, & \text{if } a_2 \rhd a_1, \text{ and} \\ D_2 + \varepsilon \cdot \text{swap}(\succ_1, \rhd)^2 - 2d - 1 & \text{otherwise.} \end{cases}$$

For $\varepsilon < \binom{m}{2}^{-2}$, the term $\varepsilon \cdot \text{swap}(\succ_1, \rhd)^2$ will be strictly smaller than 1. Hence, the value $C_{\text{SqK}}(R, \rhd)$ will be minimized for a ranking $\rhd$ such that $a_1 \rhd a_2$ and the value of $d$ is maximized. This will be the case for the ranking $\rhd^*$ such that $a_m \rhd^* a_{m-1} \rhd^* \cdots \rhd^* a_3 \rhd^* a_1 \rhd^* a_2$. Thus, $\text{SqK}(R) = \{\rhd^*\}$. Since $\text{swap}(\succ_1, \rhd^*) = \binom{m}{2} - 1$, we see that for the profile $R$ with $0 < \varepsilon < \binom{m}{2}^{-2}$, the Kemeny and Squared Kemeny rules indeed output almost reversed rankings.

# C ADDITIONAL MATERIAL FOR SECTION 6

## C.1 City Experiment Data

Table 2 presents the data we have used for the city ranking experiment in Section 6.1. The GDP per capita data is taken from Wikipedia.[8] The air quality ranking is based on the average PM 2.5 concentration for the year 2018. This was the year for which the data was the most complete, however in a few cases we had to use the data from different year (as noted in the table). The majority of the PM 2.5 concentration data comes from the World Health Organization database [WHO, 2024] (the only exception is the data for Cairo and Lagos that comes from an online article [Oğuz, 2023]). Finally, the sunniness ranking is based on the average number of hours of sunshine per year. The data for this was gathered from the Wikipedia articles about each city on 6 February 2024.

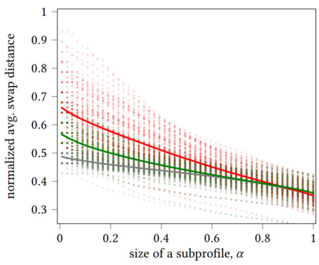
[8] https://en.wikipedia.org/wiki/List_of_cities_by_GDP, accessed: 6 February 2024

| City | GDP per capita (US$) | Avg. PM 2.5 conc. ($\mu$g/m$^3$) | Avg. Sunshine h. per year |
|---|---|---|---|
| Bangkok | 12, 670 | $23.14^{2019}$ | 2, 212 |
| Buenos Aires | 14, 024 | $10.26^{2015}$ | 2, 384 |
| Cairo | 8, 685 | $47.40^{2022,\dagger}$ | 3, 451 |
| Dubai | 47, 557 | 53.93 | 3, 570 |
| Dublin | 104, 394 | 7.89 | 1, 452 |
| Hong Kong | 52, 431 | 20.09 | 1, 829 |
| Istanbul | 14, 989 | 28.78 | 2, 181 |
| Johannesburg | 16, 033 | 22.69 | 3, 124 |
| Lagos | 3, 607 | $36.10^{2022,\dagger}$ | 1, 844 |
| Lahore | 2, 878 | $123.88^{2019}$ | 3, 034 |
| London | 66, 108 | 10.49 | 1, 675 |
| Mexico | 13, 798 | 22.00 | 2, 526 |
| Moscow | 29, 012 | $14.00^{2016}$ | 1, 731 |
| Mumbai | 10, 651 | 75.45 | 2, 612 |
| New York City | 114, 293 | 7.65 | 2, 535 |
| Paris | 63, 119 | 14.01 | 1, 717 |
| Rio de Janeiro | 15, 742 | $11.45^{2015}$ | 2, 182 |
| Rome | 40, 535 | 13.98 | 2, 724 |
| San Francisco | 157, 704 | 11.65 | 3, 062 |
| Seoul | 36, 677 | 22.93 | 2, 143 |
| Shanghai | 26, 672 | 37.66 | 1, 851 |
| Sydney | 73, 034 | $11.27^{2019}$ | 2, 639 |
| Tokyo | 51, 124 | 12.91 | 1, 927 |
| Toronto | 69, 110 | 8.00 | 2, 066 |
| Zurich | 108, 104 | 12.13 | 1, 694 |

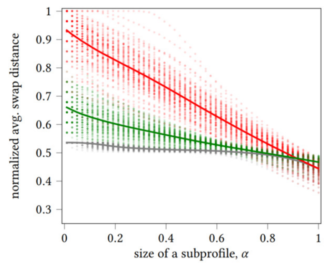
Table 2. The data used to the city ranking analysis in Section 6.1. The year in a superscript of the values PM 2.5 concentration column signifies that the data used was from a year that is different from 2018 (since for 2018 no data was available). Also, † signifies a different source of data.

## C.2 Worst-Case Average Distance

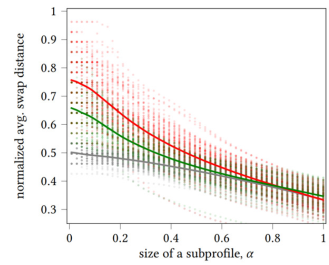
Figure 13 presents the plots described in Section 6.3 for the profiles sampled from models described in Section 6.2.



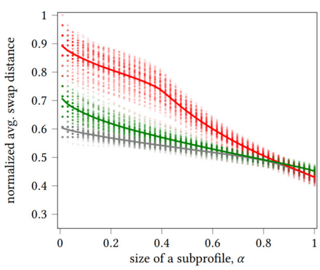

(a) Disc



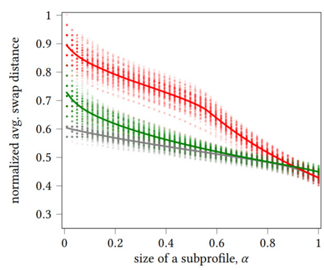

(b) Circle



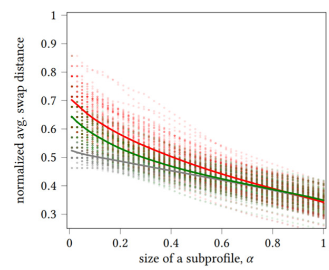

(c) Countries




(d) Mallows Mxt. ($\phi_1 = \phi_2 = 0.5$)




(e) Mallows Mxt. ($\phi_1 = 0.7, \phi_2 = 0.3$)




(f) Breakfast

Fig. 13. The maximal average distances between the subprofile of size $\alpha$ and the output of the Kemeny (red) and Squared Kemeny (green) rules, plus lower bound (gray).