

Proportionality Degree of Multiwinner Rules

Piotr Skowron

Abstract

An instance of an approval-based multiwinner election consists of a set of alternatives, a population of voters—each voter approves a subset of alternatives, and the desired committee size k ; the goal is to select a committee (a subset) of k alternatives according to the preferences of the voters. We investigate a number of election rules and ask whether the committees that they return represent the voters proportionally. In contrast to the classic literature, we employ quantitative techniques that allow to measure the extent to which the considered rules are proportional. This allows us to arrange the rules in a clear hierarchy. For example, we find that Proportional Approval Voting (PAV) has better proportionality guarantees than its sequential counterpart, and that Phragmén’s Sequential Rule is worse than Sequential PAV. Yet, the loss of proportionality for the two sequential rules is moderate and in some contexts can be outweighed by their other appealing properties.

1 Introduction

An approval-based committee election rule (an ABC rule, in short) is a function that given a set of m candidates C (the candidates are also referred to as alternatives), a population of n voters—each approving a subset of C —and an integer k representing the desired committee size, returns a k -element subset of C . Committee election rules are important tools that facilitate collective decision making in various contexts such as electing representative bodies (e.g., supervisory boards, trade unions, etc.), finding responses to database querying [10, 26], selecting validators in consensus protocols, such as the blockchain [7], suggesting collective recommendations for groups [21, 22], and managing discussions on proposals within liquid democracy [4]. Further, since committee elections are a special case of participatory budgeting (PB) [8, 14], good understanding of ABC rules is a prerequisite for designing effective PB methods.

The applicability of various ABC rules often depends on the particular context, yet an important requirement one often imposes on a committee election rule is that it should be fair to (groups of) voters. While fairness is a broad concept putting various principles under the same umbrella, in certain types of collective decision-making—specifically when the goal is to elect a committee—it is often argued that a fair rule should be *proportional*¹, as illustrated by the following example.²

Example 1. *Consider an election with 30 candidates, c_1, c_2, \dots, c_{30} , and 100 voters having the following approval-based preferences. The first 60 voters approve the subset $C_1 = \{c_1, \dots, c_{10}\}$; the next 30 voters approve $C_2 = \{c_{11}, \dots, c_{20}\}$, and the last 10 approve $C_3 = \{c_{21}, \dots, c_{30}\}$. When the size of the committee to be elected is 10, then a proportional rule should select a committee with 6 members coming from C_1 , 3 members coming from C_2 , and 1 member from C_3 .*

¹Other basic axioms describing fairness requirements are, e.g., anonymity or solid coalitions property [9].

²Indeed, proportional representation (PR) electoral systems are often argued to be more fair than non-proportional ones [15, 20]. Yet, the criticism of proportionality also appears in the literature, and it comes from the two main directions—one stems from the analysis of voting power in the elected committees [13, 25]; the other one is grounded in the arguments in favor of degressive proportionality [17]. Proportionality is also related to fairness in other domains, such as allocation of individual [27] and public goods [12].

Identifying a proportional committee in [Example 1](#) was easy due to the voters’ preferences having a very specific structure: each two approval sets were either the same or disjoint. In a more general case, when the approval sets can arbitrarily overlap, the answer is no longer straightforward. Several approaches to formalizing proportionality have been proposed in the literature [[1](#), [2](#), [6](#), [12](#), [19](#), [24](#)]. These approaches are mainly axiomatic—they formally define natural properties referring to proportionality, and classify known rules based on whether they satisfy these axioms or not. With a few notable exceptions (e.g., [[12](#), [24](#)]) the approaches from the literature are qualitative, giving only a yes/no answer to the question of “Is a given rule proportional or not?”. Hence, in essence, they are not capable of measuring the extent to which proportionality is satisfied or violated. This is a serious drawback since there is no single rule satisfying all the desired properties, and the choice of the rule boils down to a judgment call. The mechanism designer must decide on which properties she finds most critical and which tradeoff she is willing to accept. However, in order to make such a decision one primarily needs to understand these tradeoffs; in particular the extent to which properties of interest are violated.

In our study we employ a quantitative approach. For several rules we will determine their *proportionality degrees*, i.e., we will assess the extent to which they satisfy proportionality. Informally speaking, the proportionality degree of a rule is a function specifying how the rule treats groups of voters with cohesive preferences, depending on the size of these groups. At the same time, our proportionality degree has the form of a guarantee—it gives the best possible bounds on the proportionality that the rule cannot violate, no matter what are the voters’ preferences.

Our definition of proportionality degree is closely related to the concept proposed by Skowron et al. [[26](#)]. In fact, their work establishes certain bounds on the proportionality degree for a number of ABC rules. Unfortunately, these bounds cannot be used for comparing the rules, since they are not tight; not even asymptotically. Our main contribution is that we derive new almost tight bounds that allow to arrange the rules that we study in a clear hierarchy, based on how proportional they are. A tight estimation of the proportionality degree is already known for Proportional Approval Voting (PAV) [[2](#)—we generalize this result and calculate the proportionality degree for convex Thiele methods, a broad class of rules which—in particular—includes PAV. Additionally, we find close estimations of the proportionality degree for two sequential methods often considered in the literature: Sequential PAV and Phragmén’s Sequential Rule.

Our findings can be summarized as follows. As far as the proportionality guarantees are concerned, Sequential PAV is better than Phragmén’s Sequential Rule; further, PAV is better than both of the sequential rules. For reasonable committee sizes the proportionality degree of Sequential PAV is no worse than roughly 70% of the proportionality degree of PAV. The proportionality degree of Phragmén’s Sequential Rule is roughly half as much as that for PAV. On the one hand, our results suggest that PAV should be preferred whenever proportionality is the primary goal. On the other hand, they demonstrate that the loss of proportionality for Sequential PAV and Phragmén’s Sequential Rule is moderate. Thus, using these rules can be justified in cases when the decision maker considers their other distinctive properties equally important to proportionality (in [Sections 3](#) and [4](#) we briefly recall a few arguments that can speak in favour of one of the two sequential rules).

Importantly, we clarify a few misunderstandings that appear in the literature. It was commonly believed that PAV provides the strongest proportionality guarantees (as it satisfies the strong axiom of Extended Justified Representation), that Phragmén’s Sequential Rule and Phragmén’s Maximal Rule provide slightly weaker guarantees (as they both satisfy Proportional Justified Representation), and that Sequential PAV cannot be considered a proportional rule (as it fails even the basic axiom of Justified Representation). We argue these conclusions are not entirely correct, and that some of these results speak rather about

the specific choices made in the definitions of the axioms than about the nature of the voting rules. By looking at quantitative measures of proportionality we have discovered that in fact PAV, Sequential PAV, and Phragmén’s Sequential Rule give all very strong proportionality guarantees, while Phragmén’s Maximal Rule provides almost no such.

All proofs omitted from the main text are provided in [Appendix A](#).

2 The Model

For each natural number $p \in \mathbb{N}$, we set $[p] = \{1, 2, \dots, p\}$ and $[p]_0 = \{0\} \cup [p]$. For a set X by $S_p(X)$ we denote the set of all p -element subsets of X , and by $S(X)$ the powerset of X , $S(X) = \bigcup_{p=0}^{|X|} S_p(X)$.

An approval-based election is a triple (N, C, A) , where $N = \{1, 2, \dots, n\}$ is the set of voters, $C = \{c_1, c_2, \dots, c_m\}$ is the set of candidates, and A is an approval-based profile (or, in short, a profile), i.e., a function $A: N \rightarrow S(C)$ that maps each voter to a subset of C ; intuitively, $A(i)$ consists of candidates that voter i finds acceptable and is referred to as the approval set of i . Conversely, for each candidate $c \in C$ by $N(c)$ we denote the set of voters who approve c , i.e., $N(c) = \{i \in N \mid c \in A(i)\}$. Whenever we consider an election, we will implicitly assume that N and C refer to the sets of voters and candidates, respectively. Similarly, we will always assume that n and m denote the number of voters and candidates, respectively. For the sake of the simplicity of the notation we will also identify elections with approval-based profiles, assuming that the sets of voters and candidates are implicitly encoded as, respectively, the domain and the union of elements in the range of A . We denote the set of all elections by \mathcal{A} (yet, following our convention, we will treat the elements of \mathcal{A} as if these were simply profiles).

2.1 Approval-Based Committee Election Rules

An approval-based committee election rule (an ABC rule, in short) is a function \mathcal{R} that for each approval-based profile A and a positive integer k returns a nonempty set of size- k subsets of candidates, i.e., $\mathcal{R}(A, k) \subseteq S_k(C)$. We will refer to the elements of $S_k(C)$ as to size- k committees, or simply as to committees when the committee size k is known from the context. We will call the elements of $\mathcal{R}(A, k)$ winning committees. We say that a candidate c represents voter i (or that c is a representative of i) if c is a member of a winning committee and if $c \in A(i)$.

Below, we recall definitions of several prominent (classes of) approval-based committee election rules, often studied in the literature in the context of proportional representation.

Thiele Methods. For a non-increasing function $\lambda: \mathbb{R} \rightarrow \mathbb{R}$ the λ -Thiele rule is defined as follows.³ The λ -score that a committee $W \in S_k(C)$ gets from a voter $i \in N$ is equal to $sc_\lambda(W, i) = \sum_{j=1}^{|A(i) \cap W|} \lambda(j)$. The total λ -score of W is the sum of the scores that it garners from all the voters: $sc_\lambda(W) = \sum_{i \in N} sc_\lambda(W, i)$. The λ -Thiele rule returns the committees with the highest λ -score.

Proportional Approval Voting (PAV). PAV is the λ_{PAV} -Thiele rule for $\lambda_{\text{PAV}}(i) = 1/i$.

Using the harmonic sequence to compute the score ensures that PAV satisfies particularly appealing properties pertaining to proportionality [1, 6, 19].

Sequential Proportional Approval Voting (Seq-PAV). This is an iterative rule that involves k steps. It starts with an empty committee $W = \emptyset$ and in each step it adds to

³For the definition of the rules it is sufficient to have λ specified only on the set of natural numbers, yet later on it will be sometimes convenient for us to work with the continuous version of λ .

W the candidate that increases the PAV score of W most. Thus, intuitively, each voter starts with the same voting power (equal to 1), which can decrease over time. When the voter gets her i -th representative in the committee, then her voting power decreases from $1/i$ to $1/(i+1)$.

Phragmén’s Sequential Rule. This is also an iterative rule that is usually defined as a load balancing procedure. Each candidate $c \in C$ is associated with one unit of load; if c is selected this unit of load needs to be distributed among the voters who approve c . In each step the rule selects the candidate that minimizes the load assigned to the maximally loaded voter. Formally, let $\ell_i(j)$ be the total load assigned to voter i after the j -th iteration ($\ell_i(0) = 0$ for all $i \in N$). In each step j the rule selects a candidate c and finds a distribution $\{\delta_i(j)\}_{i \in N}$ of one additional unit of load such that: (i) $\sum_{i \in N} \delta_i(j) = 1$, (ii) $\delta_i(j) \geq 0$ for each $i \in N$, (iii) $\delta_i(j) = 0$ for each $i \notin N(c)$, and (iv) $\max_{i \in N} (\ell_i(j-1) + \delta_i(j))$ is minimized. (Note that minimizing the maximum load (iv) is the criterion for choosing both c and δ .) Finally, for each $i \in N$ the rule updates the total load, $\ell_i(j) = \ell_i(j-1) + \delta_i(j)$. For more discussion on the rule and its properties we refer to the work of Brill et al. [5] and to the survey by Janson [16].

Phragmén’s Maximal Rule. This method is similar to Phragmén’s Sequential Rule, but in this case the rule does not proceed sequentially—the whole optimization is performed only once, globally. The criterion for choosing the winning committee and the corresponding load distribution is that the total load assigned to the maximally loaded voter should be minimized.

2.2 Proportionality Degree of ABC Rules

This section defines and explains the notion of proportionality used in our comparative study.

Given an approval-based profile A and a desired committee size $k \in \mathbb{N}$, we say that a group of voters $V \subseteq N$ is ℓ -large if $|V| \geq \ell \cdot \frac{n}{k}$. The satisfaction of the group V from a committee $W \in S_k(C)$ is defined as the average number of representatives that a voter from V has in committee W :

$$\text{sat}_V(A, W) = \frac{1}{|V|} \sum_{i \in V} |W \cap A(i)|.$$

Definition 1. *We say that a function $g: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ is a k -proportionality guarantee for an ABC rule \mathcal{R} if for each approval-based profile $A \in \mathcal{A}$, each ℓ -large group of voters $V \subseteq N$, and for each winning committee $W \in \mathcal{R}(A, k)$, it holds that:*

$$\text{sat}_V(A, W) \geq \min \left(\left| \bigcap_{i \in V} A(i) \right|, g(\ell, k) \right).$$

Note that the condition in **Definition 1** must hold for all values of n and $m \geq k$; in particular, g cannot depend on n , m nor on the specific structural properties of approval profiles.

Let us give an informal explanation of **Definition 1**; intuitively, this definition says that an ℓ -large group of voters is guaranteed $g(\ell, k)$ representatives in the size- k committee elected by a rule \mathcal{R} , no matter what the preferences of the other voters are. The average satisfaction of at least $g(\ell, k)$ is guaranteed only to ℓ -large groups that have coherent enough preferences, that is for groups that agree on some $g(\ell, k)$ common candidates. If a group agrees on less candidates (that is, if $|\bigcap_{i \in V} A(i)| < g(\ell, k)$), then the group is guaranteed to achieve a respectively lower satisfaction. Indeed, it is clear that if each member of a group V approves

different candidates, then V cannot be satisfied by any rule; in fact we will call such V simply a *set* instead of a *group* to indicate that there is no agreement between the voters in V , so they cannot be related or grouped based of their preferences. Thus, [Definition 1](#) specifies how the rule treats cohesive groups of certain sizes.

One intuitively expects that a proportional rule should have a guarantee of $g(\ell, k) = \ell$. Indeed, a group V such that (i) $|V| \geq \ell \cdot \frac{n}{k}$, and (ii) $|\bigcap_{i \in V} A(i)| \geq \ell$, is large enough to deserve ℓ representatives in the elected committee, and choosing ℓ candidates that all members of the group agree on is feasible. Unfortunately, such a guarantee is not possible to achieve. From the recent results of Aziz et al. [\[2\]](#) it follows that there exists no rule with the proportionality guarantee satisfying $g(\ell, k) > \ell - 1$. The same work proves that PAV matches this lower bound (they state this result in a less general form, but the proof directly extends to our definition of proportionality degree).

Theorem 1 (Aziz et al. [\[2\]](#)). *PAV has the proportionality guarantee of $g(\ell, k) = \ell - 1$.*

In order to facilitate our further discussion we define two related concepts.

Definition 2. *Let $\text{guar}(\mathcal{R})$ be the set of all k -proportionality guarantees for a rule \mathcal{R} . The k -proportionality degree of an ABC rule \mathcal{R} is the function:*

$$d_{\mathcal{R}}(\ell, k) = \sup_{g \in \text{guar}(\mathcal{R})} g(\ell, k).$$

In other words, the k -proportionality degree is the best possible k -proportionality guarantee one can find for the rule. The proportionality degree of \mathcal{R} is the function $d_{\mathcal{R}}(\ell) = \min_k d_{\mathcal{R}}(\ell, k)$, i.e., it defines the best possible guarantee that holds irrespectively of the size of the committee.

The idea of measuring proportionality of multiwinner rules as the average number of representatives that cohesive groups of voters get was first proposed by Sánchez-Fernández et al. [\[24\]](#) (they call the average number of representatives the *average satisfaction* of a voter). Though, the first estimation of the proportionality degree for a multiwinner rules appears only in the subsequent work of Aziz et al. [\[2\]](#) (see [Theorem 1](#), above); the analysis there uses a swap-optimality argument—they show that if there existed a committee W witnessing that the proportionality guarantee of PAV is lower than $d_{\text{PAV}}(\ell) = \ell - 1$, then it would be possible to find a pair of candidates, $c \in W$ and $c' \notin W$, such that replacing c with c' in W would improve the PAV score of W . The same argument applies to a local-search heuristic for PAV yielding a polynomial-time rule with virtually the same proportionality degree as PAV—this result solved an open question from 2015. Next, Skowron et al. [\[26\]](#) analyzed the proportionality degree of a number of ABC rules (though, the definition of the proportionality degree appears there only implicitly), focusing on those that satisfy the committee enlargement monotonicity. However, their estimations are not asymptotically tight, and are not sufficient to compare the studied rules. E.g., for Phragmén’s Sequential Rule, they showed that $d_{\text{Phrag}}(\ell, k) \geq \frac{\ell(\ell+1)}{5k}$; we will strengthen this result by showing that $d_{\text{Phrag}}(\ell) = \frac{\ell}{2} \pm \frac{1}{2}$ (note that $k \geq \ell$, and that for “small” groups k is much greater than ℓ , and so the estimation given in [\[26\]](#) is highly inaccurate; specifically, it does not allow to obtain a positive bound on $d_{\text{Phrag}}(\ell)$ nor to compare Phragmén’s Sequential Rule with virtually any rule known in the literature).

In [Appendix B](#) we compare the notion of the proportionality guarantee/degree with other concepts pertaining to proportionality, studied in the literature.

3 Proportionality Degree of Phragmén’s Sequential Rule

In this section we establish the proportionality degree of Phragmén’s Sequential Rule. However, we first provide an alternative definition of the rule that will be more convenient to work with. According to our new definition the voters gradually earn virtual money (credits) which they then use to buy committee members. Specifically, we assume that each voter earns money with the constant speed of one credit per unit time (time is continuous, not discrete). Buying a candidate costs n credits, and a voter pays only for a candidate that she approves of. We say that a candidate c is *electable* if c is approved by the voters who altogether have at least n credits. In the first time moment when there exists an electable candidate c the rule adds this candidate to the committee (ties are broken arbitrarily) and resets to zero the credits of all voters who approve c —intuitively, these voters pay the total amount of n for adding c to the committee. The rule stops when k candidates are selected.

The so-described process is equivalent to Phragmén’s Sequential Rule. It is apparent that in the original definition of Phragmén’s Sequential Rule we can assume that each candidate is associated with n units of load instead of one. When a candidate c is selected then its load is distributed so that all the voters who approve c have the same total load, which is the maximum load among all the voters. The same candidate would be selected by the above described process, and each voter v approving c would pay for c the number of credits which is equal to the difference between its current total load and the previous one (just before c was selected). We provide the formal proof of the equivalence of the two definitions in [Appendix C](#).

Using the new definition allows us to obtain a new accurate estimation of the proportionality degree for Phragmén’s Sequential Rule. Informally speaking, our guarantee says that Phragmén’s Sequential Rule can be at most twice less proportional than PAV. The crucial element of the proof is the analysis of a certain potential function. Intuitively, this function measures how unfair the committee iteratively built by Phragmén’s Sequential Rule can be; we will show that (i) this unfairness cumulates, and that (ii) eventually the accumulated unfairness must be used to compensate the voters who got less representatives than they deserved.

Theorem 2. *The proportionality degree of Phragmén Seq. Rule satisfies $d_{\text{Phrag}}(\ell) \geq \frac{\ell-1}{2}$.*

[Proposition 1](#) below upper-bounds the proportionality degree of Phragmén’s Sequential Rule.

Proposition 1. *The k -proportionality degree of Phragmén’s Sequential Rule satisfies the following inequality: For each $k, \ell \in \mathbb{N}$, $\ell < \frac{k}{2}$ and k divisible by ℓ we have $d_{\text{Phrag}}(\ell, k) \leq \frac{\ell}{2} \cdot \frac{2k-2\ell+2}{2k-3\ell}$.*

Next, observe that for large enough k the proportionality guarantee from [Proposition 1](#) can be arbitrarily close to $\ell/2$. Thus, we obtain the following corollary, which shows that the guarantee from [Theorem 2](#) is almost tight (up to an additive constant of $1/2$).

Corollary 1. *The proportionality degree of Phragmén’s Seq. Rule satisfies $d_{\text{Phrag}}(\ell) \leq \frac{\ell}{2}$.*

Let us now discuss the consequences of our results from the perspective of a decision maker facing the problem of choosing the right rule. First, Phragmén’s Sequential Rule offers a considerably lower proportionality guarantee than PAV, hence the latter should be recommended whenever proportionality is the primary concern. Second, the worst-case loss of proportionality for Phragmén’s Sequential Rule is moderate. Thus, in some cases this loss can be compensated by other appealing properties of the rules. For example,

Phragmén’s Sequential Rule satisfies committee enlargement monotonicity⁴, which makes it applicable when the goal is to compute a representative *ranking* of alternatives (the recent work of Skowron et al. [26] discusses several domains where finding proportional rankings is critical). Further, as discussed by Janson [16] (see, e.g., Examples 13.5 there), Phragmén’s Sequential Rule has the following appealing property:

Definition 3 (Strong Unanimity). *An ABC rule \mathcal{R} satisfies strong unanimity if for each approval-based profile A such that there exists a candidate c who is approved by all the voters, it holds that $\mathcal{R}(A, k) = \mathcal{R}(A^{-c}, k - 1) \cup \{c\}$, where A^{-c} denotes the profile obtained from A by removing c from the approval sets of all the voters.*

It easily follows from the definition that Phragmén’s Sequential Rule satisfies strong unanimity. This property is very natural, PAV does not satisfy it (this is an argument often raised by the critics of PAV, e.g., by Phragmén in his original works). **Theorem 2** shows that strong unanimity and committee enlargement monotonicity can be satisfied with only moderate loss of proportionality compared to PAV. As such, it allows to view Phragmén’s Sequential Rule as an appealing alternative for PAV (depending on how important the decision maker considers particular axioms).

3.1 Comparing Phragmén’s Sequential Rule and Phragmén’s Maximal Rule

Phragmén’s Maximal Rule was first introduced as an “optimal” alternative for its sequential counterpart. Thus, it is interesting to see that the global optimization embedded in the definition of the rule, does not translate to its better properties. In fact, it appears that Phragmén’s Maximal Rule is far from being optimal (it is not Pareto optimal and its utilitarian efficiency is much worse than those of Phragmén’s Sequential Rule) [18]. Below, we will complement this arguments showing that the proportionality degree of the rule is much worse than those of its sequential counterpart. Our results suggest that there is no reason to choose Phragmén’s Maximal Rule over Phragmén’s Sequential Rule.

Proposition 2. *The proportionality degree of Phragmén Maximal Rule is $d_{\max\text{-Phrag}}(\ell) \leq 1$.*

4 Proportionality Degree of Sequential PAV

In this section we will assess the proportionality degree of Sequential PAV, with the aim of comparing Seq-PAV with Phragmén’s Sequential Rule. Our method here is quite different from the one we used in the previous section. Specifically, instead of proving a bound on the proportionality degree of Seq-PAV directly, we will show how to construct an algorithm that given a desired committee size k finds in polynomial time an upper bound on the k -proportionality degree of the rule. We will compute these bounds for $k \leq 200$ and will argue that they are fairly accurate. Since for many natural applications of ABC rules the intended committee size is much smaller than 200, our results give a direct answer to the decision maker who is interested in estimating the proportionality degree for Seq-PAV. Yet, our result is more generic and allows a mechanism designer to make a comparison of rules having a specific committee size in mind (indeed, a committee size is usually known and fixed before the rule is used).

⁴Informally speaking, a rule \mathcal{R} satisfies committee enlargement monotonicity (sometimes, also referred to as simply committee monotonicity) if increasing the committee size from k to $k + 1$ results only in adding an additional candidate to the winning committees (the candidates that were selected by the rule for the committee size equal to k are still selected for the committee size equal to $k + 1$). For the precise definition, see [11].

Designing such an algorithm is not straightforward—the main challenge lies in reducing the size of the space that needs to be searched. Indeed, even for a fixed committee size there is an infinite number of possible profiles (even if we fix the number of voters n and the number candidates m we still have exponentially many, namely 2^{mn} , possible preference profiles). Thus, we will use several observations that will allow us to reformulate the problem and to compactly represent it as an instance of Linear Programming (LP).

4.1 The High-Level Strategy of the Algorithm

The high-level idea of our approach can be summarized as follows. First, in [Subsection 4.2](#) we describe a different optimization function that accurately estimates the proportionality degree of Seq-PAV (at first this optimization function can appear loosely related to the concept of proportionality degree). This optimization function is somehow easier to work with and in [Subsection 4.3](#) we show that it can be solved by a carefully constructed linear programming (LP): this LP takes the committee size as input and returns the proportionality degree of Seq-PAV. This algorithm is exponential in the committee size but does not depend on the number of voters nor the number of candidates, thus it allows us to derive almost exact values of the proportionality degree of Sequential PAV for small values of k . (Up to this point, we can formally prove that the so-constructed LP returns accurate estimations of the proportionality degree up to an additive constant of 1.)

Next, we will show a more complex LP which we argue is a relaxation of the first exact one. The second LP works in polynomial time in the committee size k , and also its formulation does not depend on n nor on m . This LP also provides formal, analytical proportionality guarantees for Seq-PAV, thus these values can be used to formally compare Seq-PAV with other rules. Indeed, we show that the so-obtained guarantees (i.e., the lower bounds on the proportionality degree) for Seq-PAV are considerably better than the proportionality degree for Phragmén’s Sequential Rule. While these results allow us to compare Seq-PAV with the Phragmén’s sequential rule, we still do not know how tight are the guarantees returned by our second LP. While the formal analysis of the second LP is much harder, we observe that for small values of k the two LP-s return very similar guarantees (for $k \leq 20$ they differ by at most 2%) and we observe that the plot of the guarantees returned by our second LP for different values of k flattens very quickly (e.g., the guarantees for $k = 50$ and $k = 200$ equal to 0.7085ℓ and 0.694ℓ , respectively). This allows us to conjecture that the guarantees returned by our second LP are fairly accurate, at least for reasonable committee sizes.

In [Appendix E](#) we discuss why obtaining analytical bounds that accurately estimate the proportionality degree of Seq-PAV for all values of k is a hard task. We suggest it as an interesting and challenging open question.

4.2 Revising the Optimization Objective

Let us introduce some additional notation. Let A and k be an approval-based profile and a desired committee size, respectively. Let $\mathcal{R}_{\text{seq-PAV}}(A, k)$ be a multiset of tied winning committees returned by Sequential Proportional Approval Voting for A and k (here, we use a multiset because the same committee could be obtained by different tie-breaking decisions during the execution of the rule). For each winning committee $W \in \mathcal{R}_{\text{seq-PAV}}(A, k)$ let $c_{\text{last}}(A, W)$ denote the candidate that Sequential Proportional Approval Voting has added to W as the last committee member. Let $\Delta_{\text{seq-PAV}}(A, k, W)$ denote the average, per voter, marginal increase of the PAV score due to adding $c_{\text{last}}(A, W)$ to W :

$$\Delta_{\text{seq-PAV}}(A, k, W) = \frac{1}{n} \left(\text{score}_{\text{PAV}}(A, W) - \text{score}_{\text{PAV}}(A, W \setminus \{c_{\text{last}}(A, W)\}) \right).$$

We denote the maximum possible average marginal increase as $\Delta_{\text{seq-PAV}}(A, k)$:

$$\Delta_{\text{seq-PAV}}(A, k) = \max_{W \in \mathcal{R}_{\text{seq-PAV}}(A, k)} \Delta_{\text{seq-PAV}}(A, k, W).$$

Further, we define $\Delta_{\text{seq-PAV}}(k)$ as:

$$\Delta_{\text{seq-PAV}}(k) = \sup_{A \in \mathcal{A}} \Delta_{\text{seq-PAV}}(A, k).$$

Our first lemma shows a close relation between the proportionality guarantee of Sequential Proportional Approval Voting and the value $\Delta_{\text{seq-PAV}}(k)$. This is very useful, as the definition of $\Delta_{\text{seq-PAV}}(k)$ is not based on ℓ -large cohesive groups, and thus it is much easier to handle.

Lemma 1. *For each $k, \ell \in \mathbb{N}_+$ the k -proportionality degree of Sequential PAV satisfies:*

$$d_{\text{seq-PAV}}(\ell, k) \geq \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)} - 1.$$

For each $k, \ell \in \mathbb{N}_+$ the proportionality degree of Sequential PAV satisfies:

$$d_{\text{seq-PAV}}(\ell) \leq \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)}.$$

As an immediate corollary of [Lemma 1](#) we obtain an almost tight estimation of the proportionality guarantee of Sequential Proportional Approval Voting.

Corollary 2. *The proportionality guarantee of Sequential Proportional Approval Voting is $d_{\text{seq-PAV}}(\ell) = \inf_k \frac{\ell}{k \cdot \Delta_{\text{seq-PAV}}(k)} - \epsilon$ for some $\epsilon \in [0, 1]$.*

In the remaining part of this section we will assess the expression $h_{\text{seq-PAV}}(k) = k \cdot \Delta_{\text{seq-PAV}}(k)$, for different values of k . According to [Lemma 1](#) this will give us the lower bound on the k -proportionality degree and an upper bound on the proportionality degree of Sequential PAV.

4.3 An Exact Linear Program for Assessing Proportionality Degree

For each committee size k we can compute $h_{\text{seq-PAV}}(k)$ by solving an appropriately constructed linear program; such an LP for each committee size k finds a profile A for which $k \cdot \Delta_{\text{seq-PAV}}(A, k)$ is maximal. Designing such an LP however requires some care, since ideally its size should not depend on the number of voters nor the number of candidates. Our first observation is that we can consider only the profiles with k candidates (all of which form a winning committee). Indeed, if we take a profile A that maximizes $k \cdot \Delta_{\text{seq-PAV}}(A, k)$, then we can remove from A all candidates which are not members of the winning committee. After this removal we obtain a profile that still witnesses the maximality of $k \cdot \Delta_{\text{seq-PAV}}(A, k)$. Thus, we will assume that in the profile that we look for there are k candidates, and we will represent them as integers from $[k]$. Further, without loss of generality we will assume that the candidates are added to the winning committee in the order of their corresponding numbers: candidate 1 is added first, candidate 2 is added second, etc.

Second, we observe that we do not need to represent each voter, we can rather cluster the voters into the groups having the same approval sets and for each possible approval set $T \subseteq [k]$ we are only interested in the proportion of the voters who have this approval set. Such a proportion will be denoted by the variable x_T . Now we can use a Linear Program

k	lower-bound	k	lower-bound	k	lower-bound	k	lower-bound
1	1.0 ℓ -1	6	0.8169 ℓ -1	11	0.7773 ℓ -1	16	0.7589 ℓ -1
2	1.0 ℓ -1	7	0.8064 ℓ -1	12	0.7719 ℓ -1	17	0.7563 ℓ -1
3	0.8888 ℓ -1	8	0.7979 ℓ -1	13	0.7684 ℓ -1	18	0.7540 ℓ -1
4	0.8571 ℓ -1	9	0.7888 ℓ -1	14	0.7647 ℓ -1	19	0.7522 ℓ -1
5	0.8372 ℓ -1	10	0.7825 ℓ -1	15	0.7616 ℓ -1	20	0.7503 ℓ -1

Table 1: A lower-bound on the k -proportionality degree of Sequential PAV.

k	lower-bound	k	lower-bound	k	lower-bound	k	lower-bound
1	1.0 ℓ -1	6	0.8131 ℓ -1	11	0.7643 ℓ -1	16	0.7441 ℓ -1
2	1.0 ℓ -1	7	0.7952 ℓ -1	12	0.7594 ℓ -1	17	0.7416 ℓ -1
3	0.8888 ℓ -1	8	0.7871 ℓ -1	13	0.7548 ℓ -1	18	0.7396 ℓ -1
4	0.8461 ℓ -1	9	0.7771 ℓ -1	14	0.7512 ℓ -1	19	0.7371 ℓ -1
5	0.8307 ℓ -1	10	0.7705 ℓ -1	15	0.7476 ℓ -1	20	0.7348 ℓ -1

Table 2: A lower-bound on the k -proportionality degree of Sequential PAV, as computed by the Linear Program from Figure 3 in Appendix D.

formulation given by Sánchez-Fernández et al. [24] to compute $h_{\text{seq-PAV}}(k)$.⁵ In Figure 1 we give the LP; for a logical expression E , we set $\mathbf{1}[E]$ to be 1 if E is true, and 0 otherwise.⁶

$$\begin{aligned}
& \text{maximize} && k \cdot \sum_{T \in S([k])} x_T \cdot \frac{\mathbf{1}[k \in T]}{|T|} \\
& \text{subject to:} \\
& \text{(a):} && \sum_{T \in S([k])} x_T = 1 \\
& \text{(b):} && \sum_{T \in S([k])} \frac{x_T \cdot \mathbf{1}[i \in T]}{|T \cap [i]|} \geq \sum_{T \in S([k])} \frac{x_T \cdot \mathbf{1}[j \in T]}{|T \cap ([i-1] \cup \{j\})|} && i, j \in [k]; j > i \\
& \text{(c):} && x_T \geq 0 && T \in S([k])
\end{aligned}$$

Figure 1: Linear programming (LP) formulation for computing $h_{\text{seq-PAV}}(k) = k \cdot \Delta_{\text{seq-PAV}}(k)$.

The expression that we maximize is exactly $k \cdot \Delta_{\text{seq-PAV}}(A, k)$. Indeed, $\frac{\mathbf{1}[k \in T]}{|T|}$ is the marginal increase of the PAV score coming from a voter who approves T , as a result of adding candidate k to committee $[k-1]$. The constraint (a) ensures that the proportions of clustered voters sum up to 1. The constraint (b) ensures that in the i -th step of Sequential PAV, the marginal increase of the PAV score due to adding i to committee $[i-1]$ is at least as large as due to adding $j > i$ to $[i-1]$. These constraints ensure that the candidates are indeed added in the order $1, 2, \dots, k$.

We computed the above program for $k \leq 20$; the resulting lower bounds for the k -proportionality degree of Sequential PAV are given in Table 1.

⁵Sánchez-Fernández et al. [24] used a very similar LP to compute the smallest committee size k for which Seq-PAV violates JR. Our so-far analysis shows that a very similar LP can be used to obtain much stronger guarantees that lead to roughly tight estimates of the proportionality degree of Seq-PAV.

⁶Note that the LP is guaranteed to return a rational solution, which can be converted in an exact profile.

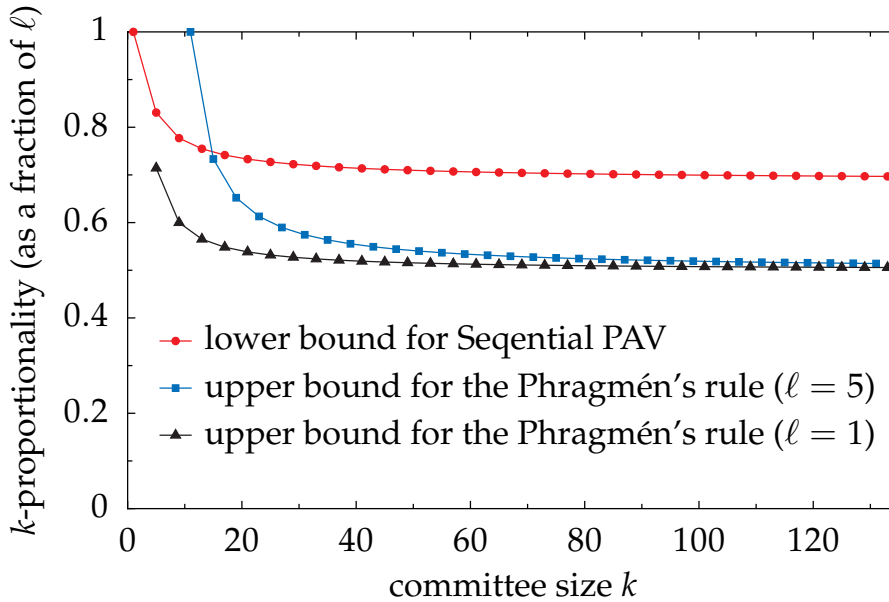


Figure 2: The lower bounds on the k -proportionality degree of Sequential PAV as computed by the LP from Figure 3. The plot compares these lower-bounds with the bounds for Phragmén’s Sequential Rule, as given in Proposition 1.

Let us compare the values from Table 1 with the bounds for Phragmén’s Sequential Rule from Theorem 2 and Proposition 1. Since the bounds for Phragmén’s Sequential Rule are more accurate for smaller groups of voters (when k is significantly larger than ℓ), let us consider the committee size $k = 10$, and an ℓ -large group consisting of 10% of the voters (thus, $\ell = 1$). In this case, Proposition 1 says that the k -proportionality degree of Phragmén’s Sequential Rule cannot be higher than 0.625ℓ , while the k -proportionality degree of Sequential PAV equals at least 0.7825ℓ . This proves that Sequential PAV has better proportionality guarantees than Phragmén’s Sequential Rule for $k = 10$; a similar comparison can be performed for $10 \leq k \leq 20$.

4.4 A Relaxed LP for Assessing Proportionality Degree

Unfortunately, the LP from Figure 1 is exponential in k , so we were able to provide accurate bounds only for $k \leq 20$. In Appendix D we describe another LP, that runs in polynomial time in k . The new LP does not compute the exact value of $h_{\text{seq-PAV}}(k) = k \cdot \Delta_{\text{seq-PAV}}(k)$, but rather an upper bound for $h_{\text{seq-PAV}}(k)$ (and thus the lower bound on the proportionality degree). However, for small values of k ($k \leq 20$) this lower bound is very accurate. One can observe that these values are very close to those from Table 1, computed by the exact LP from Figure 1. For example, for $k = 20$, these values are 0.7085 and 0.694, respectively. In Figure 2, we depict these lower-bounds for $k \leq 200$, and we observe they are considerably better than the bounds for Phragmén’s Sequential Rule. Additionally, by analyzing Table 1 and Table 2 we see that the differences between the values computed by the two LP-s increase with k ; this allows us to conjecture that the actual k -proportionality degree of Sequential PAV can be even better than the bounds presented in Figure 2 (note that the bounds from Figure 2 were computed with the relaxed LP, thus these bounds are firm but possibly not optimal as if they were computed by the exact LP).

Summarizing, Sequential PAV should be preferred over PAV and over Phragmén’s Sequential Rule, when proportionality and committee enlargement monotonicity are the primary requirements (PAV has a better proportionality degree but fails committee enlargement monotonicity). On the other hand, Sequential PAV does not satisfy strong unanimity— if a decision maker considers this a serious flaw, then Phragmén’s Sequential Rule is a good alternative.

5 Conclusion

A number of results have been deferred to the appendix. In particular, in [Appendix F](#) we establish bounds on the proportionality degrees for Thiele methods. In [Appendix G](#) we focussed on estimating the utilitarian efficiency of the considered rules. Together with the results estimating proportionality degree, this allows us to express the studied rules as implementing certain tradeoffs between proportionality and utilitarian efficiency.

In summary, this paper quantifies the level of proportionality and the utilitarian efficiency for a number of multiwinner rules. We provide general tools that allow to estimate the proportionality degree and the utilitarian efficiency for a large subclass of voting rules. Our results show, in particular, a spectrum of rules implementing various tradeoffs between proportionality and utilitarian efficiency.

Our conclusions can be summarized as follows. Phragmén’s Sequential Rule is roughly half as proportional as PAV. Sequential PAV lies, in terms of proportionality, between Phragmén’s Sequential Rule and PAV. Further, some rules such as the λ -Thiele rule for $\lambda(i) = 1/\sqrt{i}$ offer a significantly better utilitarian efficiency than PAV, at the cost of a moderate loss of proportionality.

We consider the following open questions particularly interesting and important: can we find a rule that combines the virtues of Phragmén’s Sequential Rule and of PAV? Such a rule should in particular (i) satisfy Pareto efficiency (which PAV satisfies and Phragmén’s Sequential Rule violates [\[18\]](#)), (ii) satisfy strong unanimity (which Phragmén’s Sequential Rule satisfies and PAV violates), and (iii) have a high proportionality degree. It is tempting to suggest a rule that first takes all unanimously approved candidates and complements the committee by running PAV, yet such rule looks a bit ad hoc, and its definition is specifically tailored for strong unanimity.⁷ Thus, finding stronger properties that better capture the idea exposed by strong unanimity is essential.

Second, can we use Phragmén’s Sequential Rule to compare any two committees? The definition of the rule allows only for finding winning committees. Comparing is, however, sometimes essential, e.g., when there are some external constraints put on the committee and when the goal is to find the best possible committee subject to these constraints. While Phragmén proposed a few other (seemingly similar) rules based on global optimization goals (e.g., Phragmén’s Maximal Rule), these rules offer much worse proportionality and utilitarian efficiency, so the question is still open.

Third, what is the maximum degree of proportionality that can be guaranteed by a rule that satisfies committee-enlargement monotonicity?

Acknowledgments

The author was supported by Poland’s National Science Center grant UMO-2019/35/B/ST6/02215.

⁷For profiles where there are candidates approved by “almost” all the voters, the rule would behave just like PAV. This is not intended as strong unanimity is only an example of a more complex set of requirements one would impose on an election rule.

References

- [1] H. Aziz, M. Brill, V. Conitzer, E. Elkind, R. Freeman, and T. Walsh. 2017. Justified Representation in Approval-Based Committee Voting. *Social Choice and Welfare* 48, 2 (2017), 461–485.
- [2] H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. 2018. On the Complexity of Extended and Proportional Justified Representation. In *Proceedings of the 32nd Conference on Artificial Intelligence (AAAI-2018)*.
- [3] M. Balinski and H. P. Young. 1982. *Fair Representation: Meeting the Ideal of One Man, One Vote*. Yale University Press. (2nd Edition [with identical pagination], Brookings Institution Press, 2001).
- [4] J. Behrens, A. Kistner, A. Nitsche, and B. Swierczek. 2014. *The Principles of Liquid-Feedback*.
- [5] M. Brill, R. Freeman, S. Janson, and M. Lackner. 2017. Phragmén’s Voting Methods and Justified Representation. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*. 406–413.
- [6] M. Brill, J.-F. Laslier, and P. Skowron. 2018. Multiwinner Approval Rules as Apportionment Methods. *Journal of Theoretical Politics* 30, 3 (2018), 358–382.
- [7] Jeff Burdges, Alfonso Cevallos, Peter Czaban, Rob Habermeier, Syed Hosseini, Fabio Lama, Handan Kiliç Alper, Ximin Luo, Fatemeh Shirazi, Alistair Stewart, and Gavin Wood. 2020. Overview of Polkadot and its Design Considerations. *CoRR* abs/2005.13456 (2020). arXiv:2005.13456 <https://arxiv.org/abs/2005.13456>
- [8] Y. Cabannes. 2004. Participatory budgeting: a significant contribution to participatory democracy. *Environment and Urbanization* 16, 1 (2004), 27–46.
- [9] M. Dummett. 1984. *Voting Procedures*. Oxford University Press.
- [10] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. 2001. Rank aggregation methods for the Web. In *Proceedings of the 10th International World Wide Web Conference*. 613–622.
- [11] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. 2017. Properties of Multiwinner Voting Rules. *Social Choice and Welfare* 48, 3 (2017), 599–632.
- [12] B. Fain, K. Munagala, and N. Shah. 2018. Fair Allocation of Indivisible Public Goods. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC-2018)*. 575–592.
- [13] D. Felsenthal and M. Machover. 1998. *The Measurement of Voting Power*. Edward Elgar Publishing.
- [14] A. Goel, A. Krishnaswamy, S. Sakshuwong, and T. Aitamurto. 2016. Knapsack Voting: Voting mechanisms for Participatory Budgeting. (2016). Manuscript.
- [15] B. Grofman. 2016. Perspectives on the Comparative Study of Electoral Systems. *Annual Review of Political Science* 19 (2016), 1–23.
- [16] S. Janson. 2016. *Phragmén’s and Thiele’s election methods*. Technical Report arXiv:1611.08826 [math.HO]. arXiv.org. <http://arxiv.org/abs/1611.08826>

- [17] Y. Koriyama, J.-F. Laslier, A. Macé, and R. Treibich. 2013. Optimal Apportionment. *Journal of Political Economy* 121, 3 (2013), 584–608.
- [18] M. Lackner and P. Skowron. 2020. Utilitarian welfare and representation guarantees of approval-based multiwinner rules. *Artificial Intelligence* 288 (2020), 103366.
- [19] Martin Lackner and Piotr Skowron. 2021. Consistent Approval-Based Multi-Winner Rules. *Journal of Economic Theory* 192, 105173.
- [20] A. Lijphart and B. Grofman. 1984. *Choosing an Electoral System: Issues and Alternatives*. Praeger, New York.
- [21] T. Lu and C. Boutilier. 2011. Budgeted Social Choice: From Consensus to Personalized Decision Making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*. 280–286.
- [22] T. Lu and C. Boutilier. 2015. Value Directed Compression of Large-Scale Assignment Problems. In *Proceedings of the 29th Conference on Artificial Intelligence (AAAI-2015)*. 1182–1190.
- [23] F. Pukelsheim. 2014. *Proportional Representation: Apportionment Methods and Their Applications*. Springer.
- [24] L. Sánchez-Fernández, E. Elkind, M. Lackner, N. Fernández, J. A. Fisteus, P. Basanta Val, and P. Skowron. 2017. Proportional justified representation. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI-2017)*. 670–676.
- [25] L. Shapley and M. Shubik. 1954. A Method for Evaluating the Distribution of Power in a Committee System. *American Political Science Review* 48, 3 (1954), 787–792.
- [26] P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. 2017. Proportional Rankings. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI-2017)*. 409–415.
- [27] H. Steinhaus. 1948. A Note on Preference Aggregation. *Econometrica* 16, 1 (1948), 101–104.

Piotr Skowron
University of Warsaw
Warsaw, Poland
Email: p.skowron@mimuw.edu.pl

A Proofs Omitted From the Main Text

A.1 Proof of **Theorem 2**

Theorem 2. *The proportionality degree of Phragmén Sequential Rule satisfies $d_{\text{Phrag}}(\ell) \geq \frac{\ell-1}{2}$.*

Proof. In the proof we will be using the alternative definition of Phragmén's Sequential Rule, using the concept of virtual money (credits).

Let W be a committee returned by Phragmén's Sequential Rule. For the sake of contradiction, let us assume that there exists an ℓ -large group of voters V with the average satisfaction lower than $\min(|\bigcap_{i \in V} A(i)|, \frac{\ell-1}{2})$. We set $n_V = |V|$.

Observe that the rule stops after at least k time units. Indeed, the total amount of credits earned by all the voters in the first k time units is equal to kn . This allows to buy at most k candidates.

The total amount of credits collected by the voters from V in the first k time units is equal to $n_V k$. The number of credits left to V after the whole committee is elected is at most equal to n (otherwise, these credits would have been spent earlier for buying an additional candidate approved by all the voters from V ; such a candidate would exist as $\text{sat}_V(A, W) < |\bigcap_{i \in V} A(i)|$). Thus, the voters from V spent at least $n_V k - n$ credits for buying candidates they approve.

We now move to the central argument of the proof—we will estimate how much on average a voter from V pays for buying a committee member. We analyze the following potential function.

Let $p_t(i)$ denote the number of credits held by voter $i \in V$ at time t ; further, let $p_t(\text{av}) = \frac{1}{n_V} \sum_{i \in V} p_t(i)$. For a time t we define the potential value ϕ_t as:

$$\phi_t = \sum_{i \in V} (p_t(i) - p_t(\text{av}))^2.$$

We will now analyze how the potential value changes over time. First, observe that the potential value remains unchanged when the number of credits of each voter is incremented (earning credits does not change the potential value). Next, we analyze what happens when the voters use their credits to pay for the committee members that they approve. Consider the time moment t when a new committee member c is selected. The voters who approve c pay for her with their credits; furthermore, a voter who pays uses all her available credits. We consider these voters separately, one by one. Consider a voter $j \in V$ and assume that her number of credits decreases to 0 (since we consider the voters separately, the number of credits held by each other voter remains unchanged). In such a case the average $p_t(\text{av})$ decreases by $\frac{p_t(j)}{n_V}$. We assess the change in the potential value:

$$\begin{aligned} \Delta_\phi &= \sum_{i \in V, i \neq j} \left(p_t(i) - \left(p_t(\text{av}) - \frac{p_t(j)}{n_V} \right) \right)^2 + \left(0 - \left(p_t(\text{av}) - \frac{p_t(j)}{n_V} \right) \right)^2 - \sum_{i \in V} (p_t(i) - p_t(\text{av}))^2 \\ &= \sum_{i \in V} \left(p_t(i) - \left(p_t(\text{av}) - \frac{p_t(j)}{n_V} \right) \right)^2 - \sum_{i \in V} (p_t(i) - p_t(\text{av}))^2 \\ &\quad + \left(p_t(\text{av}) - \frac{p_t(j)}{n_V} \right)^2 - \left(p_t(j) - \left(p_t(\text{av}) - \frac{p_t(j)}{n_V} \right) \right)^2 \\ &= \sum_{i \in V} \frac{p_t(j)}{n_V} \left(2p_t(i) - 2p_t(\text{av}) + \frac{p_t(j)}{n_V} \right) - p_t(j)^2 + 2p_t(j) \left(p_t(\text{av}) - \frac{p_t(j)}{n_V} \right) \end{aligned}$$

Now, observe that $\sum_{i \in V} (2p_t(i) - 2p_t(\text{av})) = 0$, thus:

$$\Delta_\phi = \frac{p_t(j)^2}{n_V} + p_t(j) \left(2p_t(\text{av}) - \frac{2p_t(j)}{n_V} - p_t(j) \right) = p_t(j) \left(2p_t(\text{av}) - \frac{p_t(j)(n_V + 1)}{n_V} \right).$$

We further observe that in each time t we have $p_t(\text{av}) \leq \frac{n}{n_V} \leq \frac{k}{\ell}$ as otherwise the sum of credits within the group would be greater than n , and such credits would be earlier spent on buying a committee member who is approved by all the voters within the group.

Let us now interpret the above calculations. Intuitively, we will argue that a voter from V will on average pay at most $\frac{2n_V}{n_V+1} \cdot \frac{k}{\ell}$ for a committee member. Let us set $x_{t,j} = p_t(j) - \frac{2n_V}{n_V+1} \cdot \frac{k}{\ell}$. Then,

$$\Delta_\phi = p_t(j) \left(2p_t(\text{av}) - 2\frac{k}{\ell} - \frac{x_{t,j}(n_V + 1)}{n_V} \right) \leq -\frac{x_{t,j}(n_V + 1)}{n_V} \cdot p_t(j).$$

If $x_{t,j} > 0$, then ϕ decreases by at least $2|x_{t,j}| \cdot \frac{k}{\ell}$. Similarly, if $x_{t,j} \leq 0$, then ϕ increases by at most $2|x_{t,j}| \cdot \frac{k}{\ell}$. Since the potential value is always non-negative, we infer that the values of $x_{t',j'}$ for $j' \in V$ and t' are on average lower than or equal to 0. Recall that $p_t(j)$ in the definition of $x_{t,j}$ is equal to how much voter j pays for the selected candidate. Thus, the voters from V on average pay at most $\frac{2n_V}{n_V+1} \cdot \frac{k}{\ell} < \frac{2k}{\ell}$ for a committee member. Consequently, the average number of committee members that the voters from V approve equals at least:

$$\frac{n_V k - n}{\frac{2k}{\ell} \cdot n_V} = \frac{k}{\frac{2k}{\ell}} - \frac{n}{\frac{2k}{\ell} \cdot n_V} \geq \frac{\ell}{2} - \frac{n}{\frac{2k}{\ell} \cdot \frac{n\ell}{k}} = \frac{\ell - 1}{2}.$$

This leads to a contradiction and so completes the proof. \square

A.2 Proof of Proposition 1

Proposition 1. *The k -proportionality degree of Phragmén's Sequential Rule satisfies the following inequality: For each $k, \ell \in \mathbb{N}$, $\ell < \frac{k}{2}$ and k divisible by ℓ we have $d_{\text{Phrag}}(\ell, k) \leq \frac{\ell}{2} \cdot \frac{2k-2\ell+2}{2k-3\ell}$.*

Proof. Let us fix two natural numbers, $\ell, k \in \mathbb{N}$; $\ell < \frac{k}{2}$ and k is divisible by ℓ . Let $x = \frac{2(k-\ell)}{\ell} - 1$ and let L be an integer divisible by ℓ , x , and by k . We construct a profile with $n = \frac{Lk}{\ell}$ voters and $2k + \ell$ candidates $C = \{b_1, \dots, b_k, c_1, \dots, c_k, d_1, \dots, d_\ell\}$, as follows:

$$\begin{aligned} N(b_1) &= \{1, 2, \dots, L/x\} \cup \{L+1, L+2, \dots, n - L/x\}, \\ N(b_2) &= \{L/x+1, \dots, 2L/x\} \cup \{L+1, L+2, \dots, n - 2L/x\}, \\ &\dots \\ N(b_x) &= \{L - L/x + 1, \dots, L\} \cup \{L+1, \dots, n - L\}, \\ N(b_{x+1}) &= \{1, 2, \dots, L/x\} \cup \{L+1, \dots, n - L\}, \\ N(b_{x+2}) &= \{L/x+1, \dots, 2L/x\} \cup \{L+1, \dots, n - L\}, \\ &\dots \\ N(b_k) &= (\text{a cyclic shift of } \{1, 2, \dots, L/x\}) \cup \{L+1, \dots, n - L\} \\ N(c_1) &= \dots = N(c_k) = \{n - L, n - L + 1, \dots, n\} \\ N(d_1) &= \dots = N(d_\ell) = \{1, \dots, L\}. \end{aligned}$$

In particular, candidate b_1 is approved by $n - L$ voters in total, b_2 is approved by $n - L - L/x$ voters, and b_3 by $n - L - 2L/x$ voters. Each candidate from $\{b_x, \dots, b_k\}$ is approved by

the same $n - 2L$ voters from $\{L + 1, \dots, n\}$ and by some L/x voters from $\{1, \dots, L\}$, which cyclically shift. The voters from $\{1, \dots, L\}$ who approve b_i , $i \geq x$ are those who are right after the voters who approve b_{i-1} -th candidate, unless those who approve b_{i-1} form the last segment, i.e., $\{L - L/x + 1, \dots, L\}$. In such a case, the voters from $\{1, \dots, L\}$ who approve b_i are exactly $\{1, 2, \dots, L/x\}$.

Let $t = \frac{k}{k-\ell}$; this ensures that $t(n - L) = \frac{k}{k-\ell} (n - n\frac{\ell}{k}) = n$. In our profile Phragmén's Sequential Rule selects b_1 first at time t . Indeed, at this time the group of voters $N(b_1)$ collects $(n - L)t = n$ credits. At time $2t$ each voter from $\{L/x + 1, \dots, 2L/x\}$ has already $2t$ credits, and each voter from $\{L + 1, L + 2, \dots, n - 2L/x\}$ has t credits; altogether, they have n credits, thus b_2 is selected second. By a similar reasoning, we infer that in the first x steps candidates b_1, \dots, b_x will be selected by Phragmén's Sequential Rule, and the last one of them will be selected at time xt . At this time the rule would also select one candidate from $\{c_1, \dots, c_k\}$. Indeed, at time xt the voters from $n - L, n - L + 1, \dots, n$ have the following number of credits:

$$\frac{L}{x} \cdot t \cdot (1 + 2 + \dots + x) = \frac{L}{x} \cdot \frac{k}{k-\ell} \cdot \frac{x(x+1)}{2} = \frac{n\ell(x+1)}{2(k-\ell)} = n.$$

Let us now analyze what happens in the next steps. First, let us consider how Phragmén's Sequential Rule would behave if there were no candidates from $\{c_1, \dots, c_k\}$. At time $(x+1)t$ voters from $\{1, 2, \dots, L/x\}$ have xt credits each. Similarly, each voter from $\{L/x + 1, \dots, 2L/x\}$ has $(x-1)t$ credits, each voter from $\{2L/x + 1, \dots, 3L/x\}$ has $(x-2)t$ credits, etc. The amount of credits held by the voters from $\{1, 2, \dots, L\}$ altogether at time $(x+1)t$ is:

$$\frac{L}{x} \cdot t_1 \cdot (1 + 2 + \dots + x) = n.$$

At the same time voters from $N(b_{x+1})$ have n credits altogether, thus, b_{x+1} can be selected next, before any candidate from $D = \{d_1, \dots, d_k\}$ is chosen⁸. Similarly, at time $(x+2)t$ each voter from $\{1, 2, \dots, L/x\}$ has t credits, each from $\{L/x + 1, \dots, 2L/x\}$ has xt credits, each from $\{2L/x + 1, \dots, 3L/x\}$ has $(x-1)t$ credits, etc.; altogether they have at most n credits, and the voters from $N(b_{x+2})$ have exactly n ; thus b_{x+2} can be selected next. Through a similar reasoning we conclude that in the first k steps Phragmén's Sequential Rule would select candidates $\{b_1, \dots, b_k\}$.

Now, consider the candidates from $\{c_1, \dots, c_k\}$. After candidates b_1, \dots, b_x are selected, candidates from $\{c_1, \dots, c_k\}$ are approved only by voters who do not approve other remaining candidates. Thus, their selection does not interfere with the relative order of selecting the other candidates. Further, observe that over time $(x+1)t$ the last L voters collect the following number of credits:

$$(x+1)tL = \frac{2(k-\ell)}{\ell} \cdot \frac{k}{k-\ell} \cdot n\frac{\ell}{k} = 2n.$$

Thus, every $\frac{(x+1)t}{2}$ time moments the rule will select one candidate from $\{c_1, \dots, c_k\}$. Consequently, in the first $t(k - \lfloor \frac{2k-2x}{x+3} \rfloor)$ time moments, the rule will select at least $\lfloor \frac{2k-2x}{x+3} \rfloor + 1$ candidates from $\{c_1, \dots, c_k\}$. Indeed, the first candidate will be selected after the first xt time moments. After the remaining $t(k - x - \lfloor \frac{2k-2x}{x+3} \rfloor)$ time moments the number of candidates from $\{c_1, \dots, c_k\}$ that will be selected is equal to:

$$\frac{t(k - x - \lfloor \frac{2k-2x}{x+3} \rfloor)}{\frac{(x+1)t}{2}} \geq \frac{2(k - x - \frac{2k-2x}{x+3})}{x+1} = \frac{2k-2x}{x+3} \geq \lfloor \frac{2k-2x}{x+3} \rfloor.$$

⁸Here we assume adversarial tie breaking, yet the construction can be strengthened so that it does not depend on the particular tie-breaking mechanism.

Next, observe that:

$$\left\lfloor \frac{2k - 2x}{x + 3} \right\rfloor + 1 \geq \left\lfloor \frac{2k - \frac{4k-6\ell}{\ell}}{\frac{2k-3\ell}{\ell} + 3} \right\rfloor + 1 = \left\lfloor \frac{2k\ell - 4k + 6\ell}{2k} \right\rfloor \geq \ell - 1.$$

Consequently, the winning committee W has at most $k - \ell + 1$ candidates from $\{b_1, \dots, b_{k-1}, b_k\}$ and the other committee members are from $\{c_1, \dots, c_{k-1}, c_k\}$.

The group of voters $V = \{1, \dots, L\}$ is ℓ -cohesive. Let us assess their average number of representatives. Observe that, except for candidates from $\{c_1, \dots, c_{k-1}, c_k\}$, each candidate from the selected committee is approved by exactly $\frac{L}{x}$ voters of V . Thus:

$$\frac{1}{|V|} \sum_{i \in V} |W \cap A(i)| = \frac{1}{L} \cdot \frac{L}{x} \cdot (k - \ell + 1) = \frac{\ell}{2} \cdot \frac{2k - 2\ell + 2}{2k - 3\ell}.$$

□

A.3 Proof of Proposition 2

Proposition 2. *The proportionality degree of Phragmén Maximal Rule is $d_{\max\text{-Phrag}}(\ell) \leq 1$.*

Proof. Consider the profile constructed as follows. The voters are divided into k equal-size groups: $N = N_1 \cup N_2 \cup \dots \cup N_k$; intuitively $n \gg k$. There is a set W_1 of k candidates approved by all except k voters—from each group N_i we pick one voter who does not approve of any candidate from W_1 . Further, there are k candidates, $W_2 = \{c_1, \dots, c_k\}$, such that for each $i \in [k]$, $N(c_i) = N_i$. Clearly, Phragmén Maximal Rule selects W_2 as this committee allows to distribute the load evenly among the voters. With W_2 each voter will have a single representative. On the other hand, for each ℓ the group $N(W_1)$ is $(\ell - 1)$ -large and has more than ℓ commonly approved candidates. Intuitively, for this profile W_1 should be chosen (in particular, it would be the single committee returned by Phragmén's Sequential Rule). □

A.4 Proof of Lemma 1

Lemma 1. *The k -proportionality degree of Sequential PAV satisfies:*

$$d_{\text{seq-PAV}}(\ell, k) \geq \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)} - 1.$$

For each $k \in \mathbb{N}_+$ the proportionality degree of Sequential PAV satisfies:

$$d_{\text{seq-PAV}}(\ell) \leq \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)}.$$

Proof. First, we prove that the k -proportionality guarantee of Sequential Proportional Approval Voting satisfies $d_{\text{seq-PAV}}(\ell, k) \geq \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)} - 1$. For the sake of contradiction let us assume that this is not the case, and that there exists an approval-based profile A with n voters, a committee size k , an ℓ -large group of voters V , and a winning committee $W \in \mathcal{R}_{\text{seq-PAV}}(A, k)$ such that:

$$\frac{1}{|V|} \sum_{i \in V} |W \cap A(i)| < \min \left(\left| \bigcap_{i \in V} A(i) \right|, \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)} - 1 \right).$$

By the pigeonhole principle, there exists a candidate $c \notin W$ who is approved by all voters from V . Let us now estimate by how much adding c to W increases the PAV score. Fixing

the average number of representatives that the voters from V have in W , it is straightforward to check that the increase would be the smallest when each voter in V has roughly the same number of representatives. Thus, adding c to W would increase the PAV score by more than

$$\frac{|V|}{\ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)} - 1 + 1} = |V| \cdot \frac{k}{\ell} \cdot \Delta_{\text{seq-PAV}}(k) \geq n \Delta_{\text{seq-PAV}}(k).$$

Clearly, adding c to any subset of W would result in at least the same increase of the PAV score. Since Sequential Proportional Approval Voting always selects a candidate that increases the PAV score of the committee most, and since c was not selected, we infer that $\Delta_{\text{seq-PAV}}(A, k, W) > \frac{1}{n} \cdot n \Delta_{\text{seq-PAV}}(k) = \Delta_{\text{seq-PAV}}(k)$, a contradiction.

Second, let us fix $\epsilon > 0$ and $k \in \mathbb{N}$. We will construct an approval-based profile A_{hard} that witnesses that $d_{\text{seq-PAV}}(\ell) \leq \ell \cdot \frac{1}{k \cdot \Delta_{\text{seq-PAV}}(k)} + \epsilon$. Let A be an approval-based profile and W be a size- k committee winning in A such that $\Delta_{\text{seq-PAV}}(A, k, W) > \Delta_{\text{seq-PAV}}(k) - \epsilon_1$, for some ϵ_1 ; the value of ϵ_1 will become clear later on. Let us fix an integer L —intuitively, L is a large number; the exact value of L will also become clear from the further part of the proof. Let $n = |A|$. We construct A_{hard} by appending L independent copies of A (we clone both the voters and the candidates) and $y = \frac{Ln\ell}{Lk-\ell}$ voters who all approve some ℓ candidates, not approved by any voter from any copy of A —let us denote this set of y voters by V . Let $n' = |A_{\text{hard}}|$; clearly $n' = Ln + \frac{Ln\ell}{Lk-\ell} = \frac{L^2nk}{Lk-\ell}$. We set the required committee size to $k' = Lk$. Observe that V is ℓ -cohesive. Indeed, all voters from V approve common ℓ candidates; further the relative size of V , y/n' , is equal to

$$\frac{y}{n'} = \frac{Ln\ell}{Lk-\ell} \cdot \frac{Lk-\ell}{L^2nk} = \frac{\ell}{Lk} = \frac{\ell}{k'}.$$

Now, we show that the voters from V have on average less than $\frac{\ell}{k \cdot \Delta_{\text{seq-PAV}}(k)} + \epsilon$ representatives in some winning committee for A_{hard} . Towards a contradiction, let us assume that this is not the case. Since the voters in V are identical, this means that each such a voter has more than $\frac{\ell}{k \cdot \Delta_{\text{seq-PAV}}(k)} + \epsilon$ representatives in each winning committee. Thus, for any winning committee, when the last representative of the voters from V was added, the PAV-score of the committee increased by at most:

$$\Delta = y \cdot \frac{k \cdot \Delta_{\text{seq-PAV}}(k)}{\ell + \epsilon k \cdot \Delta_{\text{seq-PAV}}(k)}$$

The above expression can be written as

$$\Delta = y \cdot \left(\frac{k \cdot \Delta_{\text{seq-PAV}}(k)}{\ell} - \epsilon_2 \right),$$

where ϵ_2 is some parameter dependent on ϵ and k . Further, observe that:

$$\begin{aligned} \lim_{L \rightarrow \infty} \Delta &= \lim_{L \rightarrow \infty} y \cdot \left(\frac{k \cdot \Delta_{\text{seq-PAV}}(k)}{\ell} - \epsilon_2 \right) = \lim_{L \rightarrow \infty} \frac{Ln\ell}{Lk-\ell} \cdot \left(\frac{k \cdot \Delta_{\text{seq-PAV}}(k)}{\ell} - \epsilon_2 \right) \\ &= \frac{n\ell}{k} \cdot \left(\frac{k \cdot \Delta_{\text{seq-PAV}}(k)}{\ell} - \epsilon_2 \right) = n \Delta_{\text{seq-PAV}}(k) - \epsilon_2 \cdot \frac{n\ell}{k}. \end{aligned}$$

Thus, there exists large enough L and small enough ϵ_1 such that:

$$\Delta < n \Delta_{\text{seq-PAV}}(k) - n \epsilon_1 < n \Delta_{\text{seq-PAV}}(A, k, W).$$

These are exactly the values of L and ϵ_1 that we use in our construction. In other words, the increase of the PAV score due to adding the last representative of V is lower than the increase of the PAV score due to adding the last committee member to W . Thus, clearly there exists a winning committee that consists of some copies of the candidates from W and less than $\frac{\ell}{k \cdot \Delta_{\text{seq-PAV}}(k)} + \epsilon$ candidates from those approved by the voters from V , a contradiction. This completes the proof. \square

B Proportionality degree versus other notions of proportionality

In this section we compare the notion of the proportionality guarantee/degree with other concepts pertaining to proportionality, studied in the literature.

Extended Justified Representation (EJR). EJR [1] requires that each ℓ -large group of voters V with $|\bigcap_{i \in V} A(i)| \geq \ell$ must contain a voter who approves at least ℓ members of the winning committee(s). This property is very natural and interesting, yet it also has certain drawbacks. On the one hand, it seems quite weak—it aims at analyzing how voting rules treat *groups* of voters, yet for each group it only enforces the requirement that there must exist *some* voter in the group who is well represented. On the other hand, it appears very strong—if each voter from V approves $\ell - 1$ members of the winning committee, then V already witnesses that the rule violates the property. Indeed, PAV is almost the only natural known rule satisfying EJR.⁹ Our approach, on the other hand, provides a fine-grained guarantee on how well a given rule represents certain groups of voters, and focuses on the average satisfaction of the voters within the group rather than on the satisfaction of the most satisfied voter.

Yet, there is also a relation between EJR and the proportionality degree. Clearly, if for some $\alpha \in [0, 1]$, the proportionality guarantee of a rule \mathcal{R} satisfies $g_{\mathcal{R}}(\ell, k) \geq \alpha \ell$, then—by the pigeonhole principle— \mathcal{R} also satisfies an α -approximation of EJR. For the reverse direction a weaker implication has been shown by Sánchez-Fernández et al. [24]: if a rule \mathcal{R} satisfies EJR, then it has the proportionality guarantee of \mathcal{R} is $g_{\mathcal{R}}(\ell, k) = \frac{\ell-1}{2}$.¹⁰

Lower Quota. Another approach to investigating proportionality of ABC rules is to analyze how they behave for structured preferences; e.g., when the voters and the candidates can be divided into disjoint groups so that each group of voters approves exactly one group of candidates. Such profiles can be represented as party-list elections, and so the classic notions of proportionality for apportionment methods [3, 23] apply. For instance, EJR generalizes the definition of lower-quota. Informally speaking, for party-list elections lower quota requires that if a population of at least $\ell \cdot n/k$ voters vote for a single party, then this party gets at least ℓ seats in the elected committee. Brill et al. [6] discussed how different ABC rules behave for such restricted preferences, and Lackner and Skowron [19] proved that under certain assumptions the behavior of ABC rules uniquely extends from party-list profiles to general preferences. Our study answers whether the considered rules still behave proportionally for general preferences.

⁹Other rules satisfying EJR are either very similar to PAV (e.g., the local search algorithm for PAV) or very technical and specifically tailored to satisfy the particular property [2].

¹⁰The main argument is based on applying the EJR property recursively: Assume a rule satisfies EJR and consider an ℓ -large group of voters V with $|\bigcap_{i \in V} A(i)| \geq \ell$. Then by EJR, we know that there exists $v \in V$ with at least ℓ representatives. Now, observe that $V \setminus \{v\}$ is an $(\ell - 1)$ -large group and that it is cohesive. Thus, there exists $v' \in V \setminus \{v\}$ with at least $\ell - 1$ representatives, etc.

C Equivalence of the Two Definitions of the Phragmén's Rule

In this section we prove that the money-based procedure described in the first part of [Section 3](#) is equivalent to Phragmén's Sequential Rule.

As we already indicated, we can assume that the original definition of the Phragmén's rule each candidate is associated with n units of load instead of one. Recall that by $\ell_i(j)$ we denote the total load assigned to voter i after the j -th iteration of the Phragmén's rule. Let $\ell_{\max}(j)$ denote the total load of the voter who carries the maximal load after the j -th iteration of the rule, that is $\ell_{\max}(j) = \max_{i \in N} \ell_i(j)$. For each $j \in [k] \cup \{0\}$ we will prove by induction the following two statements:

1. The j -th candidate added by the Phragmén's rule to the winning committee is the same as the j -th candidate added to the winning committee by the money-based procedure.
2. The value $\ell_{\max}(j) - \ell_i(j)$ is equal to the number of credits that voter i is left with after the j -th candidate is added to the winning committee.

Clearly, our induction hypothesis is true for $j = 0$, since at that point each of the two rules had not yet added any candidate to the committee. Assume that the induction hypothesis is true for $j = x$. We will show that it also holds for $j = x + 1$. Let t_j denote the time moment in the money-based procedure that corresponds to adding the j -th candidate to the committee. For each Δ in time $t_x + \Delta$ voter i has the total amount of $\ell_{\max}(x) + \Delta - \ell_i(x)$ money (this follows by the inductive assumption, and by the fact that the voters earn money with the constant speed of one credit per time unit). Thus, the money-based procedure adds candidate c the $(j + 1)$ -th committee member in time $t_x + \Delta$ if:

$$\sum_{i \in N(c)} (\ell_{\max}(x) + \Delta - \ell_i(x)) = n$$

and if for each other candidate c' we have

$$\sum_{i \in N(c')} (\ell_{\max}(x) + \Delta - \ell_i(x)) < n$$

Thus, it finds the smallest possible value $v = \ell_{\max}(x) + \Delta$ such that there exists a candidate c for whom:

$$\sum_{i \in N(c)} (v - \ell_i(x)) = n$$

On the other hand, the Phragmén's rule aims at minimizing $\ell_{\max}(x+1)$. Thus, $\ell_{\max}(x+1)$ will be the minimal value v' for which there exists a candidate c and a load distribution $\{\delta_i\}_{i \in N}$ such that:

$$\sum_{i \in N(c)} (\delta_i + \ell_i(x)) = v', \quad \text{and} \quad \sum_{i \in N(c)} \delta_i = n.$$

These equalities can be reformulated as:

$$\sum_{i \in N(c)} (v' - \ell_i(x)) = \sum_{i \in N(c)} \delta_i = n.$$

Thus, we see that the minimization of v' corresponds to the minimization of v , and so $\ell_{\max}(x+1) = v$. As a result the two rules will pick the same candidate c to the committee, proving the first condition of our inductive hypothesis. Further, the money-based procedure will add this candidate after Δ time units from adding the previous one, where:

$$\Delta = v - \ell_{\max}(x) = \ell_{\max}(x+1) - \ell_{\max}(x).$$

Thus, for each $i \notin N(c)$ we will have that voter i is left with the following amount of money:

$$\ell_{\max}(x) - \ell_i(x) + \Delta = \ell_{\max}(x+1) - \ell_i(x) = \ell_{\max}(x+1) - \ell_i(x+1).$$

Similarly, if $i \in N(c)$, then the number of credits left is equal to zero, but also $\ell_{\max}(x+1) = \ell_i(x+1)$; this proves the second condition of our inductive hypothesis, and so it completes the proof.

D A Relaxed LP for Assessing Proportionality Degree

In this section we describe a new LP that allowed us to compute a lower-bound on the k -proportionality degree of Sequential PAV for $k \leq 200$; we will see that for larger values of k this bound is still considerably better than the upper-bounds for Phragmén's Sequential Rule.

In the new LP we use the same observations as before: (i) we consider profiles with k candidates only, and (ii) we cluster the voters into groups, and for each group we have a variable denoting how large it is as a fraction of all the voters. Specifically, we have the following variables. For each $i \in [k]$ we have a variable a_i that denotes the fraction of voters who approve exactly i from all candidates in total. For $i \in [k], j, p \in [k]_0$ and $p \leq i, j$, we have variable $b_{i,j,p}$ that denotes the fraction of the voters who approve i candidates in total, and that after the j -th step of Sequential PAV approve exactly p from the already selected committee members. Finally, for $i, j, p \in [k]$ and $p \leq i, j$, we have variable $c_{i,j,p}$ that denotes the fraction of voters who approve i candidates in total, and that as a result of the j -th step of Sequential PAV the number of their representatives increased from $p-1$ to p . Finally, for $j \in [k]$ we have a variable d_j that denotes the increase of the PAV score after the j -th step of Sequential PAV, multiplied by $1/n$. The new LP is given in [Figure 3](#).

Let us now explain the constraints in the LP. Constraints (a1)–(a2) and (c1) are intuitive. Constraints (b1)–(b2) say that before Sequential PAV starts the committee is empty (and so every voter in each group has no representatives) and that after Sequential PAV finishes, all the candidates are selected (and so every voter in each group has as many representatives, as the number of approved candidates). Constraint (c2) says that the voters whose number of representatives increased from $p-1$ to p in the j -th step, must have had $p-1$ representatives after the $(j-1)$ -th step of Sequential PAV. Constraints (d1)–(d4) give the natural relations between b - and c - variables. Constraint (e1) encodes the definition of the d -variables. Constraint (e2) is the most tricky, and it is an incarnation of the pigeonhole principle. Consider the expression on the right-hand side of (e2). Consider the voters who approve i candidates in total, and who before the j -th step of Sequential PAV, approve p already elected members. The fraction of such voters is $b_{i,j-1,p}$. Each such a voter approves $i-p$ not yet elected candidates. Adding each such a candidate would increase the score that each such a voter assigns to a committee by $\frac{1}{p+1}$. Thus, the sum in the right-hand side of (e2) gives the total, over all not yet selected candidates, increase of the PAV score due to adding such candidates to the current committee; the whole sum is normalized by $1/n$. There are $k-j+1$ not yet selected candidates. Thus, by the pigeonhole principle, adding at least one of such candidate increases the normalized PAV score by at least the value which

$$\begin{aligned}
& \text{maximize} && k \cdot d_k \\
& \text{subject to:} && \\
& \text{(a1)} : a_i \geq 0 && i \in [k] \\
& \text{(a2)} : \sum_{i \in [k]} a_i = 1 && \\
& \text{(b1)} : b_{i,0,0} = b_{i,k,i} = a_i && i \in [k] \\
& \text{(b2)} : b_{i,k,p} = 0 && p < i \\
& \text{(c1)} : c_{i,j,p} \geq 0 && i, j, p \in [k], p \leq i, j \\
& \text{(c2)} : c_{i,j,p} \leq b_{i,j-1,p-1} && i, j, p \in [k], p \leq i, j \\
& \text{(d1)} : b_{i,j,j} = c_{i,j,j} && i, j \in [k], j \leq i \\
& \text{(d2)} : b_{i,j,i} = b_{i,j-1,i} + c_{i,j,i} && i, j \in [k], i \leq j \\
& \text{(d3)} : b_{i,j,0} = b_{i,j-1,0} - c_{i,j,1} && i, j \in [k] \\
& \text{(d4)} : b_{i,j,p} = b_{i,j-1,p} - c_{i,j,p+1} + c_{i,j,p} && i, j, p \in [k], p \leq i-1, j-1 \\
& \text{(e1)} : d_j = \sum_{i \in [k]} \sum_{p \leq i, j} \frac{c_{i,j,p}}{p} && j \in [k] \\
& \text{(e2)} : d_j \geq \frac{1}{k-j+1} \cdot \sum_{i \in [k]} \sum_{p \leq i, j-1} (i-p) \frac{b_{i,j-1,p}}{p+1} && j \in [k]
\end{aligned}$$

Figure 3: Linear programming (LP) formulation for computing an upper bound for $h_{\text{seq-PAV}}(k)$.

is in the right-hand side of (e2). Since Sequential PAV selects the candidate which increases the (normalized) PAV score most, we get Constraint (e2).

The LP from Figure 3 gives a set of constraints that each preference profile must satisfy. However, some solutions to the LP do not encode any valid profile; intuitively, the LP finds a solution in a larger space than the space of all valid preference profiles. This is why it only provides an upper bound for $h_{\text{seq-PAV}}(k)$, in contrast to the LP from Figure 1, which computes its exact value. Nevertheless, the value computed by the LP from Figure 3 is accurate enough to formulate interesting conclusions.

E Hardness of Estimating the Proportionality Degree of Seq-PAV

Let us argue that obtaining a good estimation of the proportionality degree for Seq-PAV requires considerably different insights than those used for estimating the proportionality degree of PAV.

Recall that in order to define PAV in Subsection 2.1 we first formulated a scoring function $\text{sc}_{\lambda_{\text{PAV}}} : S(C) \rightarrow \mathbb{R}$ that assigns a numeric value to each committee; then, PAV selects a size- k committee W that maximizes $\text{sc}_{\lambda_{\text{PAV}}}(W)$. The proof of Theorem 1 [2] (a tight estimation of the proportionality degree for PAV) crucially relies on one particular property: that the marginal contributions of all committee members sum up to at most n , i.e., that for each

k	lower-bound	k	lower-bound	k	lower-bound	k	lower-bound
1	1.0 ℓ	4	0.5 ℓ	7	0.2857 ℓ	10	0.2 ℓ
2	1.0 ℓ	5	0.4 ℓ	8	0.25 ℓ	11	0.1818 ℓ
3	0.6666 ℓ	6	0.3333 ℓ	9	0.2222 ℓ	12	0.1666 ℓ

Table 3: The estimation of the k -proportionality degree of the sequential algorithm for maximizing f that is non-negative, monotonic, and that it satisfies the marginal contributions property.

k	lower-bound	k	lower-bound	k	lower-bound	k	lower-bound
1	1.0 ℓ	4	0.8141 ℓ	7	0.7667 ℓ	10	0.7246 ℓ
2	1.0 ℓ	5	0.8372 ℓ	8	0.7492 ℓ	11	0.7150 ℓ
3	0.8888 ℓ	6	0.7888 ℓ	9	0.7358 ℓ	12	0.7066 ℓ

Table 4: The estimation of the k -proportionality degree of the sequential algorithm for maximizing f that is non-negative, monotonic, submodular, and that it satisfies the marginal contributions property.

committee W :

$$(\text{sc}_{\lambda_{\text{PAV}}}(W) - \text{sc}_{\lambda_{\text{PAV}}}(W \setminus \{c\})) \leq n. \quad (1)$$

We ask whether this, and other natural properties of $\text{sc}_{\lambda_{\text{PAV}}}$ are sufficient to prove accurate estimations for the proportionality degree of Sequential PAV. We ask: Assume we are given a function $f: S(C) \rightarrow \mathbb{R}$ that is non-negative, monotonic, and that satisfies the marginal contributions property (**Inequality (1)**); Has the sequential algorithm for maximizing f comparably good proportionality degree to Seq-PAV? If so, it would suggest that one could focus on the aforementioned three properties of Seq-PAV in the attempt to obtaining tight proportionality guarantees.

We first constructed an LP where we represent the function f through a large collection of variables, specifying the values of f for all committees of size equal to at most 12. The three properties (being non-negative, monotonic, and the marginal contributions property) can be easily encoded in the LP. Then, we used a similar technique to the one described in **Subsections 4.2** and **4.3** to assess the proportionality degree of the sequential algorithm for maximizing f . The values computed by our LP are given in **Table 3**. These values are much lower than the ones obtained by the LPs from **Subsection 4.3** and **appendix D** (cf., **Table 1** and **Table 2**). Hence, the aforementioned three properties are not sufficient to derive a good estimation of the proportionality degree for Seq-PAV.

Next, we additionally encoded that f is submodular. The values computed with this additional assumption are provided in **Table 4**. Here, the estimations of the proportionality degree are much better, yet still considerably worse than those derived by the relaxed LP from **Appendix D**. This shows that in order to obtain a tight estimation of the proportionality degree of Seq-PAV one needs some new insights and to explore other properties of $\text{sc}_{\lambda_{\text{PAV}}}(W)$ (the “standard” properties and the one that was used for deriving the proportionality degree of PAV are not sufficient).

F Proportionality Degree of Thiele Methods

In this section we establish bounds on the proportionality degrees for a large subclass of Thiele methods. We first obtain a generic result that covers the whole class of rules, and

then we show how the general result can be used to assess the proportionality degrees for concrete rules in the class; for the selected concrete rules, we will show that our estimations are accurate. The proof of the general result is based on an analysis of how the total Thiele score changes as a result of adding/removing certain committee members. In order to estimate such changes, we analyze a function that aggregates the changes of scores due to adding/removing committee members, employ the pigeonhole principle, and use the Jensen's and the Chebyshev's sum inequalities.

Theorem 3. *Let $\lambda: \mathbb{R} \rightarrow \mathbb{R}_+$ be a non-increasing, convex function, and let $g: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ be a function satisfying $g(\ell, k) \leq k$, and the following equality:*

$$(k - g(\ell, k))\lambda(1 + g(\ell, k)) = \frac{k - \ell}{\ell} \cdot \max_{x \in [k]} x\lambda(x) \quad \text{for each } \ell, k \in [m], \ell \leq k. \quad (2)$$

Then, the λ -Thiele rule has the k -proportionality guarantee of g .

Proof. Let W be a committee winning according to the λ -Thiele rule. Consider an ℓ -large group of voters V , and for the sake of contradiction, assume that $\frac{1}{|V|} \sum_{i \in V} |W \cap A(i)| < \min(|\bigcap_{i \in V} A(i)|, g(\ell, k))$. From the pigeonhole principle we infer that there exists a candidate $c \notin W$ who is approved by each voter from V (and, possibly, by some voters outside of V). Let us now estimate the change of the λ -score due to replacing a committee member $c' \in W$ with c :

$$\Delta(c, c') \geq \sum_{i \in V: c' \notin A(i)} \lambda(|A(i) \cap W| + 1) - \sum_{i \notin V: c' \in A(i)} \lambda(|A(i) \cap W|).$$

Let us now assess the following sum:

$$\begin{aligned} \sum_{c' \in W} \Delta(c, c') &\geq \sum_{i \in V, c' \in W \setminus A(i)} \lambda(|A(i) \cap W| + 1) - \sum_{i \notin V, c' \in W \cap A(i)} \lambda(|A(i) \cap W|) \\ &\geq \sum_{i \in V} (k - |A(i) \cap W|) \cdot \lambda(|A(i) \cap W| + 1) - \sum_{i \notin V} |A(i) \cap W| \cdot \lambda(|A(i) \cap W|) \\ &\geq \sum_{i \in V} (k - |A(i) \cap W|) \cdot \lambda(|A(i) \cap W| + 1) - (n - |V|) \cdot \max_{x \in [k]} x\lambda(x). \end{aligned}$$

From the Jensen's inequality, we get that:

$$\begin{aligned} \sum_{i \in V} \frac{(k - |A(i) \cap W|)}{\sum_{j \in V} (k - |A(j) \cap W|)} \cdot \lambda(|A(i) \cap W| + 1) &\geq \lambda \left(\frac{\sum_{i \in V} (k - |A(i) \cap W|) \cdot (|A(i) \cap W| + 1)}{\sum_{j \in V} (k - |A(j) \cap W|)} \right) \\ &= \lambda \left(1 + \frac{\sum_{i \in V} (k - |A(i) \cap W|) \cdot |A(i) \cap W|}{\sum_{i \in V} (k - |A(i) \cap W|)} \right) \\ &\quad \text{(from Chebyshev's sum inequality, and by monotonicity of } \lambda) \\ &\geq \lambda \left(1 + \frac{\sum_{i \in V} (k - |A(i) \cap W|) \cdot \sum_{i \in V} |A(i) \cap W|}{|V| \sum_{i \in V} (k - |A(i) \cap W|)} \right) \\ &= \lambda \left(1 + \frac{\sum_{i \in V} |A(i) \cap W|}{|V|} \right) \geq \lambda(1 + g(\ell, k)). \end{aligned}$$

Consequently, we get that:

$$\begin{aligned} \sum_{c' \in W} \Delta(c, c') &\geq \sum_{i \in V} (k - |A(i) \cap W|)\lambda(1 + g(\ell, k)) - (n - |V|) \cdot \max_{x \in [k]} x\lambda(x) \\ &> |V|(k - g(\ell, k))\lambda(1 + g(\ell, k)) - (n - |V|) \cdot \max_{x \in [k]} x\lambda(x). \end{aligned}$$

Since for each $c' \in W$ we have that $\Delta(c, c') \leq 0$, it holds that:

$$(n - |V|) \cdot \max_{x \in [k]} x \lambda(x) > |V|(k - g(\ell, k))\lambda(1 + g(\ell, k)),$$

from which it follows that:

$$\frac{k - \ell}{\ell} \cdot \max_{x \in [k]} x \lambda(x) > (k - g(\ell, k))\lambda(1 + g(\ell, k)),$$

a contradiction. This completes the proof. \square

We first argue that Equality (2) has always a unique solution, i.e., that for each positive, non-increasing, convex function λ [Theorem 3](#) provides a single k -proportionality guarantee g . Observe that $(k - g(\ell, k))$ is a decreasing function of $g(\ell, k)$, and $\lambda(1 + g(\ell, k))$ is a non-increasing function of $g(\ell, k)$. As a result, for $g(\ell, k) \leq k$ the left-hand side of (2) is a decreasing function of $g(\ell, k)$. For $g(\ell, k) = k$ the left hand side of (2) is lower or equal to the right-hand side; for $g(\ell, k) = -k^2$ we have the opposite, thus for each ℓ and k there exists a unique value $g(\ell, k)$ that satisfies Equality (2).

[Proposition 3](#), below provides a necessary condition for g to be a k -proportionality degree for λ . Thus, by finding a function $g(\ell, k)$ that violates this condition we obtain an upper-bound on the proportionality degree of the rule. We will get the most accurate upper-bound if we treat this condition as an equality and solve it for g . The equations in [Theorem 3](#) and [Proposition 3](#) differ slightly, so we cannot formally prove that our estimation given in [Theorem 3](#) is always tight. Yet, we solved them for a number of representative Thiele rules (in particular, for $\lambda(i) = 1/i$, $\lambda(i) = 1/\sqrt{i}$, $\lambda(i) = (1/i)^{2/3}$, and $\lambda(i) = 1/i^2$) and verified that the obtained lower and upper bounds are close.

Proposition 3. *Let $\lambda: \mathbb{R} \rightarrow \mathbb{R}$ be a non-increasing, convex function. The k -proportionality guarantee g of the λ -Thiele rule must satisfy the following inequality:*

$$(k - g(\ell, k))\lambda(g(\ell, k)) \geq \frac{k - \ell}{\ell} \cdot \max_{x \in [k]} x \lambda(x + 1) \quad \text{for each } \ell, k \in [m], \ell \leq k.$$

Proof. Let $y = \operatorname{argmax}_{i \in [k]} i \lambda(i)$, and for the sake of contradiction, let us assume that for some $\ell, k \in [m]$, and for some $\epsilon > 0$ it holds that:

$$(k - g(\ell, k))\lambda(g(\ell, k)) + \epsilon < \frac{k - \ell}{\ell} \cdot y \lambda(y + 1).$$

To simplify the notation we set $x = g(\ell, k)$. Clearly, $x \leq k$. We rewrite the above inequality:

$$(k - x)\lambda(x) + \epsilon < \frac{k - \ell}{\ell} \cdot y \lambda(y + 1).$$

We will construct an instance of an election witnessing that g cannot be a proportionality guarantee of the λ -Thiele rule. Let $C = B \cup D$ be the set of candidates, where $B = \{b_1, b_2, \dots, b_k\}$ and $D = \{d_1, d_2, \dots, d_k\}$. We divide the voters into two disjoint groups, V and V' , such that $|V| = n \cdot \frac{\ell}{k}$ and $|V'| = n - |V|$. Without loss of generality, let us assume that n is divisible by k^2 . Each candidate $b \in B$ is approved by all the voters from V . Candidate d_1 is approved by the first $|V| \frac{x}{k}$ voters from V and by the first $|V'| \frac{y}{k}$ voters from V' . The approval set of candidate d_i is constructed by taking the $|V| \frac{x}{k}$ voters from V that appear right after the voters from $V \cap N(d_{i-1})$ and by taking $|V'| \frac{y}{k}$ voters from V' that appear right after the voters from $V' \cap N(d_{i-1})$, taking the cyclic shift if necessary. Finally, from the approval set of the last candidate d_k we remove one arbitrary voter.

We will show that D is an optimal committee for this instance. For the sake of contradiction assume that an optimal committee W contains $z \geq 1$ candidates from B . Then, the voters from V and V' approve on average $\frac{x}{k}(k-z) + z$ and $\frac{y}{k}(k-z)$ committee members, respectively. Since λ is convex, given a constraint on the total number of committee members approved, the score of a committee is maximized when the voters approve roughly the same number of candidates in W . Thus, we can assume that each voter from V approves $\lfloor \frac{x}{k}(k-z) \rfloor + z$ or $\lceil \frac{x}{k}(k-z) \rceil + z$ members of W and, analogously, that each voter from V' approves $\lfloor \frac{y}{k}(k-z) \rfloor$ or $\lceil \frac{y}{k}(k-z) \rceil$ members of W . Consequently, if we replace one candidate from B with a candidate from D in W , then the score of the committee will change by:

$$\begin{aligned} \Delta &\geq |V'| \frac{y}{k} \lambda \left(\left\lceil \frac{y}{k}(k-z) \right\rceil + 1 \right) - \left(|V| - |V'| \frac{x}{k} + 1 \right) \lambda \left(\left\lfloor \frac{x}{k}(k-z) \right\rfloor + z \right) \\ &\geq |V'| \frac{y}{k} \lambda \left(\left\lceil \frac{y}{k} \cdot k \right\rceil + 1 \right) - \left(|V| - |V'| \frac{x}{k} + 1 \right) \lambda \left(\left\lfloor \frac{x}{k} \cdot k \right\rfloor \right) \\ &= |V'| \frac{y}{k} \lambda(y+1) - \left(|V| - |V'| \frac{x}{k} + 1 \right) \lambda(x) \\ &= \left(n - n \frac{\ell}{k} \right) \frac{y}{k} \lambda(y+1) - n \frac{\ell}{k} \left(1 - \frac{x}{k} + \frac{k}{n\ell} \right) \lambda(x). \end{aligned}$$

Since $\Delta \leq 0$, we get that:

$$0 \geq (k-\ell)y\lambda(y+1) - \ell(k-x) \cdot \lambda(x) - \frac{k^2}{n} \lambda(x),$$

which is equivalent to:

$$(k-x) \cdot \lambda(x) + \frac{k^2}{n\ell} \lambda(x) \geq \frac{k-\ell}{\ell} y \lambda(y+1).$$

By taking n large enough, so that $\frac{k^2}{n\ell} \lambda(x) \leq \epsilon$, we get a contradiction. Consequently, we get that D forms a winning committee. Now, observe that V is ℓ -large, and that $|\bigcap_{i \in V} A(i)| \geq x$. The average number of representatives of V is however, lower than x . Thus, g is not a proportionality guarantee for the λ -Thiele method. This gives a contradiction, and completes the proof. \square

Let us give an example application of [Theorem 3](#), by considering $\lambda_{\sqrt{\text{PAV}}}(i) = 1/\sqrt{i}$. In this case, the equality from [Theorem 3](#) becomes:

$$\frac{k-\ell}{\ell} \cdot \sqrt{k} = \frac{k-g(\ell, k)}{\sqrt{1+g(\ell, k)}}. \quad (3)$$

Let us calculate a proportionality guarantee $g_{\sqrt{\text{PAV}}}$ for $\lambda_{\sqrt{\text{PAV}}}$ -Thiele method by solving (3), and picking the solution for $g(\ell, k)$ that is no greater than k . We solved the equality analytically, yet the formulas for the bounds have two lines in displayed equations. We decided it is more informative if we present these guarantees in a form of plots rather than closed formulas. We plot $g_{\sqrt{\text{PAV}}}$ and the guarantees for three other rules in [Figure 4](#) (the plots for upper and lower bounds are in fact indistinguishable for these rules). Interestingly, we can see that the loss of proportionality in comparison to PAV—especially for reasonably large cohesive groups of voters—is moderate. This can justify using rules such as the $\lambda_{\sqrt{\text{PAV}}}$ -Thiele method in some contexts. For instance, as we will see in the next section, such rules guarantee much better utilitarian efficiency.

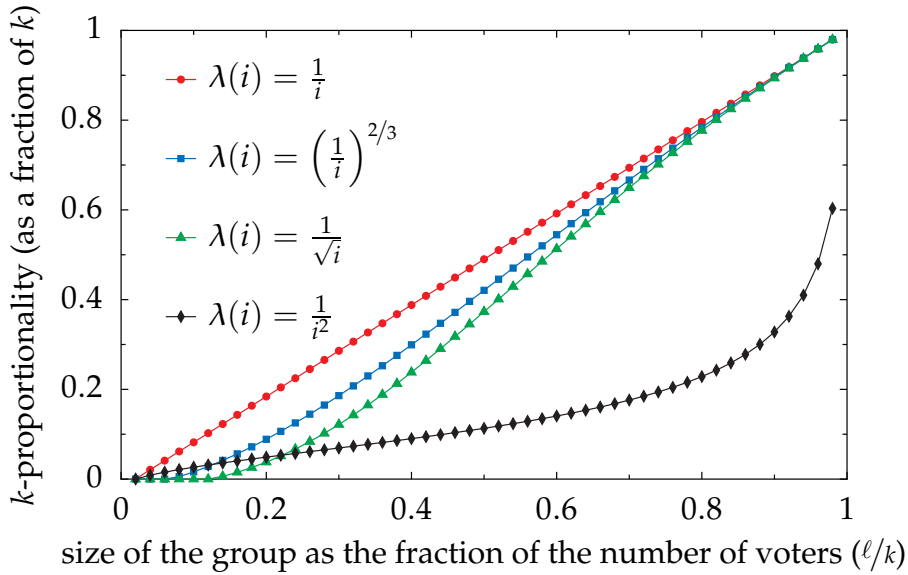


Figure 4: Lower bounds on the k -proportionality degree of selected λ -Thiele rules.

G The Tradeoff Between Proportionality and Utilitarian Efficiency

In the previous sections we assessed the proportionality degree for a number of committee election rules. Here, our goal is estimate the utilitarian efficiency of the considered rules. Together with our previous results, this will allow us to express the studied rules as implementing certain tradeoffs between proportionality and utilitarian efficiency.

We define the utilitarian efficiency of a rule as a lower bound on the ratio between the total number of approvals that the candidates from the winning committee get from all the voters, to the total number of approvals that any committee can possibly get. Formally:

Definition 4 (Utilitarian Efficiency). *For a given committee size k an ABC rule \mathcal{R} has a utilitarian efficiency guarantee of α if for each winning committee $W \in \mathcal{R}(A, k)$ we have:*

$$\sum_{i \in N} |A(i) \cap W| \geq \alpha \cdot \max_{W_{\text{opt}} \in S_k(C)} \sum_{i \in N} |A(i) \cap W_{\text{opt}}|.$$

The next theorem gives a generic tool for calculating the utilitarian efficiency for a large subclass of Thiele methods. In particular, [Theorem 4](#) generalizes known results from the literature—an asymptotically tight utilitarian guarantee is already known for PAV and for the p -geometric rule [\[18\]](#).

Theorem 4. *Let $\lambda: \mathbb{N} \rightarrow \mathbb{R}$ be a non-increasing, convex function. For a committee size k , the λ -Thiele rule has the utilitarian efficiency guarantee of $\frac{\alpha}{1+\alpha} > \alpha - \alpha^2$, for α defined as:*

$$\alpha\lambda(1) = \lambda(1 + k\alpha).$$

Proof. Let $W_{\mathcal{R}}$ and W_{opt} denote, respectively, the committee selected by rule \mathcal{R} , and the committee maximizing the total utility of the voters. For each voter $i \in N$, let $w_i = |A(i) \cap W_{\mathcal{R}}|$. Without loss of generality, we can assume that $W_{\mathcal{R}} \neq W_{\text{opt}}$. Let c be the

candidate that is approved by most voters among the candidates from $W_{\text{opt}} \setminus W_{\mathcal{R}}$, and let $n_c = |N(c)|$. For each candidate $c' \in W_{\mathcal{R}}$ let $\Delta(c, c')$ denote the change of the total λ -score due to replacing c' with c in committee $W_{\mathcal{R}}$. Since $W_{\mathcal{R}}$ is optimal according to \mathcal{R} , we have that $\Delta(c, c') \leq 0$ for each $c' \in W_{\mathcal{R}}$. Thus:

$$\begin{aligned}
0 &\geq \sum_{c' \in W_{\mathcal{R}}} \Delta(c, c') = \sum_{c' \in W_{\mathcal{R}}} \left(\sum_{i \in N(c) \setminus N(c')} \lambda(w_i + 1) - \sum_{i \in N(c') \setminus N(c)} \lambda(w_i) \right) \\
&= \sum_{i \in N(c)} (k - w_i) \lambda(w_i + 1) - \sum_{i \notin N(c)} w_i \lambda(w_i) \\
&\geq \sum_{i \in N(c)} k \lambda(w_i + 1) - \sum_{i \in N} w_i \lambda(\max(w_i, 1)) \\
&\geq \sum_{i \in N(c)} k \lambda(w_i + 1) - \lambda(1) \sum_{i \in N} w_i.
\end{aligned}$$

Thus:

$$\lambda(1) \sum_{i \in N} w_i \geq \sum_{i \in N(c)} k \lambda(w_i + 1).$$

Since λ is convex, by the Jensen's inequality we get that:

$$\lambda(1) \sum_{i \in N} w_i \geq kn_c \lambda \left(1 + \frac{1}{n_c} \sum_{i \in N(c)} w_i \right).$$

Now, let us consider two cases. If $(\sum_{i \in N} w_i) \geq \alpha kn_c$, then the ratio of total utilities of W_{opt} and $W_{\mathcal{R}}$ is at most equal to:

$$\tau = \frac{\sum_{i \in N} w_i + kn_c}{\sum_{i \in N} w_i} \leq 1 + \frac{1}{\alpha} = \frac{1 + \alpha}{\alpha}.$$

Thus, $1/\tau$ equals at least $\alpha/1+\alpha$. Otherwise, i.e., when $(\sum_{i \in N} w_i) < \alpha kn_c$, we have that:

$$\alpha \lambda(1) > \lambda \left(1 + \frac{1}{n_c} \sum_{i \in N(c)} w_i \right) \geq \lambda \left(1 + \frac{1}{n_c} \sum_{i \in N} w_i \right) \geq \lambda(1 + k\alpha),$$

a contradiction. Thus, we get that only the first case is possible, and so the ratio of total utilities of $W_{\mathcal{R}}$ and W_{opt} is at least $\alpha/1+\alpha$. \square

Finally, the subsequent proposition provides an upper bound on utilitarian efficiency guarantees of Thiele rules. In particular, these upper bounds confirm that for many rules the guarantees given in [Theorem 4](#) are asymptotically tight up to a multiplicative factor of 2 (as we will show later on, usually we have $\alpha = o(1)$, which gives the asymptotic tightness).¹¹

Proposition 4. *Let $\lambda: \mathbb{N} \rightarrow \mathbb{R}$ be a decreasing, convex function. Let α satisfy $\alpha \lambda(1) = \lambda(k\alpha)$. For a committee size k , the utilitarian efficiency guarantee of the λ -Thiele rule is below $2\alpha - \alpha^2$.*

¹¹Similarly as for the proportionality—since the equations in [Theorem 4](#) and [Proposition 4](#) differ slightly, we cannot claim that our estimation is asymptotically tight for every Thiele method. Yet, solving the two equations for a number of representative rules shows the lower and the upper bounds are asymptotically the same.

Proof. We construct an instance of election with a set of $2k$ candidates $C = B \cup D$, where $B = \{b_1, \dots, b_k\}$, and $D = \{d_1, \dots, d_k\}$. The voters are divided into two groups, V and $N \setminus V$, such that $|V|k\alpha = n - |V|$. Each voter from V approves all the candidates from B . The voters from $N \setminus V$ are divided into k equal-sized groups; each group approves one candidate from D so that the approval sets of any two groups do not overlap. First, we will show that the elected committee W for this instance has at most αk candidates from B . Indeed, if this were not the case, then by replacing in W one candidate from B with a candidate from D we would change its score by:

$$\Delta > \frac{n - |V|}{k} \cdot \lambda(1) - |V|\lambda(k\alpha) = |V|\alpha\lambda(1) - |V|\lambda(k\alpha).$$

Since $\Delta \leq 0$, we get that $\lambda(k\alpha) > \alpha\lambda(1)$, a contradiction. Thus, W has at most $k\alpha$ candidates from B . Now, we assess the ratio τ of total utilities of W and B :

$$\tau \leq \frac{\alpha k \cdot |V| + (k - \alpha k) \cdot \frac{n - |V|}{k}}{|V|k} = \alpha + \frac{(n - |V|)(1 - \alpha)}{|V|k} = \alpha + \alpha(1 - \alpha) = 2\alpha - \alpha^2.$$

□

The equation in [Proposition 4](#), $\alpha\lambda(1) = \lambda(k\alpha)$, has a unique solution: the left-hand side (LHS) is increasing, right-hand side (RHS) is decreasing, $\text{LHS} < \text{RHS}$ at $\alpha = 1/k$, and $\text{LHS} > \text{RHS}$ at $\alpha = 1$.

E.g., from [Theorem 4](#) and [Proposition 4](#) it follows that the utilitarian efficiency of λ -Thiele methods for $\lambda(i) = \frac{1}{i}$, $\lambda(i) = \frac{1}{\sqrt{i}}$ and $\lambda(i) = \frac{1}{i^2}$ is, respectively, $\Theta\left(\frac{1}{\sqrt{k}}\right)$, $\Theta\left(\frac{1}{\sqrt[3]{k}}\right)$, and $\Theta\left(\frac{1}{k^{2/3}}\right)$.