

Approval-Based Committee Voting in Practice: A Case Study of (Over-)Representation in the Polkadot Blockchain

Niclas Boehmer¹, Markus Brill², Alfonso Cevallos³, Jonas Gehrlein³, Luis Sánchez-Fernández⁴, and Ulrike Schmidt-Kraepelin⁵

¹Harvard University, ²University of Warwick, ³Web3 Foundation, ⁴Universidad Carlos III de Madrid, ⁵TU Eindhoven

Abstract

We provide the first large-scale data collection of real-world approval-based committee elections. These elections have been conducted on the Polkadot blockchain as part of their *Nominated Proof-of-Stake* mechanism and contain around one thousand candidates and tens of thousands of (weighted) voters each. We conduct an in-depth study of application-relevant questions, including a quantitative and qualitative analysis of the outcomes returned by different voting rules. Besides considering proportionality measures that are standard in the multiwinner voting literature, we pay particular attention to less-studied measures of overrepresentation, as these are closely related to the security of the Polkadot network. We also analyze how different design decisions such as the committee size affect the examined measures.

1. Introduction	2
1.1. Our Contributions	2
1.2. Related Work	3
2. Preliminaries	4
2.1. Approval-Based Committee Voting	4
2.2. Application Background	5
3. A First View on Instances and Committees	6
3.1. Description of the Data	6
3.2. Overlap Between Committees	9
4. Preventing Underrepresentation	9
5. Preventing Overrepresentation	12
6. Analyzing Design Decisions	15
6.1. Choosing the Committee Size	15
6.2. Selecting Candidates Multiple Times	15
7. Discussion and Conclusion	16
8. Bibliography	18

A. Additional Material for Section 4	22
B. Additional Material for Section 5	23
B.1. Computing Candidate Groups with Minimum Approval Weight	23
B.2. Proof of Theorem 1	23
C. Additional Material for Section 6	28
D. Kusama Elections	30

1. Introduction

Approval-based committee (ABC) voting describes the task of selecting a subset of candidates based on approval-style preferences of voters over candidates. A central concern in such elections is to ensure that voters’ opinions are “proportionally” reflected in the selected set of candidates, i.e., in the *committee*. Formally capturing proportional representation as axioms, and designing rules that are guaranteed to satisfy these axioms, is a very active area of research; see the book by Lackner and Skowron [2022]. While this effort has resulted in a good theoretical understanding of the axiomatic aspects of different voting rules, there are only very few studies that examine the actual behavior of such rules on specific real-world or synthetically generated voting instances [Elkind et al., 2017, Szufa et al., 2022, Faliszewski et al., 2023b, Mehra et al., 2023]. Nevertheless, these few empirical works have already proved useful, as they found that different voting rules often produce similar outcomes and that most voting rules tend to significantly outperform their worst-case proportionality guarantees. This already motivated the development of new proportionality axioms [Skowron and Górecki, 2022, Brill and Peters, 2023]. Nevertheless, as proportionality axioms seem to lose their discriminative power in practice, the problem arises of how to measure the proportionality of outcomes and how to quantify the differences between proportional rules.

One reason for the shortage of empirical works might be the lack of real-world data. Accordingly, previous empirical works often either resorted to synthetically generated data or converted data from other voting applications such as ordinal elections or participatory budgeting to fit the ABC voting setting. Despite the fact that previous research has named a multitude of potential applications of ABC voting ranging from political elections [Brill et al., 2018] to recommender systems [Strevniotis and Chalkiadakis, 2022b,a, Gawron and Faliszewski, 2022] to forest management [Pommerening et al., 2020], there are only very few applications where ABC elections have been implemented. A so-far mostly unexplored exception are blockchain protocols that conduct ABC elections on a day-to-day basis. Specifically, these elections occur in blockchains using the Nominated Proof-of-Stake (NPoS) protocol. In this system, a subset of stakeholders, called *validators*, are elected to run the consensus protocol, which is crucial for the integrity of the blockchain. The problem of selecting the validators can be modeled as an ABC voting problem, and, indeed, on the *Polkadot* network (<https://polkadot.network>), a proportional ABC voting rule is used [Burges et al., 2020, Cevallos and Stewart, 2021].¹

1.1. Our Contributions

We complement the mostly theoretical literature on ABC voting in various directions, thereby contributing tools and insights for future empirical works.

¹Polkadot currently uses Phragmén’s sequential rule, but considers switching to the Phragmms rule [Cevallos and Stewart, 2021].

- We compile the first collection of real-world ABC elections consisting of 496 elections from the Polkadot blockchain. These elections contain between 18 202 and 48 025 voters and between 920 and 1080 candidates.
- We conduct an in-depth study of the behavior of different voting rules, with a particular focus on their application in Polkadot. We compare the outcomes and analyze whether outcomes under- or overrepresent voters. Notably, security concerns in Polkadot provide a novel view on (and motivation for) the goal of preventing overrepresentation. Our empirical findings, summarized in Section 7, contribute to the ongoing discussion in voting and blockchain research regarding the selection of voting rules. These insights offer compelling justifications for the use of more sophisticated, proportional voting rules, similar to those implemented in Polkadot.
- As part of our analysis, we initiate the study of quantitative measures for over- and underrepresentation that are needed to distinguish and describe the behavior of voting rules on real-world data. We do so by adapting known proportionality axioms and by introducing a new measure regarding the prevention of overrepresentation.
- We consider design decisions concerning the chosen committee size and the question whether (copies of) the same candidate can be selected multiple times. We make recommendations in light of our formulated desiderata.

All collected elections and the code for our experiments are available at github.com/n-boehmer/ABC-practice-Polkadot. Another blockchain network that follows the NPoS protocol is *Kusama* (<https://kusama.network>). In Appendix D, we present 1520 elections conducted in the Kusama network, with roughly 2000 candidates and 10 000 voters each, and verify that most of our empirical findings also hold for the Kusama elections.

1.2. Related Work

There is a large literature that develops (axiomatic) measures to assess the proportionality of a committee [Aziz et al., 2017, Sánchez-Fernández et al., 2017, Lackner and Skowron, 2022, Brill and Peters, 2023, Skowron, 2021, Peters et al., 2021]. Questions typically examined in *empirical* works on ABC voting concern the similarity of outcomes returned by different voting rules [Reichert and Elkind, 2023, Faliszewski et al., 2023b, Elkind et al., 2017], how often voting rules satisfy certain proportionality axioms [Faliszewski et al., 2023b, Mehra et al., 2023], and how often such axioms are satisfied by randomly selected committees [Bredereck et al., 2019, Brill and Peters, 2023]. Complementary to our work, Szufa et al. [2022] also provided tools for more empirical studies of ABC rules by proposing several synthetic models for generating data as well as a framework for visualizing the data and experimental results. On a more general note, empirical works are more common in the context of ordinal voting. In terms of methodological contributions, the *PrefLib* [Mattei and Walsh, 2013, 2017] and *Pabulib* [Faliszewski et al., 2023a] databases, which contain large collections of real-world preference and participatory budgeting data and the recently developed map of elections framework [Szufa et al., 2020, Boehmer et al., 2021] are notable projects.

Regarding the usage of voting in blockchains, there are also some examples beyond Polkadot and Kusama. However, most of them use non-proportional voting rules. For instance, the EOS network uses the voting rule *approval voting* [Grigg, 2017], where the candidates with the highest number of approvals get selected. Notably, the usage of this rule has led to a series of complaints regarding centralization issues [Chong, 2019, Garg, 2019] (arguably related to the non-proportionality of this rule).

2. Preliminaries

In this section, we formally introduce approval-based committee (ABC) elections and describe *Polkadot*, a blockchain network that carries out such elections every day and serves as the main source of data for this paper. Due to the general homogeneity of our data observed in Section 3, in our following analysis we mostly report average values, as the variance is usually quite low.

2.1. Approval-Based Committee Voting

For any $n \in \mathbb{N}$, we define $[n] = \{1, \dots, n\}$. A (*weighted*) ABC election $E = (C, V, A, w, k)$ is defined by a set $C = \{c_1, \dots, c_m\}$ of candidates, a set $V = [n]$ of voters, an *approval profile* A , a *weight function* $w : V \mapsto [0, 1]$ that maps each voter to its *voting weight*, and the size k of the committee to be selected, where a *committee* is a subset of the candidates. The approval profile A consists of a subset $A_v \subseteq C$ for every voter $v \in V$, containing all candidates v approves of. We allow voters to have different *voting weights* (as in Polkadot), which are captured by the weight function w . Without loss of generality, we assume that $\sum_{v \in V} w(v) = 1$. For a subset $V' \subseteq V$ of voters, we let $w(V') := \sum_{v \in V'} w(v)$ denote the sum of their voting weights. For a candidate $c \in C$, we let V_c be the set of *supporters* of c , i.e., $V_c = \{v \in V : c \in A_v\}$. For a candidate $c \in C$, we define $w(c) := w(V_c)$ to be the *approval weight* of c . Extending this notation, for a set $C' \subseteq C$ of candidates their approval weight $w(C') := \sum_{v \in V : A_v \cap C' \neq \emptyset} w(v)$ is the summed voting weight of voters approving at least one candidate from C' . The average *satisfaction* of a voter group $V' \subseteq V$ with a committee W is $\frac{1}{w(V')} \sum_{v \in V'} w(v) \cdot |W \cap A_v|$. Given a committee $W \subseteq C$, for each $\ell \in [k]$, and each non-selected candidate $c \in C \setminus W$, an ℓ -*supporting group* (of c) is a subset of c 's supporters $V' \subseteq V_c$ with $w(V') \geq \frac{\ell}{k}$.

When discussing the legitimacy of a committee $W \subseteq C$, we often use *vote assignments* $\alpha : V \times W \rightarrow [0, 1]$. The idea is that any candidate $c \in W$ needs to be backed by voters supporting them; however, voters should not be counted multiple times across different candidates. This leads to the following constraints of a vote assignment: (i) for any $v \in V$ and $c \in W$, $\alpha(v, c) > 0$ implies that $c \in A_v$, and (ii) for all $v \in V$ it holds that $\sum_{c \in A_v \cap W} \alpha(v, c) \leq w(v)$. Regarding the second constraint, we often additionally assume (implicitly or explicitly) that $\sum_{c \in A_v \cap W} \alpha(v, c) = w(v)$ as long as $A_v \cap W \neq \emptyset$. For a candidate $c \in W$, we are interested in its *backing weight* (w.r.t. α), defined by $\sum_{v \in V_c} \alpha(v, c)$.

An ABC voting rule takes as input an election $E = (C, V, A, w, k)$ and outputs a size- k committee. Below, we define the six rules that we study in this paper; for all six of them, the adaptation to weighted votes is straightforward. While the formulations of the rules allow for ties, we use lexicographic tie breaking whenever necessary, i.e., if two candidates can be added to the committee, we add the one with the smaller index.

Approval Voting (AV) AV selects k candidates with highest approval weight.

Satisfaction Approval Voting (SAV) The rule was introduced by [Brams and Kilgour \[2014\]](#). Under SAV, the score of a candidate $c \in C$ is defined by $\sum_{v \in V_c} \frac{w(v)}{|A_v|}$. The rule then selects k candidates with the highest score.

The next four rules work in a sequential fashion, i.e., they start with the empty committee and add candidates one by one until k candidates have been selected. We often informally refer to them as the *proportional rules*.

Sequential Proportional Approval Voting (seq-PAV) The rule was introduced by [Thiele \[1895\]](#) and analyzed in detail by [Janson \[2016\]](#) and [Aziz et al. \[2017\]](#). The *PAV score* of a committee W is $sc(W) = \sum_{v \in V} w(v) \cdot (\sum_{i=1}^{|A_v \cap W|} \frac{1}{i})$. The rule starts with the empty

committee and in each round adds a candidate with maximum marginal contribution, i.e., $\arg \max_{c \in C} \text{sc}(W \cup \{c\}) - \text{sc}(W)$.

Sequential Phragmén (seq-Phragmén) The rule was introduced by Phragmén [1894] and analyzed in detail by Janson [2016] and Brill et al. [2017]. In this rule, voters earn virtual money at a constant speed proportional to their voting weight. At the beginning, each voter has no money. As soon as the supporters of a candidate jointly own 1 unit of money, the candidate is added to the committee and the money of all its supporters is reset to zero. All other voters keep their money and the process is repeated k times.

Method of Equal Shares (MES) The rule was introduced by Peters and Skowron [2020]. At the beginning, each voter $v \in V$ has a budget b_v of $k \cdot w(v)$ units of virtual money. A candidate $c \in C$ is called q -affordable for some $q \in [0, 1]$ if the supporters of c can together pay one unit of virtual money without any of them paying more than q units, i.e., $\sum_{v \in V_c} \min(b_v, q) \geq 1$. In each round, a not yet selected candidate c which is q -affordable for a minimum value of q is added to the committee and for each voter $v \in V_c$ we set $b_v = \max(0, b_v - q)$. MES might produce outcomes containing less than k candidates, in which case we complete the outcome by running seq-Phragmén. In this case, the budget of a voter at the end of the execution of MES is their starting budget for seq-Phragmén.

Phragmms The rule was introduced by Cevallos and Stewart [2021] as a combination of seq-Phragmén and the maximin support method Sánchez-Fernández et al. [2022].² The overarching goal of the rule is to select a committee W together with a vote assignment α that guarantees high backing weight for any candidate in W . The rule repeats the following two steps iteratively: (i) Given (W, α) , compute a score for any candidate in $c \in C \setminus W$ that corresponds to the highest backing weight t that can be given to c without decreasing the backing weight of any candidate in W below t while only doing local changes to the vote assignment. Then, add the candidate with highest score to W . (ii) Compute a new vote assignment for W that is *balanced*.³

We do not include the maximin support method in our analysis due to its prohibitive computational complexity.

2.2. Application Background

The Polkadot blockchain [Wood, 2016, Burdges et al., 2020] implements a variation of Proof-of-Stake (PoS) as its consensus mechanism to determine the addition of new blocks to the blockchain. These systems rely on a restricted set of *validators* who are granted the exclusive privilege to append new blocks, notably avoiding the reliance on energy-intensive computing power which characterizes Proof-of-Work (PoW) systems, thus making it an environmentally friendly alternative. This is different from Proof-of-Work blockchains such as Bitcoin [Nakamoto, 2008], where everyone can propose new blocks. For the integrity of PoS networks, it is vital that validators adhere to established rules when creating new blocks. Crucially, the network remains secure as long as fewer than one-third of these validators behave maliciously [Lamport et al., 2019].

Polkadot operates as a permissionless network, which means that everyone can become a validator candidate. To address the arising selection problem, the network allows token holders to act as voters, referred to as *nominators*, and screen and evaluate the available candidates,

²We do not include the maximin support method in our analysis due to its prohibitive computational complexity.

³A balanced vote assignment maximizes the sum of the backing weights of the candidates, while minimizing the sum of squared backing weights. For details, see Cevallos and Stewart [2021].

which is a generally quite challenging task [Gehrlein et al., 2023]. In particular, each token holder can submit a ballot approving up to 16 validator candidates. Aggregating these casted ballots, a committee of 297 active validators is selected in each era (day) by Polkadot’s *Nominated Proof-of-Stake (NPoS)* election algorithm, which uses the seq-Phragmén rule.

Contrasting with other ABC elections, elections held on the Polkadot network are characterized by three unique aspects: Firstly, a voter’s voting weight corresponds to their *stake*, that is, the aggregate of tokens in their possession. Secondly, if a voter endorses a candidate who subsequently gets selected, the voter’s stake is held as collateral. If this chosen candidate breaches protocol—for instance, by proposing blocks that are against the rules—the voter’s stake may be seized, an action known as *slashing* (see Section 5). Lastly, a voter continuously receives rewards (in the form of network tokens) for approving candidates who have been elected and diligently perform their duties. These latter two characteristics align the economic incentives of voters with the network’s interests, ensuring that voters include only candidates regarded as trustworthy on their ballot.

In the following sections, we discuss several desiderata that the Polkadot designers formulated for their voting rule, including underrepresentation (Section 4) and overrepresentation (Section 5) concerns. There are also several other desiderata that are beyond the scope of the paper, such as the running time of the algorithm and verification concerns [Cevallos and Stewart, 2021].

3. A First View on Instances and Committees

We start by describing characteristics of our collected dataset and analyzing the overlap between committees that are selected by the voting rules mentioned in Section 2.1.

3.1. Description of the Data

In Polkadot, every day is viewed as an *era*, and one election is conducted at the end of each era, implying that our 496 collected elections have a time-based ordering.⁴ In particular, the studied elections cover eras 398 (July 5, 2021) through 1078 (May 16, 2023) and are generated from openly accessible data maintained and distributed by the *Web3 Foundation*. However, the data stream contains some gaps, i.e., in 185 eras from the above interval, elections were not correctly stored. The committee size in the elections is 300 and voters are only allowed to approve of 16 candidates, i.e., $k = 300$ and $|A_v| \leq 16$ for all $v \in V$.⁵ Notably, it is possible to map voters and candidates of different elections to each other, as they all provide unique IDs.

In Figure 1, we present the sizes of the collected elections. Each election contains between 18 202 and 48 025 voters and between 920 and 1080 candidates. Thus, the number of voters exhibited a stronger fluctuation over time than the number of candidates. Nevertheless, we see here that neither of the two parameters changes too much from one era to the next (the few “jumps” in the plots are mostly due to missing elections). In terms of the average number of candidates a voter approves, there is a monotonic decrease from around 9.7 to around 7.5 over time.

How Much do Elections Change over Time?

We take a closer look at how much elections change over time. For this, we make use of the fact that in Polkadot each voter and each candidate is identified by a stash address, which typically

⁴This also makes the data suitable for testing models that cover collective decision making over time [Lackner, 2020, Boehmer and Niedermeier, 2021]. Currently, such data is very rare even in the context of ordinal single-winner elections [Mattei and Walsh, 2017, Boehmer and Schaar, 2023].

⁵The actual committee size currently used by Polkadot is 297. We use $k = 300$ for easier readability of our results.

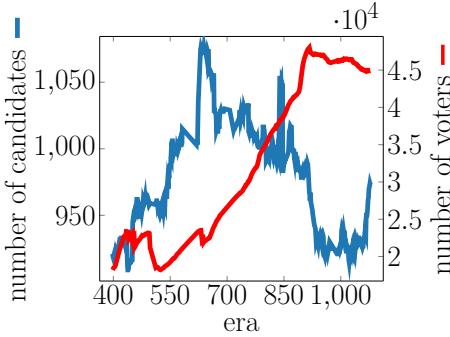


Figure 1: Number of candidates (blue) and voters (red) in our 496 elections.

	AV	SAV	seq-PAV	Phragmms	seq-Phrag.	MES
AV	—	266.190	263.583	266.772	262.595	262.163
SAV	266.190	—	282.042	278.677	276.933	277.206
seq-PAV	263.583	282.042	—	286.562	288.317	288.119
Phragmms	266.772	278.677	286.562	—	288.091	286.887
seq-Phrag.	262.595	276.933	288.317	288.091	—	295.123
MES	262.163	277.206	288.119	286.887	295.123	—

Table 1: Average overlap between committees returned by different voting rules. The committee size is 300.

does not change over time. Accordingly, voters and candidates have the same identity in different elections.

We examine four different quantities comparing two elections $E = (C, V, A, w, k)$ and $E' = (C', V', A', w', k)$. For each of them, the smaller the value, the smaller the change.

Relative voter set change This quantity captures how much the set of voters changed between the two elections: $\frac{|V \Delta V'|}{|V| + |V'|}$.

Relative weight change This quantity captures how much the weight of voters that are present in both elections changes: $\frac{\sum_{v \in V \cap V'} |w(v) - w'(v)|}{\sum_{v \in V \cap V'} w(v)}$

Relative opinion change This quantity captures how much the voters that are present in both elections change their opinions about the candidates that are also present in both elections. We take the average of the following expression over all voters $v \in V' \cap V$ (for two subsets of candidates: $\frac{|(A_v \cap C') \Delta (A'_v \cap C)|}{|(A_v \cap C')| + |(A'_v \cap C)|}$).

Relative candidate set change This quantity captures how much the set of candidates changes: $\frac{|C \Delta C'|}{|C| + |C'|}$.

For each of these four quantities, Figure 2 depicts the change between two consecutive elections in our dataset. The picture for all of them is similar: Going from one election to the next, typically no large change occurs, which is also quite intuitive as without any active actions from participants, nothing will change. Note that most of the spikes visible in Figure 2 are due to gaps in our data. However, comparing the first election in our dataset to the last one, indeed quite some changes occurred: the relative voter set change is 0.76, the relative candidate set change is 0.39, the relative weight change is 0.4, and the relative opinion change is 0.46.

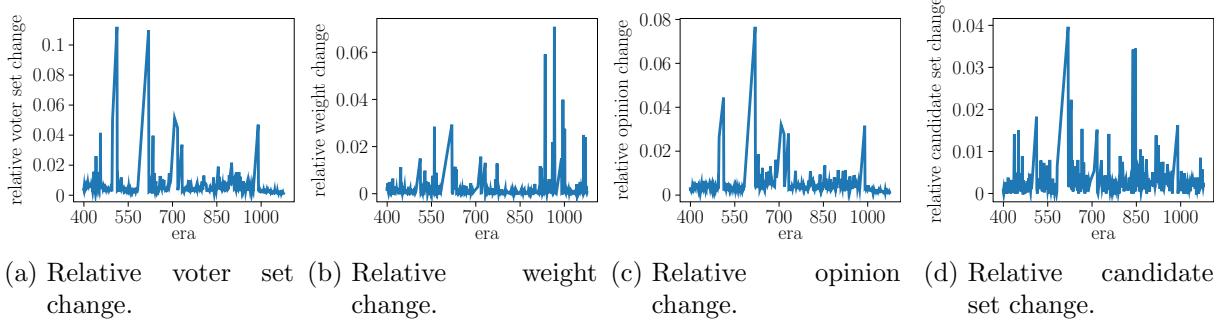


Figure 2: Some statistics capturing by how much elections change from one era to the next.

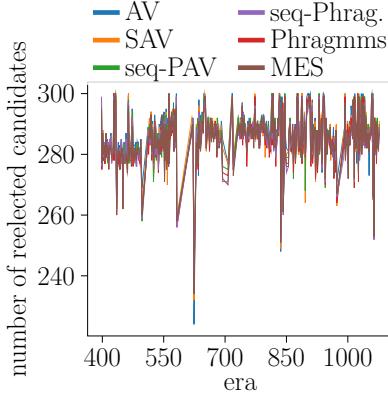


Figure 3: Number of candidates that are also part of the winning committee in the next election.

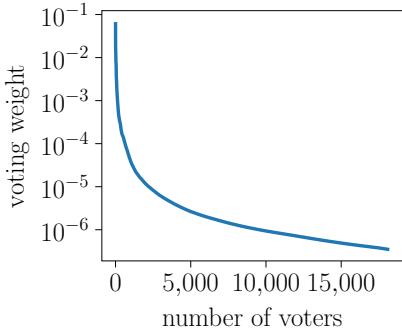
How Much do Committees Change Over Time?

In Figure 3, we show the overlap between winning committees in two successive elections (recall that we can map candidates from different elections to each other by their stash address). In general, we observe that, in most cases, only around 20 candidates get replaced when moving from one election to the next (the average value for all rules is between 13 and 15). Recalling that we have seen in Figure 2 that elections change only marginally in many steps, this might still be viewed as surprisingly many changes (especially, recalling that the difference between the outcomes produced by the proportional rules on the same election is lower). Beyond that, it is worth noting that the graphs show some eras with significantly smaller overlaps. Again, most of these are due to the gaps in our data, which naturally lead to smaller overlap, as a larger period of time has passed between one election and its successors.

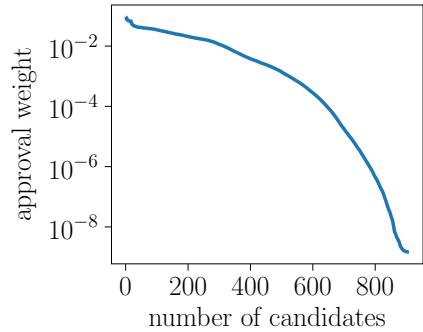
Another trend that stands out from Figure 3 is that the rules behave remarkably similarly, which means that all of them make roughly the same number of adjustments to account for changes in the election. This observation stands in partial contrast to previous theoretical and empirical results from the literature that different rules have a different sensitivity to random or adversarial changes in the election [Bredereck et al., 2021, Boehmer et al., 2023a, 2022]. Yet, voters in Polkadot elections certainly do not change their votes randomly.

Election Properties

In the following, we describe some structural properties of the Polkadot elections. The plot in Figure 4a shows the averaged order statistics of the weight of voters: For each election, we sort the voter weights decreasingly and depict the averaged values here (we cut off the plot at the minimum number of voters in our elections). It stands out that the weight of voters is distributed



(a) Distribution of voting weight.



(b) Distribution: Candidate's approval weight.

Figure 4: Some statistics regarding the 496 collected Polkadot elections. The y -axes are logarithmic.

very unevenly. On average, the top-31 voters combined have more weight than the remaining voters together (these powerful voters are often called “whales”). In Figure 4b, we repeat this analysis with the approval weight of candidates. Here, again, the distribution is far from uniform, yet slightly less imbalanced than for the voting weights: Some candidates have a substantially higher approval weight than others; however, typically around 300 candidates have an approval weight of at least 0.01. This mismatch between Figure 4a and Figure 4b can be explained by the fact that voters can approve multiple candidates and that voters with high weights often have disjoint approval sets (voters with a high weight typically represent a company or organization which also adds some candidates in the election; accordingly each company lets their voters vote for their own candidates in order to ensure that they get selected and increase their monetary rewards).

3.2. Overlap Between Committees

In Table 1, we show the similarity between different voting rules in terms of the average overlap of their computed committees. We observe a high consensus of all rules, thereby confirming previous findings on synthetic data [Reichert and Elkind, 2023, Faliszewski et al., 2023b]; in fact, across all elections, two rules never disagree on more than 51 out of the $k = 300$ candidates. However, there are some differences: The four considered proportional rules have a particularly high agreement. This is most extreme for seq-Phragmén and MES, which have an average overlap of 295 and never produce outcomes that differ in more than 11 candidates. In contrast, AV returns committees that are furthest away from the outcomes of the other rules, while outcomes produced by SAV are in general slightly closer to the ones returned by the proportional rules. Moreover, in Appendix 3.1, we analyze how much winning committees change over time and observe that, with respect to this aspect, all our rules behave remarkably similarly to each other. In most cases, winning committees in successive time steps differ by at most 20 candidates.

4. Preventing Underrepresentation

In Polkadot, voters are financially rewarded when (some of) their approved candidates appear in the selected committee. Thus, it is important to ensure that no voter groups are underrepresented in the selected committee, as otherwise voters might feel disengaged and stop participating in the protocol. In general, unrepresentative outcomes can also lead to power centralization, which is very dangerous for the system. In voting theory, underrepresentation is typically prevented by requiring proportional representation. Intuitively, this means that a group of voters with a

	PAV score	JR violations	EJR+ violations	priceability violations	priceability gap
AV	2.502	28.347	32.343	124.341	6.21
SAV	2.547	25.143	32.321	116.692	7.14
seq-PAV	2.585	0	0	12.127	0.53
Phragmms	2.578	0 (Guarantee)	0	0.000	-0.12
seq-Phrag.	2.578	0 (Guarantee)	0	0.000	-0.16
MES	2.579	0 (Guarantee)	0 (Guarantee)	0.000	-0.17

Table 2: Measures related to underrepresentation. Entries marked with a “(Guarantee)” are guaranteed to be zero.

summed voting weight of $\frac{\ell}{k}$ should be allowed to select ℓ of the committee members. In this section, we consider different forms of measuring the representation (and satisfaction) of voters.

PAV Score

We start with the PAV score of the computed committees (see Table 2), which is usually regarded as a simple proportionality measure. Unsurprisingly, seq-PAV, which greedily optimizes this value in a sequential fashion, performs best. The other three proportional rules all have a very similar performance and perform only marginally worse than seq-PAV (i.e., by about 0.2%), while still outperforming AV and SAV.

JR and EJR+

In addition, proportionality is typically judged in terms of binary proportionality axioms, which are often defined in terms of cohesive groups, i.e., subgroups of voters of a certain size that jointly approve some number of candidates. However, because the respective notions tend to be computationally intractable to check, we follow the approach of [Brill and Peters \[2023\]](#) and focus on notions regarding non-selected candidates instead. The intuition here is as follows: if a non-selected candidate is approved by “many” voters who are currently “underrepresented” in the committee, then the candidate should have been added to the committee. Two axioms that implement this general intuition are JR [[Aziz et al., 2017](#)] and EJR+ [[Brill and Peters, 2023](#)].⁶

Definition 1. For a given election $E = (C, V, A, w, k)$ and committee $W \subseteq C$, a non-selected candidate $c \in C \setminus W$ violates EJR+ if there is an ℓ -supporting group V' of c such that all voters from V' approve less than ℓ candidates from W . If c violates EJR+ for $\ell = 1$, c violates JR.

MES is guaranteed to output committees satisfying EJR+, while committees returned by Phragmms and seq-Phragmén always satisfy JR but may fail EJR+, and the other rules fail even JR [[Brill and Peters, 2023](#), [Cevallos and Stewart, 2021](#), [Lackner and Skowron, 2022](#)]. However, in our instances the behavior of the rules is quite different: All four proportional rules return committees satisfying EJR+ for all tested instances. In contrast, AV violates JR and EJR+ on all instances, while SAV violates JR in all but 36 instances and EJR+ in all but 26 instances. To get a more differentiated view, in Table 2 we present the average number of non-selected candidates violating JR/EJR+ in our elections. Given the reported numbers, one can conclude that the committees returned by AV and SAV are typically quite far away from satisfying the two axioms. Moreover, as for both rules the number of candidates violating JR and EJR+ are quite similar to each other, it follows that if a candidate violates EJR+, then they often violate JR as well.

⁶EJR+ implies JR as well as other established proportionality notions such as EJR [[Aziz et al., 2017](#)], PJR [[Sánchez-Fernández et al., 2017](#)], and IPSC [[Aziz and Lee, 2021](#)].

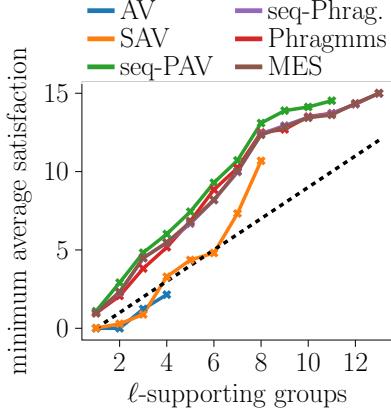


Figure 5: Minimum average satisfaction of ℓ -supporting groups. The dashed line is the function $f(\ell) = \ell - 1$. Lines stop in case no ℓ -supporting group of this size exists.

Average Satisfaction ℓ -Supporting Groups

EJR+ and JR only check for ℓ -supporting groups in which *all* voters are unsatisfied with the current assignment. Weakening this approach, one can also search for ℓ -supporting groups with low *average* satisfaction. This view is reflected in the notion of representativeness recently introduced by Brill and Peters [2023], which is related to the proportionality degree [Skowron, 2021]. In fact, Brill and Peters [2023] proved that for each rule satisfying EJR+ (such as MES), each ℓ -supporting group is guaranteed to have an average satisfaction of at least $\frac{\ell-1}{2}$ with the committee. In each election, we compute for each $\ell \in [k]$, the ℓ -supporting group that has the lowest average weighed satisfaction. In Figure 5, we show this value for varying $\ell \in [k]$ (averaged over all elections where an ℓ -supporting group exists). On average, all four proportional rules produce outcomes in which all ℓ -supporting groups have an average satisfaction clearly above $\ell - 1$. Thus, they outperform their known worst-case guarantees from the literature, and even consistently outperform the best possible guarantee of $\ell - 1$ [Brill and Peters, 2023]. Interestingly, seq-PAV performs slightly better than the other proportional rules. SAV and AV perform substantially worse, yet still acceptable for $\ell \geq 4$.

Priceability

Complementing the above analysis, we propose a quantitative measure based on the notion of *priceability* [Peters and Skowron, 2020], which checks whether voters have an equal influence on the outcome:

Definition 2 (Peters and Skowron, 2020). *A committee $W \subseteq C$ is priceable in a given election $E = (C, V, A, w, k)$ if there exists a price system (p, f) with price $p \in [0, 1]$ and a cost assignment $f : V \times W \mapsto [0, 1]$ such that:*

1. $\sum_{c \in W} f(v, c) \leq w(v)$ for all $v \in V$,
2. $f(v, c) = 0$ for all $v \in V$ and $c \in W \setminus A_v$,
3. $\sum_{v \in V} f(v, c) = p$ for all $c \in W$, and
4. $\sum_{v \in V_c} (w(v) - \sum_{c \in W} f(v, c)) \leq p$ for all $c \in C \setminus W$.

Seq-Phragmén, Phragmms, and MES always return priceable committees [Peters and Skowron, 2020, Cevallos and Stewart, 2021]. A natural way to turn this axiom into a quantitative measure

is to compute the price system that minimizes the *priceability gap*, i.e., the gap between the determined price and the maximum spare money of the supporters of a non-selected candidate, i.e.,

$$\min_{c \in C \setminus W} \sum_{v \in V_c} (w(v) - \sum_{c \in W} f(v, c)) - p.$$

This captures by how much the candidate is exceeding or missing to be affordable by their supporters. Using an LP, we computed the price system minimizing this value. In the second-to-last and last columns of Table 4, we show the average number of candidates whose supporters' money exceeds the set price and the average priceability gap (normalized by the set price), respectively. Moreover, in Figure 8 we show the average normalized money by which candidate's supporters exceed the set price (where we first sort candidates decreasingly by this value and then take the average).

SAV and AV return committees that are far away from being priceable: In all elections, they returned committees violating priceability with an average priceability gap of 6.21 and 7.14 and typically around 100 candidates whose supporters own more than the set price: In fact, in more than 75% of the elections, there are non-selected candidates whose supporter's money exceeds the price by more than 600%. This is a clear indicator that certain voters are much better represented and had a much larger influence on the returned outcome than others in the selected committee.

For seq-PAV, the picture is mixed: The rule returns priceable outcomes in 327 of the elections, yet in case the returned outcomes are not priceable, the priceability gap is sometimes as high as 4, which leads to an average priceability gap of 0.53 and on average 12 candidates whose supporters exceed the set price.

For the other rules, it is guaranteed that the returned committees are priceable. Accordingly, their priceability gap is negative. In fact, there is a small difference between these rules: For MES and seq-Phragmén non-selected candidates are slightly further away from being affordable by their supporters than for Phragmms.

5. Preventing Overrepresentation

One of the major concerns of blockchain designers is the security of the chain: If a certain fraction of participants collude and together execute some malicious action, they can seize control over the chain and threaten the integrity of the whole system. To protect against such attacks in Nominated Proof-of-Stake, it is vital to ensure that groups of candidates can only get selected if their joint set of supporters has a sufficient stake. In other words, we want to prevent overrepresentation [Cevallos and Stewart, 2021].⁷ In the following, we explore three perspectives on overrepresentation.

Minimum Approval Weight

The first measure we consider was introduced by Cevallos and Stewart [2021]. In order to make it as difficult as possible for an attacker to get ℓ committee members selected, they propose to maximize $\min_{W' \subseteq W : |W'| = \ell} w(W')$, i.e., the minimum approval weight of a group of ℓ committee

⁷The critical threshold of elected malicious candidates depends on the type of consensus and is $\frac{k}{3}$ in Byzantine fault-tolerant consensus [Pease et al., 1980], as used in Polkadot, and $\frac{k}{2}$ in Nakamoto consensus [Stifter et al., 2018]. However, already a small number of malicious agents might pose certain inconveniences for the system [Cevallos and Stewart, 2021].

	maximin support value	min. appr. weight of winner	cost of “replacing” ℓ		
			$\ell = 1$	$\ell = \frac{k}{3}$	$\ell = \frac{k}{2}$
AV	0.0015	0.0110	0.0110	0.14	0.27
SAV	0.0018	0.0024	0.0015	0.24	0.40
seq-PAV	0.0024	0.0028	?	?	?
Phragmms	0.00272	0.0028	0.0027	0.40	0.79
seq-Phrag.	0.00270	0.0028	0.0027	0.40	0.79
MES	0.00269	0.0028	?	?	?

Table 3: Measures related to overrepresentation.

members.⁸ Figure 6a depicts these values for a varying value of ℓ .⁹ For all examined rules, the values are generally quite high and close to the dashed line (which corresponds to the function $\frac{1}{300}x$). This is reassuring, as it means that groups of candidates in the selected committee are also backed by an appropriate amount of stake. Considering the differences between the rules, we see that the four proportional rules perform best (for $\ell \geq 15$), with seq-PAV performing slightly worse than the other three. AV performs worst; yet, the generally small difference between AV and the proportional rules is quite remarkable given that we have seen in Section 4 that AV tends to underrepresent voter groups (and thus runs the risk of overrepresenting others).

Maxmin Support Value

As an aggregate version of this measure, Cevallos and Stewart [2021] also proposed to consider the minimum *average* approval weight of a committee member, where the minimum is taken over groups of different sizes, i.e., $\min_{W' \subseteq W} \frac{1}{|W'|} w(W')$. Interestingly, Cevallos and Stewart [2021] proved that this value is equivalent to the *maximin support (MMS) value*, which was introduced in a different context (see Definition 3 below). In Table 3, we see that, on average, the four proportional rules achieve substantially higher MMS values than AV and SAV. In particular, for the four proportional rules, the MMS value is very close to the minimum approval weight of a selected candidate (also in Table 3), which constitutes a natural upper bound for it. Taking a closer look at the four proportional rules, seq-PAV performs slightly worse, while Phragmms, seq-Phragmén, and MES all produce very similar values. Particularly interesting is the comparison between seq-Phragmén and Phragmms: The main argument in favor of a potential switch from the former to the latter is the fact that the latter provides a constant-factor approximation guarantee for the MMS value [Cevallos and Stewart, 2021]. Among the 252 instances in our dataset where seq-Phragmén and Phragmms select different committees, Phragmms outperforms seq-Phragmén in 209 cases. (In the remaining 43 instances, seq-Phragmén wins.) However, the difference between the two rules is always at most 0.0002 and thus negligible.

Stake Lost

For our next measure, we take a closer look at how slashing works in Polkadot. As explained in Section 2.2, when committee members misbehave, some of their supporters lose stake. To determine the amount of stake a voter loses in this event, the election mechanism in Polkadot

⁸For $\ell = 1$, this value is the minimum approval weight of a selected candidate and is maximized by AV (see Table 3).

⁹Computing these values is NP-hard, which is why we resorted to an ILP (see Appendix B.1 for details). As solving a single instance of the ILP took sometimes more than one day, in Figure 6a we only averaged over 15 instances, uniformly spaced over our dataset, and considered values of ℓ from $\{1, 15, 30, 45, \dots, 300\}$.

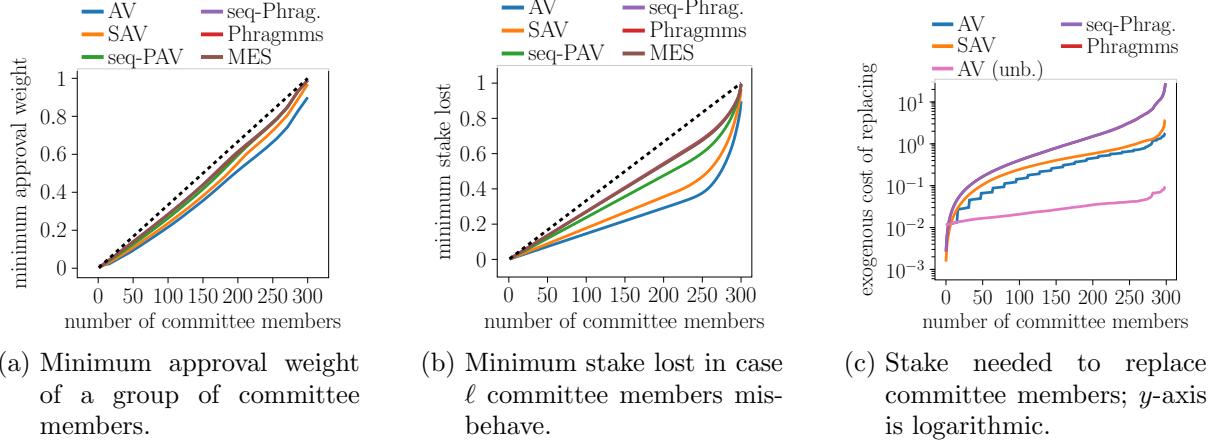


Figure 6: Different figures for metrics to prevent overrepresentation. The dashed line is the function $\frac{1}{300}\ell$.

not only outputs a committee W , but also a vote assignment α , where $\alpha(v, c)$ specifies how much of a voter v 's stake is assigned to committee member $c \in A_v$ (see Section 2.1). Following an approach proposed by Sánchez-Fernández et al. [2022], this vote assignment is chosen so as to maximize the backing weight of the least-backed committee member.

Definition 3. Given an election $E = (C, V, A, w, k)$ and a committee W , the maximin support value of W is given by $\max_{\alpha} \min_{c \in W} \sum_{v \in V_c} \alpha(v, c)$, where the maximum is taken over all possible vote assignments for W . A vote assignment maximizing this quantity is called a maximin assignment.

If a committee member c misbehaves in Polkadot, each supporter $v \in V_c$ loses (up to) $\alpha(v, c)$ stake [Burges et al., 2020]. Accordingly, the maximin support value expresses the minimum amount of stake that is slashed if a single member of the committee acts maliciously. For a given committee, a maximin assignment can be computed via an LP [Sánchez-Fernández et al., 2022]. Using this assignment, in Figure 6b we plot the minimum amount of stake assigned to a group of ℓ committee members, for $\ell \in [k]$. This corresponds to the minimum stake that will be slashed if ℓ committee members act maliciously.

Figure 6b is closely connected to Figure 6a; the main difference is that in Figure 6b we assume that there is a predefined split of the stake of supporters onto their approved candidates (as defined by the maximin assignment), whereas in Figure 6a we consider the full stake of all voters approving at least one candidate from the group. Accordingly, in Figure 6b, the dashed line is an upper bound (that is achieved if all committee members have the same backing weight). The difference between the rules is much more pronounced in Figure 6b, with the proportional rules performing much better than AV and SAV. In particular, for the critical thresholds of $k/3$ and $k/2$, the difference is above 50%. The proportional rules, in turn, are reasonably close to the ideal line.

Exogenous Cost of Replacing

For our final measure, we take an “exogenous” view and reason about how much stake a malicious agent would need to possess in order to *replace* a given number of committee members by newly added candidates (assuming that the agent is allowed to add new votes and candidates, while the remainder of the election remains unchanged). To the best of our knowledge, this view has not been explored so far. We show that, for most of our rules, the cost of replacing ℓ candidates

can be computed in constant time, provided we have access to data that is generated while computing the rules.

Theorem 1 (informal). *For AV, SAV, seq-Phragmén, and Phragmms, the minimum exogenous cost of replacing ℓ candidates can be computed in $\mathcal{O}(1)$ time, assuming we use data from executions of the rules (specified in Appendix B.2).*

For details, we refer to Appendix B.2, where we also discuss the complexity of this problem for seq-PAV and MES. Figure 6c shows the stake needed to replace ℓ committee members, for $\ell \in [k]$ (the y -axis is logarithmic and the pink line stands for AV where approval votes of unbounded length are allowed). AV has the highest cost for $\ell = 1$, since AV maximizes the minimum approval weight of a selected candidate (see Table 3). For larger ℓ , replacements under AV become cheaper, as the same stake can be used to replace multiple candidates. In contrast, for the other rules (except seq-PAV and MES), there always exists an optimal replacement in which every voter approves one candidate. However, already for $\ell = 4$ the cost of replacing starts to be more expensive for seq-Phragmén and Phragmms than for AV. Notably, external attacks for AV would be even cheaper if voters were allowed to approve an unbounded number of candidates (see the pink line in Figure 6c). For larger ℓ , the cost for seq-Phragmén and Phragmms becomes substantially higher than for AV and even SAV (see also Table 3). In particular, for both seq-Phragmén and Phragmms, replacing one-third of the committee would require 40% of the total stake of all voters. This is roughly the total stake possessed by agents not participating in the validator elections, and thus makes such an external attack highly unlikely.

6. Analyzing Design Decisions

In this section, we briefly summarize our findings regarding the influence of various design decisions on our measures. For the full analysis and relevant data, see Appendix C.

6.1. Choosing the Committee Size

In Polkadot, the committee size is determined by the network’s governance body of token holders and is thus an adjustable design choice. This observation raises the following general question: Based on the desiderata formulated in the previous sections, would it be beneficial to increase or decrease the committee size?

We focus on $k = 200$ and $k = 400$ as two alternative committee sizes (Appendix C also includes $k = 250$ and $k = 350$). Generally speaking, for $k = 200$ and $k = 400$, the differences between the rules are similar as for $k = 300$, so we only focus on the general trends in the results. Regarding underrepresentation, increasing k turns out to be clearly favorable: For $k = 400$, the minimum average satisfaction of ℓ -supporting groups increases substantially. Regarding overrepresentation, it turns out that the committee size has only a marginal influence on attacks trying to take over one-third or half of the committee: Intuitively speaking, individual candidates in a committee with more candidates are less backed; on the other hand, one-third/half of the committee corresponds to larger numbers of candidates. It turns out that both effects approximately cancel out for the cost of replacing. By contrast, the minimum approval weight values slightly drop by around 15% when increasing k to 400.

6.2. Selecting Candidates Multiple Times

Brill et al. [2020] initiated the study of ABC elections where multiple copies of a candidate can be included in the committee. Here, we analyze what impact this would have on our measures.¹⁰

¹⁰While this is currently not possible in the Polkadot network, it is very easy for candidates to create copies of

Generally speaking, allowing copies in the committee is favorable. Most notably, it leads to a more uniform distribution of candidates' backing weights in the maximin assignment, leading to a drop of their variance by around 80%. The only metrics where allowing copies leads to a worse performance regard the satisfaction of ℓ -supporting groups. This is to be expected, as allowing for copies implies that *all* candidates count as “non-selected” and, thus, our measures range over strictly more subgroups of voters. Nevertheless, the minimum average satisfaction of ℓ -supporting groups remains quite high, suggesting that allowing for copies might be a worthwhile consideration for platform designers. Remarkably, whether or not to allow copies has a stronger (positive) impact on several measures than changing the committee size to 200 or 400 (or choosing among the proportional rules). In particular, this holds for the PAV score, the cost of replacing candidates, and the minimum approval weight of groups of committee members.

7. Discussion and Conclusion

We conducted a thorough multi-criteria analysis of the behavior of different ABC voting rules in Polkadot elections. We conclude with a discussion of our data and results followed by a summary of different directions for future work.

Discussion and Summary of Results

Generally speaking, we found that the “proportional” rules (seq-PAV, seq-Phragmén, MES, and Phragmms) behave very similarly. Only seq-PAV showed a slightly different behavior: Our results suggest that the rule prioritizes the total weighted satisfaction of voters slightly more, even if this means that some voters have a greater influence on the outcome than others. One reassuring observation for Polkadot developers is that all four rules provide a good level of security. For instance, for both seq-Phragmén and Phragmms, seizing control of the network by replacing one-third of the committee with malicious candidates requires around 40% of the tokens present in the election.

In contrast, SAV and the common AV rule tend to return committees that are less similar to the ones produced by the proportional rules. In particular, they frequently violate basic representation axioms and return outcomes overrepresenting certain voter groups. The most drastic difference is that seizing control over one-third of the committee only requires 14% of tokens for AV and 24% for SAV. Nevertheless, the performance of both rules could still be viewed as acceptable with regard to the other security measures.

We have also explored the impact of various design decisions. Here, we observed that changing the committee size only marginally influences our measures, whereas allowing for selecting (copies of) candidates multiple times is a more impactful and generally beneficial design decision.

Collected Data

We have collected 496 and 1520 large approval-based committee elections from the Polkadot and Kusama network, respectively and hope that others will use our data which is available on github.com/n-boehmer/ABC-practice-Polkadot. One might argue that the fact that all voting rules substantially outperform their axiomatic guarantees on our collected elections (e.g., seq-PAV always fulfills EJR+) suggests that voters’ preferences must be quite “simple”. While measuring the “simplicity” of an election is non-trivial, we calculated the following quantity as a proxy: For any two voters with overlapping approval set, we divide the size of the intersection of their approval sets by the sum of the sizes of their approval sets. If we would be in the so-called

themselves. Indeed, entities that own a large amount of stake create multiple (up to 100) candidates in the system and even label them as copies. Thus, *de facto* the same entity can already control multiple candidates.

“party-list” setting, i.e., any pair of approval sets are disjoint or equal, then this value would be 0.5. However, we found that the average of this value is typically around 0.1, proving that our elections are far from falling into the simple party-list setting. Instead, our interpretation of the observed phenomenon is rather that most of our considered voting rules, e.g., seq-PAV, satisfy demanding proportionality axioms on most (reasonable) elections, only violating them on adversarial constructed ones unlikely to occur in real-world or synthetic data. Indeed, this view is supported by the works of [Faliszewski et al. \[2023b\]](#), [Bredereck et al. \[2019\]](#), and [Brill and Peters \[2023\]](#), who also report that voting rules substantially outperform theoretical guarantees and that even random committees often satisfy strong proportionality notions.

Future Work

Several lines of continuation of this work can be envisioned. First, it would be interesting to dive deeper into the analysis of different voting rules, trying to identify properties of “controversial” candidates (i.e., candidates that are only selected by some rules but not others).

Second, it would be intriguing to analyze the connection between over- and underrepresentation in more detail, both from a theoretical and an empirical perspective; for instance, it would be interesting to relate under- and overrepresentation axioms to each other formally, or to give bounds on the “gap” between a rule’s over- and underrepresentation performance. Moreover, the goal of preventing overrepresentation (motivated by Polkadot’s security concerns) by itself constitutes a new research direction deserving further attention and intense study.

Third, our finding that classic “binary” axioms are not able to meaningfully distinguish voting rules in practice highlights the necessity for further theoretical work on axioms. Our hope is that our work initiates a rethinking of the standard binary axiomatic analysis and a shift more into a direction of *quantitative* axioms (e.g. the *cost of replacing* candidates). Our paper takes a first step in this direction, but given the multitude of binary axiomatic works, there are clearly numerous directions remaining.

Acknowledgments

This work was initiated at Dagstuhl Seminar 22271 “Algorithms for Participatory Democracy” in July 2022 (<https://www.dagstuhl.de/22271>). We thank Jannik Peters for helpful comments. Markus Brill was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant BR 4744/2-1. Luis Sánchez-Fernández was supported by the “Generation of Reliable Synthetic Health Data for Federated Learning in Secure Data Spaces” Research Project (PID2022-141045OB-C43 (AEI/ERDF, EU)) funded by MCIN/AEI/ 10.13039/501100011033 and by “ERDF A way of making Europe” by the “European Union”. Ulrike Schmidt-Kraepelin was supported by the *Centro de Modelamiento Matemático (CMM)* (under grant FB210005, BASAL funds for center of excellence from ANID-Chile), *ANID-Chile* (grant ACT210005) as well as by the National Science Foundation under Grant No. DMS-1928930 and the Alfred P. Sloan Foundation under grant G-2021-16778 while she was in residence at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Fall 2023 semester. The main work was done while Niclas Boehmer was affiliated with TU Berlin, where he was supported by the DFG project ComSoc-MPMS (NI 369/22).

Ethics Statement

This paper acknowledges the significant environmental impact associated with blockchain technology. However, the Polkadot network employs a Proof-of-Stake (PoS) consensus mechanism that does not involve energy-intensive mining activities that are common to the more well-known

Proof-of-Work (PoW) systems. This results in an energy usage that is several orders of magnitude smaller. Polkadot has been acknowledged as one of the protocols with the lowest carbon footprint having an annual CO₂ emission equivalent to that of five average American households [Crypto Carbon Ratings Institute, 2023]. This is partly because the number of validators who participate in consensus is limited enabled by the election process studied in the paper.

8. Bibliography

- H. Aziz and B. E. Lee. Proportionally representative participatory budgeting with ordinal preferences. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5110–5118. AAAI Press, 2021. [→ p. 10]
- H. Aziz, M. Brill, V. Conitzer, E. Elkind, R. Freeman, and T. Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017. [→ p. 3, 4, 10]
- N. Boehmer and R. Niedermeier. Broadening the research agenda for computational social choice: Multiple preference profiles and multiple solutions. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1–5. ACM, 2021. [→ p. 6]
- N. Boehmer and N. Schaar. Collecting, classifying, analyzing, and using real-world ranking data. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1706–1715. ACM, 2023. [→ p. 6]
- N. Boehmer, R. Bredereck, P. Faliszewski, R. Niedermeier, and S. Szufa. Putting a compass on the map of elections. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 59–65, 2021. [→ p. 3]
- N. Boehmer, R. Bredereck, P. Faliszewski, and R. Niedermeier. A quantitative and qualitative analysis of the robustness of (real-world) election winners. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO)*, pages 7:1–7:10. ACM, 2022. [→ p. 8]
- N. Boehmer, P. Faliszewski, L. Janeczko, and A. Kaczmarczyk. Robustness of participatory budgeting outcomes: Complexity and experiments. In *Proceedings of the 16th International Symposium on Algorithmic Game Theory (SAGT)*, pages 161–178, 2023a. [→ p. 8]
- S. J. Brams and D. M. Kilgour. Satisfaction approval voting. In *Voting Power and Procedures, Studies in Choice and Welfare*, pages 323–346. Springer, 2014. [→ p. 4]
- R. Bredereck, P. Faliszewski, R. Niedermeier, and A. Kaczmarczyk. An experimental view on committees providing justified representation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 109–115, 2019. [→ p. 3, 17]
- R. Bredereck, P. Faliszewski, A. Kaczmarczyk, R. Niedermeier, P. Skowron, and N. Talmon. Robustness among multiwinner voting rules. *Artificial Intelligence*, 290:103403, 2021. [→ p. 8]
- M. Brill and J. Peters. Robust and verifiable proportionality axioms for multiwinner voting. In *Proceedings of the 24th ACM Conference on Economics and Computation (ACM-EC)*, page 301. ACM, 2023. Full version arXiv:2302.01989 [cs.GT]. [→ p. 2, 3, 10, 11, 17]

- M. Brill, R. Freeman, S. Janson, and M. Lackner. Phragmén’s voting methods and justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 406–413. AAAI Press, 2017. [→ p. 5]
- M. Brill, J.-F. Laslier, and P. Skowron. Multiwinner approval rules as apportionment methods. *Journal of Theoretical Politics*, 30(3):358–382, 2018. [→ p. 2]
- M. Brill, P. Gölz, D. Peters, U. Schmidt-Kraepelin, and K. Wilker. Approval-based apportionment. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1854–1861. AAAI Press, 2020. [→ p. 15]
- M. Brill, P. Gölz, D. Peters, U. Schmidt-Kraepelin, and K. Wilker. Approval-based apportionment. *Mathematical Programming*, 2023. [→ p. 28]
- J. Burdges, A. Cevallos, P. Czaban, R. Habermeier, S. Hosseini, F. Lama, H. K. Alper, X. Luo, F. Shirazi, A. Stewart, and G. Wood. Overview of Polkadot and its design considerations. Technical report, arXiv:2005.13456 [cs.CR], 2020. [→ p. 2, 5, 14]
- A. Cevallos and A. Stewart. A verifiably secure and proportional committee election rule. In *Proceedings of the 3rd ACM Conference on Advances in Financial Technologies (AFT)*, pages 29–42. ACM, 2021. [→ p. 2, 5, 6, 10, 11, 12, 13, 25]
- N. Chong. Crypto ratings agency downgrades eos for “serious centralization problems”. <https://www.newsbtc.com/news/eos/crypto-ratings-agency-downgrades-eos-serious-centralization-problems/>, 2019. [→ p. 3]
- Crypto Carbon Ratings Institute. Pos benchmark study 2023. carbon-ratings.com/dl/pos-report-2023, 2023. [→ p. 18]
- E. Elkind, P. Faliszewski, J.-F. Laslier, P. Skowron, A. Slinko, and N. Talmon. What do multiwinner voting rules do? An experiment over the two-dimensional Euclidean domain. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 494–501. AAAI Press, 2017. [→ p. 2, 3]
- P. Faliszewski, J. Flis, D. Peters, G. Pierczynski, P. Skowron, D. Stolicki, S. Szufa, and N. Talmon. Participatory budgeting: Data, tools and analysis. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2667–2674. ijcai.org, 2023a. [→ p. 3]
- P. Faliszewski, M. Lackner, K. Sornat, and S. Szufa. An experimental comparison of multiwinner voting rules on approval elections. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2675–2683, 2023b. [→ p. 2, 3, 9, 17]
- P. Garg. Eos voting structure encourages centralization. <https://cryptoslate.com/eos-voting-structure-encourages-centralization/>, 2019. [→ p. 3]
- G. Gawron and P. Faliszewski. Using multiwinner voting to search for movies. In *Proceedings of the 19th European Conference on Multi-Agent Systems (EUMAS)*, pages 134–151. Springer, 2022. [→ p. 2]
- J. Gehrlein, G. Miebs, M. Brunelli, and M. Kadziński. An active preference learning approach to aid the selection of validators in blockchain environments. *Omega*, 118:102869, 2023. [→ p. 6]

- I. Grigg. EOS – An introduction. <https://whitepaperdatabase.com/eos-whitepaper>, 2017. White Paper. [→ p. 3]
- S. Janson. Phragmén’s and Thiele’s election methods. Technical report, arXiv:1611.08826 [math.HO], 2016. [→ p. 4, 5]
- M. Lackner. Perpetual voting: Fairness in long-term decision making. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, pages 2103–2110. AAAI Press, 2020. [→ p. 6]
- M. Lackner and P. Skowron. *Multi-Winner Voting with Approval Preferences*. Springer, 2022. [→ p. 2, 3, 10, 28]
- L. Lamport, R. Shostak, and M. Pease. The byzantine generals problem. In *Concurrency: The Works of Leslie Lamport*, pages 203–226. 2019. [→ p. 5]
- N. Mattei and T. Walsh. Preflib: A library for preferences <http://www.preflib.org>. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*, pages 259–270. Springer, 2013. [→ p. 3]
- N. Mattei and T. Walsh. A preflib.org retrospective: Lessons learned and new directions. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 15, pages 289–305. AI Access, 2017. [→ p. 3, 6]
- K. Mehra, N. K. Sreenivas, and K. Larson. Deliberation and voting in approval-based multi-winner elections. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2853–2861, 2023. [→ p. 2, 3]
- S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008. [→ p. 5]
- M. C. Pease, R. E. Shostak, and L. Lamport. Reaching agreement in the presence of faults. *Journal of the ACM*, 27(2):228–234, 1980. [→ p. 12]
- D. Peters and P. Skowron. Proportionality and the limits of welfarism. In *Proceedings of the 21st ACM Conference on Economics and Computation (ACM-EC)*, pages 793–794. ACM, 2020. Full version arXiv:1911.11747 [cs.GT]. [→ p. 5, 11]
- D. Peters, G. Pierczyński, and P. Skowron. Proportional participatory budgeting with additive utilities. In *Advances in Neural Information Processing Systems*, volume 34, pages 12726–12737, 2021. [→ p. 3]
- E. Phragmén. Sur une méthode nouvelle pour réaliser, dans les élections, la représentation proportionnelle des partis. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 51(3):133–137, 1894. [→ p. 5]
- A. Pommerening, M. Brill, U. Schmidt-Kraepelin, and J. Haufe. Democratising forest management: Applying multiwinner approval voting to tree selection. *Forest Ecology and Management*, 478:118509, 2020. [→ p. 2]
- J. Reichert and E. Elkind. Agreement among proportional approval-based multi-winner voting rules: An empirical analysis, 2023. Presented at the 5th Games, Agents, and Incentives Workshop (GAIW). [→ p. 3, 9]
- L. Sánchez-Fernández, E. Elkind, M. Lackner, N. Fernández, J. A. Fisteus, P. Basanta Val, and P. Skowron. Proportional justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 670–676. AAAI Press, 2017. [→ p. 3, 10, 26]

- L. Sánchez-Fernández, N. Fernández, J. A. Fisteus, and M. Brill. The maximin support method: An extension of the D'Hondt method to approval-based multiwinner elections. *Mathematical Programming*, 2022. Forthcoming. [→ p. 5, 14]
- P. Skowron. Proportionality degree of multiwinner rules. In *Proceedings of the 22nd ACM Conference on Economics and Computation (ACM-EC)*, pages 820–840. ACM, 2021. [→ p. 3, 11]
- P. Skowron and A. Górecki. Proportional public decisions. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, pages 5191–5198. AAAI Press, 2022. [→ p. 2]
- N. Stifter, A. Judmayer, P. Schindler, A. Zamyatin, and E. Weippl. Agreement with satoshi – on the formalization of nakamoto consensus. Cryptology ePrint Archive, Paper 2018/400, 2018. URL <https://eprint.iacr.org/2018/400>. <https://eprint.iacr.org/2018/400>. [→ p. 12]
- E. Strevniotis and G. Chalkiadakis. Preference aggregation mechanisms for a tourism-oriented bayesian recommender. In *Proceedings of the 24th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA)*, pages 331–346. Springer, 2022a. [→ p. 2]
- E. Strevniotis and G. Chalkiadakis. Multiwinner election mechanisms for diverse personalized bayesian recommendations for the tourism domain. In *Proceedings of the Workshop on Recommenders in Tourism (RecTour)*, pages 65–82. CEUR-WS.org, 2022b. [→ p. 2]
- S. Szufa, P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Drawing a map of elections in the space of statistical cultures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1341–1349. IFAAMAS, 2020. [→ p. 3]
- S. Szufa, P. Faliszewski, L. Janeczko, M. Lackner, A. Slinko, K. Sornat, and N. Talmon. How to sample approval elections? In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 496–502, 2022. [→ p. 2, 3]
- T. N. Thiele. Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441. 1895. [→ p. 4]
- G. Wood. Polkadot: Vision for a heterogeneous multi-chain framework. <https://assets.polkadot.network/Polkadot-whitepaper.pdf>, 2016. White Paper. [→ p. 5]

Appendix

A. Additional Material for Section 4

	weighted satisfaction	max. approval weight loser	min. average satisfaction		
			$\ell = 1$	$\ell = 5$	$\ell = 10$
AV	9.055	0.011	0.000	—	—
SAV	8.724	0.022	0.011	4.359	—
seq-PAV	8.672	0.028	1.092	7.442	14.124
Phragmms	8.650	0.039	0.960	6.840	13.510
seq-Phrag.	8.588	0.039	0.990	6.678	13.468
MES	8.584	0.039	0.980	6.773	13.438

Table 4: Average values for different notions related to voter satisfaction and preventing underrepresentation.

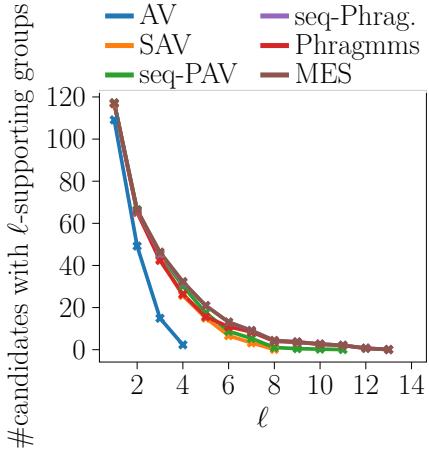


Figure 7: Number of non-elected candidates with ℓ -supporting groups. Only non-zero values are plotted.

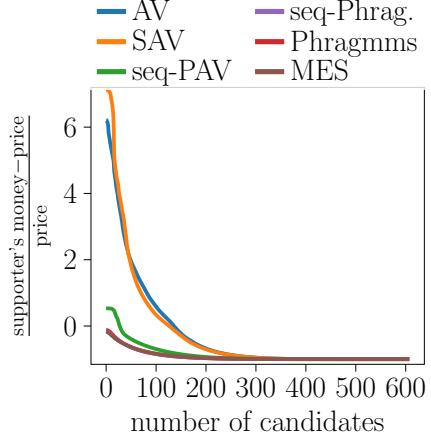


Figure 8: Average amount by which supporter's money exceeds price for priceability system.

In this section, we provide some more information regarding the satisfaction and representation of voters (see Table 4 for an overview). We start by considering the average weighted satisfaction of the voters; see the first column of Table 4. AV maximizes this objective; the other rules perform between 4% and 5.3% worse, with SAV producing the best and MES producing the worst result. In general, the observed weighted satisfaction values are remarkably high given that voters approve less than 10 candidates on average.

To give some more context on the minimum average satisfaction of ℓ -supporting groups (see Table 4), Figure 7 depicts the average number of non-selected candidates for which there is an ℓ -supporting group in our elections. In Figure 7, we observe that for all rules, the number of non-selected candidates for which an ℓ -supporting group exists quickly decreases with increasing ℓ . We see a contrast between AV and the other rules, especially with respect to the size of the largest supporting group. For AV, there are no ℓ -supporting groups for $\ell > 4$, implying that having an approval weight larger than $\frac{5}{300}$ is always sufficient to be included in the committee. For the other rules, this is not the case, as here there also exist ℓ -supporting groups for more candidates and for larger values of ℓ . While generally speaking all other rules show a similar performance, SAV

and seq-PAV tend to produce slightly less non-selected candidates with ℓ -supporting groups. This behavior is also reflected in the maximum approval weight of a non-selected candidate shown in column two of Table 4.

B. Additional Material for Section 5

B.1. Computing Candidate Groups with Minimum Approval Weight

To generate Figure 6a, we needed to compute size- ℓ candidate sets with minimum approval weight. We prove that this problem is NP-hard and give an ILP formulation, which we used in our experiments.

Proposition 1. *Given an election (C, V, A, w, k) , a committee W , and an integer $\ell \in [|W|]$, computing the size- ℓ subset $W' \subseteq W$ with minimum approval weight, i.e., $\arg \min_{W' \subseteq W : |W'|=\ell} w(W')$, is NP-hard.*

Proof. We reduce from CLIQUE on regular graphs. Given an r -regular graph $G = (U, E)$ and an integer t , we construct our election as follows. We add a candidate c_u for each $u \in U$ and a voter for each edge $e \in E$ approving the two candidates corresponding to the vertices incident to e . All voters have weight 1, we set W to be the set of all candidates and ask whether there is a size- t set $W' \subseteq W$ of candidates with $w(W') \leq r \cdot t - \binom{t}{2}$.

Assume that $U' \subseteq U$ is a size- t clique in G , then $W' := \{c_u : u \in U'\}$ fulfills $w(W') \leq r \cdot t - \binom{t}{2}$, as there are $\binom{t}{2}$ voters that approve two candidates from W' .

Assume that we are given a size- t candidate set W' with $w(W') \leq r \cdot t - \binom{t}{2}$, then $U' := \{u : c_u \in W'\}$ is a size- t clique in G . This follows as $w(W') \leq r \cdot t - \binom{t}{2}$ implies that there must be $\binom{t}{2}$ voters approving two candidates from W' , implying that there are $\binom{t}{2}$ edges with two endpoints in U' . \square

To solve the problem in our experiments, we model it as an Integer Linear Program (ILP) and solve it using Gurobi. We add a binary variable x_c for each $c \in C$ which is 1 if c is selected as part of the set and 0 otherwise. We impose that a size- ℓ candidate set is selected by requiring that:

$$\sum_{c \in C} x_c = \ell.$$

Further, we add a real variable y_v for each $v \in V$ which is intended to be 0 if v approves of any of the selected candidates and 1 otherwise. We enforce these upper bounds of 0 or 1 by adding for each voter $v \in V$ and each candidate $c \in A_v$ the following constraint:

$$x_c + y_v \leq 1,$$

while lower-bound constraints are not needed. Finally, we add the following objective function to enforce that a subset of minimum approval weight is selected:

$$\min \sum_{v \in V} (1 - y_v) \cdot w(v).$$

B.2. Proof of Theorem 1

We now formally define the minimum exogenous cost of an ℓ -replacement, which was informally introduced in Section 5. Consider some weighted multiwinner election $E = (C, V, A, w, k)$. We call any (C', V', A', w') with $|C'| = \ell$ and $C \cap C', V \cap V' = \emptyset$ an ℓ -replacement. Its weight is $w(V') = \sum_{v \in V'} w'(v)$, and we say it is *successful* for some voting rule if $C' \subseteq W$, where W is the

committee selected by the rule on the extended election $(C \cup C', V \cup V', A \cup A', w \cup w', k)$, where $w \cup w'$ denotes the natural concatenation of the two weight functions w and w' . The *minimum exogenous cost of an ℓ -replacement* is the minimum weight of a successful ℓ -replacement.

In order to compute a minimum ℓ -replacement for all $\ell \in [300]$ and all of our tested elections, which we depicted in Figure 6c, we derived algorithms for AV, SAV, seq-Phragmén, and Phragmms. When reporting the running time for any of these algorithms, it is important to point out that we assume that certain information which is naturally derived from the execution of a voting rule on the original instance, can be taken as input for our computations. Since this information may vary for each voting rule, we specify the required information in each theorem statement.

Since all considered rules are sequential, we can order the selected candidates chronologically: for a given voting rule, let c_1 be the first selected candidate, let c_2 be the second selected candidate, and so on. Note that this labeling depends on the rule under consideration.

Theorem 2. *For every $\ell \in [k]$, the minimum exogenous cost of an ℓ -replacement for AV can be computed in time $\mathcal{O}(1)$, assuming that we get as input the approval weight of the candidate $c = c_{k-\ell+1}$, i.e., $w(V_c)$.*

Proof. Let (C', V', A', w') be an ℓ -replacement for election (C, V, A, w, k) . It is successful for AV if and only if

$$w'(V'_{c'}) \geq w(V_{c_{k-\ell+1}})$$

holds for all $c' \in C'$.¹¹ Thus, an ℓ -replacement that consists, for example, of one voter with weight $w(V_{c_{k-\ell+1}})$ approving all ℓ candidates in C' , is clearly successful and weight minimal. \square

Theorem 3. *For every $\ell \in [k]$, the minimum exogenous cost of an ℓ -replacement for SAV can be computed in time $\mathcal{O}(1)$, assuming that we get as input the SAV score of the candidate $c = c_{k-\ell+1}$, i.e., $x := \sum_{v \in V_{c_{k-\ell+1}}} \frac{w(v)}{|A_v|}$.*

Proof. Let (C', V', A', w') be an ℓ -replacement for election (C, V, A, w, k) . It is successful for SAV if and only if

$$\sum_{v \in V'_{c'}} \frac{w'(v)}{|A_v|} \geq \sum_{v \in V_{c_{k-\ell+1}}} \frac{w(v)}{|A_v|} = x$$

holds for all $c' \in C'$. Thus, an ℓ -replacement that consists, for example, of ℓ voters each with weight x and approving a distinct candidate in C' , is clearly successful and weight minimal, with weight $\ell \cdot x$. \square

Theorem 4. *For every $\ell \in [k]$, the minimum exogenous cost of an ℓ -replacement for seq-Phragmén can be computed in time $\mathcal{O}(1)$, assuming that we get as input the point in time t at which candidate $c_{k-\ell+1}$ is elected.*

Proof. Let $t \in \mathbb{R}$ be the point in time at which candidate $c_{k-\ell+1}$ is elected. A necessary condition for an ℓ -replacement (C', V', A', w') to be successful is that voters in V' earned at least ℓ units of money until time t , which is the case if and only if $\sum_{v \in V'} w'(v) \geq \frac{\ell}{t}$. To see that this is indeed the minimum exogenous cost of an ℓ -replacement, consider an ℓ -replacement that consists of ℓ voters of weight $\frac{1}{t}$ each, each approving exactly one (distinct) candidate in C' . In the extended election, candidates $c_1, \dots, c_{k-\ell}$ are elected with unchanged order, and then all candidates in C' will be chosen before candidate $c_{k-\ell+1}$, hence the ℓ -replacement is successful. \square

¹¹We assume in this section that any election rule breaks ties in a way that benefits the candidates in the ℓ -replacement.

Theorem 5. For every $\ell \in [k]$, the minimum exogenous cost of an ℓ -replacement for Phragmms can be computed in time $\mathcal{O}(1)$, assuming that we get as input the value $\min\{s_1, \dots, s_{k-\ell}\}$, where s_i is the Phragmms score of candidate c_i for any $i \in [k]$.

Proof. For the definition of candidate scores in Phragmms, we refer to Cevallos and Stewart [2021]. For election (C, V, A, w, k) , recall that any ℓ -replacement (C', V', A', w') is “properly separated,” in the sense that $C \cap C' = \emptyset$ and $A \cap A' = \emptyset$. As a consequence (and similarly to the previous election rules), the outcome of Phragmms can be computed by merging its executions on each instance separately. More precisely, if c_1, \dots, c_k are the elected candidates in (C, V, A, w, k) labeled by election order, and s_i is the score of c_i at the time of its election, and similarly if $(c'_1, s'_1), \dots, (c'_\ell, s'_\ell)$ are the candidates and their scores when we run the election (V', C', A', w', ℓ) , then we claim that in an execution of Phragmms on the extended instance $(V \cup V', C \cup C', A \cup A', w \cup w', k)$, the selection order is determined by starting with the empty committee and repeating the following process k times: Let $i \in [k]$ be the smallest index of an unelected candidate in $\{c_1, \dots, c_k\}$ and let $j \in [\ell]$ be the smallest index of an unelected candidate in $\{c'_1, \dots, c'_\ell\}$; if $s_i > s'_j$, elect c_i , otherwise elect c_j . To prove the claim, it suffices to verify that for each subinstance of the extended instance, neither the updating of the scores nor the rebalancing step of the vote assignment is influenced by the other subinstance. This fact can be easily checked when going through Section 4 in Cevallos and Stewart [2021].

Having established this claim, we aim to find a weight-minimal ℓ -replacement. We first show that we can assume without loss of generality, that such a replacement has a simple structure, namely, consists of one voter who approves all of the ℓ candidates in C' . To see why, consider any ℓ -replacement of some weight $x := \sum_{v \in V'} w'(v)$, and let s'_1, \dots, s'_ℓ be the corresponding score sequence. Lemma 18 by Cevallos and Stewart [2021] states that, at every iteration i , the minimum backing weight of the vote assignment after iteration $i - 1$ is at least as the score s'_i . Moreover, since no backing weight is reduced below s'_i in iteration i , this implies an upper bound of x/i on s'_i . Considering the trivial ℓ -replacement of weight x , i.e., one voter of weight x approving all candidates in C' , we observe that the resulting score sequence corresponds exactly to $x/1, x/2, \dots, x/\ell$. Therefore, the existence of some ℓ -replacement of weight x implies the fact that there also exists such a trivial ℓ -replacement of weight x .

It remains to determine the minimum weight of some ℓ -replacement. For now, we call this weight x . Any ℓ -replacement is successful if and only if candidate c'_ℓ gets elected before candidate $c_{k-\ell+1}$. This is the case if and only if $s'_\ell = \frac{x}{\ell} \geq \min\{s_1, \dots, s_{k-\ell+1}\}$. Thus, a minimum ℓ -replacement has weight

$$z = \ell \cdot \min\{s_1, \dots, s_{k-\ell+1}\},$$

proving the claim. Lastly, we remark that due to the last argument, the following ℓ -replacement is successful as well: There are ℓ voters, each with weight $\frac{z}{\ell}$ and each approving a distinct candidate in C' . \square

Sequential PAV Maybe surprisingly, the task of computing the minimum exogenous cost for replacing ℓ candidates for seq-PAV leads to an interesting theoretical question, which may be of independent interest for future work. In the following, we report on our observations so far, and provide a linear program with which we were able to solve the question for all $\ell \leq 21$.

For a given election (C, V, A, w, k) , let c_i be the i^{th} candidate that is selected by seqPAV and let s_i be the marginal PAV score of candidate c_i at the time of selection. More precisely, this means that

$$s_i = \text{sc}(\{c_1, \dots, c_i\}) - \text{sc}(\{c_1, \dots, c_{i-1}\}),$$

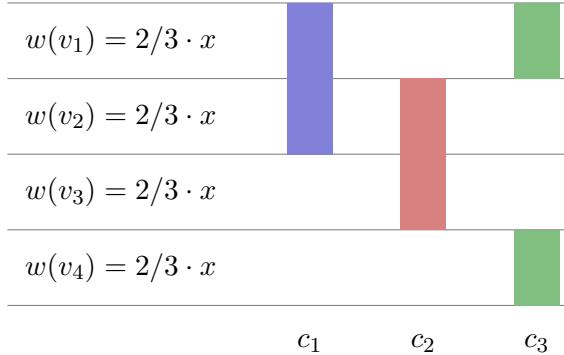


Figure 9: Illustration of an optimal 3-replacement for Seq-PAV, with a total weight $\frac{8}{3}x$, in which the 3rd candidate has marginal sequential PAV score of x . The approval ballots are interpreted as follows: $A_{v_1} = \{c_1, c_3\}$, $A_{v_2} = \{c_1, c_2\}$, $A_{v_3} = \{c_2\}$ and $A_{v_4} = \{c_3\}$.

where we recall that the function sc is defined as

$$\text{sc}(W) = \sum_{v \in V} w(v) \sum_{j=1}^{|A_v \cap W|} \frac{1}{j}.$$

Similarly, for any (potential) ℓ -replacement (C', V', A', w') , let c'_i be the candidate selected at point $i \in [\ell]$ in the election (C', V', A', w') and s'_i be the marginal PAV score of this candidate at the point of selection. Then, the ℓ -replacement is successful, if and only if $s_{k-\ell+1} < s'_\ell$. The reason for this is twofold: (1) Both sequences s_1, \dots, s_k and s'_1, \dots, s'_ℓ are monotone. (2) Clearly, when running the extended election, the two subselections behave completely independently from each other.

Hence, the question of computing the minimum exogenous cost for replacing ℓ candidates reduces to the following task: Given $\ell \in \mathbb{N}$ and $x := s_{k-\ell+1} \in \mathbb{R}$, find an election (C', V', A', w', ℓ) of minimum weight, such that the marginal contribution of the ℓ^{th} selected candidate is at least x . For all previous voting rules, we were always able to restrict ourselves to ℓ -replacements with a very simple structure, e.g., one voter approving all candidates in C' . Let us try the same by considering two natural simple election structures in the context of seq-PAV. In the first, consider one voter v with weight $w(v)$ approving all ℓ candidates in C' . Then, the score sequence corresponds to $s'_1 = w(v), \dots, s'_\ell = \frac{w(v)}{\ell}$. Hence the minimum successful weight when we restrict ourselves to this election format is $w(v) = \ell \cdot x$. Similarly, if we consider ℓ voters, all of which approve one of the candidates in C' each, then the score sequence corresponds to $s'_1 = w(v_1), \dots, s'_\ell = w(v_\ell)$, hence, the minimum weight also corresponds to $\sum_{v \in V} w(v) = \ell \cdot x$. Given those two extremes, one might be tempted to conjecture that $\ell \cdot x$ is the correct answer, however, this is not the case since there can be non-trivial synergy effects between the candidates. For intuition, we give an example for $\ell = 3$ in Figure 9.

When constructing minimum ℓ -replacements for seq-PAV, there is a non-trivial trade-off between reusing the (reduced) approval weight of already selected candidates and not violating the greedy selection order. In the following, we describe a linear program¹² that optimally solves this trade-off, albeit having exponentially many variables. Before we define the LP, we define for any $S \subseteq C'$ and $i \in [\ell]$ the value

$$t(S, i) = (|S \cap \{c_1, \dots, c_{i-1}\}| + 1)^{-1}.$$

¹²The underlying ideas of this LP are similar to those used by Sánchez-Fernández et al. [2017] to prove that seq-PAV satisfies JR when $k \leq 5$, but fails it when $k \geq 6$.

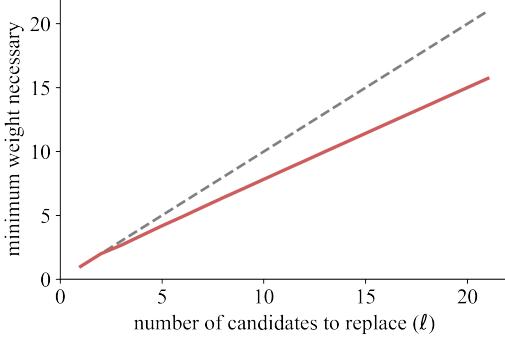


Figure 10: Minimum exogenous cost for replacing ℓ candidates when normalizing the marginal PAV score of the $k - \ell + 1^{\text{th}}$ candidate in the original election to be one. The grey dashed line indicates the identity, which corresponds to the minimum exogenous cost if we restrict ourselves to ℓ -attacks with simple structures (e.g., those with only one voter or those with voters approving one candidate each).

The value corresponds to the marginal contribution of a voter to the PAV score at the moment when candidate c_i is elected, assuming that the voter has weight one and approves all candidates in S . The LP has one variable per subset $S \subseteq C'$, called $y(S)$, which we interpret as the weight of the voter approving exactly the candidates in S . Note that this entirely describes any election with candidate set C' . We arbitrarily index the elements in C' by c_1, \dots, c_ℓ and the first set of inequalities of the LP enforces that the candidates are elected by seq-PAV in this order. That is, for any pair $i, j \in [\ell]$ with $i < j$, the marginal contribution of i (at iteration i) is at least as large as the marginal contribution of j at iteration i . Lastly, the other non-trivial constraint implies that candidate c_ℓ has marginal contribution at least x in iteration ℓ .

$$\begin{aligned}
& \min \sum_{S \subseteq C'} y(S) \\
& \sum_{S \subseteq C', c_i \in S} t(S, i) \cdot y(S) \geq \sum_{S \subseteq C', c_j \in S} t(S, j) \cdot y(S) \\
& \quad \forall i, j \in [\ell], j > i \\
& \sum_{S \subseteq C', c_\ell \in S} t(S, \ell) \cdot y(S) \geq x \\
& y(S) \geq 0 \quad \forall S \subseteq C'
\end{aligned}$$

Note that this LP is invariant to scaling x since an optimal solution $y \in \mathbb{R}^{[0,1]^\ell}$ for input $x = 1$ lets us derive an optimal solution y' for any x' by setting $y'(S) = x' \cdot y(S)$ for all $S \subseteq C'$. Hence, we restricted ourselves to the case $x = 1$ and computed the solution of the LP for all values $\ell \in [21]$. We report these values in Figure 10.

The formulation as an LP also provides us with a non-trivial structural insight. That is, there always exists an optimal ℓ -replacement, with at most $\frac{\ell(\ell-1)}{2} + 1$ voters. We obtain this bound by the following observation: The LP has 2^ℓ variables, 2^ℓ “trivial” constraints (those of form $y(S) \geq 0$) and $\frac{\ell(\ell-1)}{2} + 1$ non-trivial constraints. Now, for any extreme point of the LP, there have to be at least 2^ℓ tight constraints. Hence, at least $2^\ell - \frac{\ell(\ell-1)}{2} + 1$ of the trivial constraints have to be tight, implying an upper bound of at most $\frac{\ell(\ell-1)}{2} + 1$ non-zero variables. Nevertheless, finding these variables (corresponding to voters) may still be a challenging task.

	PAV score		max. support loser		min. average satisfaction		priceability gap	
	$\ell = 1$	$\ell = 5$						
	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP
	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$
	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$
AV	2.502 2.075 2.324	0.011 0.021 0.017	0.000 0.0038 0.0066	0.00 0.120 0	3.357 - -	- -	6.21 9.421 7.621	3.503 10.234 6.301
SAV	2.547 2.082 2.348	0.022 0.035 0.027	0.011 0.006 0.0120	4.359 3.145 0.67	7.14 - 12.02	10.234 8.886 4.464	1.854 8.886 4.464	
seq-PAV	2.585 2.284 2.452	2.803 2.718 2.673	0.028 0.044 0.042	0.084 0.0076 0.0150	1.092 0.99 1.05	0.86 5.069 2.123	7.442 6.677 6.722	5.171 14.277 11.08
Phragmms	2.578 2.271 2.441	2.777 2.712 2.666	0.039 0.043 0.043	0.084 0.0110 0.0200	0.960 0.74 0.84	0.56 1.901 1.047	6.840 6.652 6.315	5.171 14.277 11.08
seq-Phrag.	2.578 2.277 2.443	2.802 2.713 2.667	0.039 0.043 0.043	0.084 0.0110 0.0210	0.990 0.91 1.0010	0.75 2.422 1.111	6.678 6.324 6.25	4.610 9.826 9.635
MES	2.579 2.276 2.443	2.777 2.713 2.668	0.039 0.043 0.043	0.084 0.0110 0.0210	0.980 0.94 0.99	0.71 2.460 1.12	6.773 6.316 6.311	5.110 11.238 9.731
							-0.12 -0.083 -0.042	-0.033 -0.550 -0.210
							-0.16 -0.078 -0.044	-0.035 -0.580 -0.240
							-0.17 -0.077 -0.042	-0.033 -0.590 -0.24

Table 5: Average values for different notions related to preventing underrepresentation.

Lastly, we remark that the LP might be solvable in polynomial time by a standard *column generation* approach. That is, since our LP has polynomial many constraints, deriving the dual LP results in an LP with polynomial many variables and exponential many constraints. Now, if we could find an algorithm for the corresponding *separation problem* (i.e., given a solution to the dual LP, determine in polynomial time whether a constraint is violated), then we could employ the Ellipsoid method, to solve the problem in polynomial time. In fact, we derived the dual LP and considered the separation problem, however, we were not able to find a polynomial-time algorithm nor show hardness of the separation problem.

Method of Equal Shares For MES, the situation is less clear, as the initial budget of a voter depends on the total weight of voters. Thus, by adding new voters the whole execution of the rule might change. Even worse, MES is not candidate monotone with additional voters [Lackner and Skowron, 2022], which means that adding an additional supporter for a currently selected candidate can actually lead to the candidate no longer being selected.

C. Additional Material for Section 6

Table 5 (on page 28) shows some of our measures regarding the prevention of underrepresentation. We compare different committee sizes $k \in \{200, 250, 300, 350, 400\}$ and whether candidates can be selected multiple times for $k = 300$ (following the terminology of Brill et al. [2023], we call this setting “P-APP” setting; also see their work for a formal definition). Table 6 shows the same comparison for measures regarding overrepresentation.

Changing the Committee Size. All rules are similarly affected by changes in the committee size, which is why we only report general trends focusing on the reported average values. Increasing the committee size from 300 to 400 typically leads to

- a small ($\approx 7.5\%$) increase of the PAV score,

maximin value		maximin variance		min. support winner		cost of replacing				min. approval weight				
						one-third	half			one-third	half			
	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP	$k = 300$	P-APP
	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$	$k = 200$	$k = 400$
	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$	$k = 250$	$k = 350$
AV	0.0015		0.0047		0.0110		0.14	0.27	0.22		0.36			
	0.0018	0.00076	0.0060	0.0056	0.0210	0.0038	0.14	0.14	0.24	0.27	0.16	0.22	0.26	0.35
	0.0017	0.00087	0.0049	0.0057	0.017	0.0066	0.15	0.14	0.24	0.26	—	—	—	—
SAV	0.0018		0.0043		0.0024		0.24	0.4	0.24		0.38			
	0.0026	0.00068	0.0051	0.0060	0.0028	0.0017	0.18	0.25	0.29	0.48	0.18	0.23	0.29	0.38
	0.0023	0.00120	0.0043	0.0049	0.0028	0.002	0.21	0.26	0.35	0.45	—	—	—	—
seq-PAV	0.0024	0.003	0.0031	0.00063	0.0028	0.0032	?	?	?	-?	0.26	0.31	0.42	0.47
	0.0038	0.00074	0.0031	0.0066	0.0051	0.0018	?	?	?	?	0.28	0.23	0.44	0.38
	0.0030	0.0014	0.0030	0.0044	0.0032	0.002	?	?	?	?	—	—	—	—
Phragmms	0.00272	0.003	0.0023	0.00056	0.0028	0.0032	0.4	0.42	0.79	0.82	0.28	0.31	0.44	0.46
	0.0040	0.00140	0.0023	0.0040	0.0058	0.0020	0.38	0.4	0.74	0.8	0.29	0.25	0.45	0.38
	0.0033	0.002	0.0021	0.0032	0.0036	0.0027	0.39	0.4	0.77	0.79	—	—	—	—
seq-Phag.	0.0027	0.003	0.0024	0.00066	0.0028	0.0032	0.4	0.42	0.79	0.82	0.28	0.31	0.44	0.47
	0.0041	0.00120	0.0021	0.0042	0.0058	0.0020	0.39	0.4	0.75	0.8	0.29	0.25	0.45	0.38
	0.0033	0.0019	0.0021	0.0033	0.0036	0.0026	0.4	0.4	0.78	0.8	—	—	—	—
MES	0.00269	0.003	0.0024	0.00055	0.0028	0.0032	?	?	?	?	0.28	0.31	0.44	0.47
	0.0041	0.00120	0.0021	0.0042	0.0057	0.0019	?	?	?	?	0.28	0.25	0.45	0.38
	0.0032	0.0019	0.0022	0.0033	0.0036	0.0026	?	?	?	?	—	—	—	—

Table 6: Average values for different notions related to preventing overrepresentation. Entries marked with “—” were not computed due to a shortage of computation time. “Maximin variance” stands for the variance of the candidate’s backing weight in the maximin assignment.

- a substantial (typically at least $\approx 100\%$) increase of the minimum average satisfaction,
- a substantial decrease in the priceability gap,
- a decrease of the maximum support value by typically at least around 50%,
- no significant change in terms of the cost of replacing, and
- a small ($\approx 5\%$) decrease of the minimum approval weight of size- $\frac{k}{3}$ and size- $\frac{k}{2}$ subsets of the committee.

Decreasing the committee size from 300 to 200 typically leads to

- a at least 11% decrease of the PAV score,
- small drops in the average satisfaction for the proportional rules,
- a small increase in the priceability gap,
- an around $\approx 50\%$ increase of the maximin value for the proportional rules, yet no strong increase for AV,
- no significant change of the exogenous cost of replacing (except for SAV where values drop by $\approx 25\%$), and
- no significant change of the minimum approval weight of size- $\frac{k}{3}$ and size- $\frac{k}{2}$ subsets of the committee, except for AV and SAV where values drop by around 25%.

Results for $k = 350$ almost always lie in between those for $k = 300$ and $k = 400$, and results for $k = 350$ in between those for $k = 200$ and $k = 300$. In particular, all reported measures are either not strongly affected by the committee size or behave monotonically.

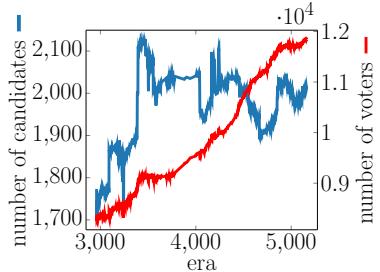


Figure 11: Number of candidates (blue) and voters (red) in our 1520 Kusama elections.

	AV	SAV	seq-PAV	Phragmms	seq-Phrag.	MES
AV	—	913.50	912.65	931.05	916.75	913.65
SAV	913.50	—	972.05	938.05	945.35	944.60
seq-PAV	912.65	972.05	—	941.40	956.85	957.30
Phragmms	931.05	938.05	941.40	—	951.85	948.10
seq-Phrag.	916.75	945.35	956.85	951.85	—	984.60
MES	913.65	944.60	957.30	948.10	984.60	—

Table 7: Average overlap between committees returned by different voting rules. The committee size is 1000.

	weighted satisfaction	PAV score	max. approval weight loser	JR violations	EJR+ violations	min. average satisfaction $\ell = 1$	average satisfaction $\ell = 5$	satisfaction $\ell = 10$	priceability violations	priceability gap
AV	14.966	3.042	0.0042	35.95	76.4	0.000	1.479	—	148.25	7.225
SAV	14.608	3.084	0.0100	20.25	43.5	0.430	3.995	15.761	137.90	9.702
seq-PAV	14.553	3.102	0.0120	0.00	0.0	2.178	9.192	17.041	64.45	4.358
Phragmms	14.586	3.090	0.0150	0.00	0.0	1.113	7.092	14.250	0.00	-0.060
seq-Phrag.	14.508	3.094	0.0150	0.00	0.0	1.526	7.620	14.514	0.00	-0.095
MES	14.476	3.094	0.0150	0.00	0.0	1.526	7.738	14.728	0.00	-0.071

Table 8: Average values for different notions related to preventing underrepresentation.

Selecting Candidates Multiple Times. AV and SAV do not admit natural translations to the P-APP setting, which is why we only report results on the four proportional rules here. Notably, for the definition of ℓ -supporting groups and priceability, all candidates are viewed as being non-selected (as we could always include an additional copy of a candidate in the committee). In general, we observe that allowing candidates to be selected multiple times leads to

- an increase in the PAV score (even when compared to the values for $k = 400$),
- a decrease in the minimum average satisfaction of ℓ -supporting groups (which, as discussed in the main body, can be explained because now also candidates which are included in the committee count as non-selected),
- an increase in the priceability gap (again same explanation as for ℓ -supporting groups),
- an increase of the maximin support value (which is roughly at the level as for $k = 250$),
- a substantial decrease of the maximin variance by around 75%, which means that the backing weight of candidates is much more uniform here, and
- a small increase in the cost of replacing and in the minimum approval weight of size- $\frac{k}{3}$ and size- $\frac{k}{2}$ subsets of the committee.

D. Kusama Elections

We now give a brief description and experimental evaluation of the elections from the Kusama network. The network follows a very similar protocol to the one for Polkadot. However, there are a few differences. First, in Kusama, multiple elections are conducted each day. Because of this, we were able to collect 1520 elections conducted on the Kusama network. Moreover, in Kusama

	maximin support value	min. appr. weight of winner	cost of “replacing” ℓ		
			$\ell = 1$	$\ell = \frac{k}{3}$	$\ell = \frac{k}{2}$
AV	0.00047	0.00420	0.0042	0.22	0.38
SAV	0.00055	0.00079	0.00047	0.24	0.41
seq-PAV	0.00062	0.00079	?	?	?
Phragmms	0.00080	0.00082	0.0008	0.43	0.85
seq-Phrag.	0.00078	0.00081	0.00078	0.43	0.86
MES	0.00079	0.00082	?	?	?

Table 9: Measures related to overrepresentation.

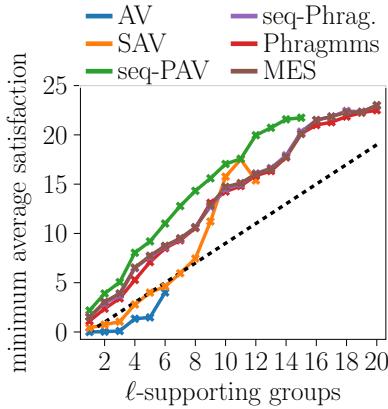


Figure 12: Minimum average satisfaction of ℓ -supporting groups. The dashed line is the function $f(\ell) = \ell - 1$. Lines stop in case no ℓ -supporting group of this size exists.

the maximum ballot length is 24, so each voter can approve up to 24 candidates (and not only up to 16 as in Polkadot). The committee size in Kusama is $k = 1000$, more than three times as large as in Polkadot.

Description of the Data Analogous to Figure 1, Figure 11 shows the size of the Kusama elections. Compared to the Polkadot elections, the Kusama elections have more candidates and fewer voters. Moreover, as for the Polkadot data, we see a steady increase in the number of voters over time and a smaller fluctuation in the number of candidates.

We also repeated our core experiments presented in the main body on the Kusama elections. However, due to the larger size and increased computation time for the Kusama elections (which have more candidates and most critically a much higher k), we only ran the experiments on a subset of 20 elections, which are uniformly distributed over the available set of elections.

Overlap between Committees Table 7 depicts the average overlap of the committees computed by the different voting rules. As for Polkadot, the outcomes returned by the rules are typically very similar to each other (the difference is typically less than 10%). However, there are small differences in how the rules relate to each other here. In particular, the proportional rules show a slightly more diverse behavior and the separation between AV and SAV is less pronounced. In particular, while MES and seq-Phragmén still produce very similar committees, outcomes returned by seq-PAV are clearly closest to the ones returned by SAV. This different behavior of seq-PAV on the Kusama data is also reflected in some of our further experiments. Moreover,

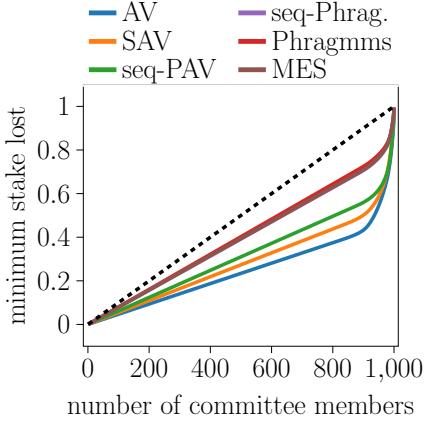


Figure 13: Minimum stake lost in case ℓ committee members misbehave.

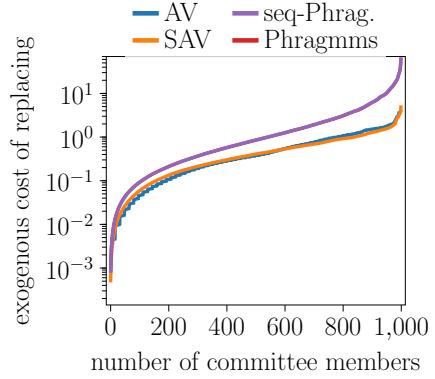


Figure 14: Stake needed to replace committee members; y-axis is logarithmic.

for Kusama, Phragmms is at a roughly similar distance from all other rules (including AV). Especially the similarity between Phragmms and AV is surprising here.

Preventing Underrepresentation Analogous to Tables 3 and 4, Table 8 shows some measures regarding the prevention of underrepresentation. Moreover, as in Figure 5, Figure 12 shows the minimum average satisfaction of ℓ -supporting groups in our elections.

The general trends are similar as for Polkadot, which is why we only point to the differences here, which mostly affect seq-PAV. Regarding the satisfaction of ℓ -supporting groups, as for Polkadot, the proportional rules substantially outperform the best possible guarantee of $\ell - 1$. The only differences are that the gap between seq-PAV and the other proportional rules is more pronounced and that SAV shows a better performance here.

Concerning priceability, seq-PAV shows a worse performance on the Kusama than on the Polkadot elections.

Lastly, for both AV and SAV, the average number of candidates violating EJR+ is more than twice as high as for JR. This indicates that on the Kusama elections, there is a practical difference between these two notions of proportionality.

Preventing Overrepresentation Analogous to Table 3, Table 9 shows some measures related to the prevention of overrepresentation. The trends in the results are again similar as for Polkadot. The biggest difference is that seq-PAV performs worse with respect to the maximin support value and produces values closer to those of AV than those of Phragmms. Another (small) difference is that AV performs slightly better with respect to the exogenous cost of replacing $\frac{k}{3}$ and $\frac{k}{2}$ committee members (see also Figure 14 for the minimum exogenous cost of replacing committee members).

As Figure 6b, Figure 13 depicts the minimum stake lost in Kusama, i.e., the minimum amount of stake assigned to a group of ℓ committee members in the maximin assignment. The results are very similar as for Polkadot, with the only difference being that seq-PAV performs slightly worse and is now more similar to AV and SAV than to the other proportional rules.