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

Niclas Boehmer[1], **Markus Brill**[2], **Alfonso Cevallos**[3], **Jonas Gehrlein**[3], **Luis Sánchez-Fernández**[4], **Ulrike Schmidt-Kraepelin**[5]

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

nboehmer@g.harvard.edu, markus.brill@warwick.ac.uk, alfonso@web3.foundation, gehrlein.jonas@protonmail.com, luiss@it.uc3m.es, u.schmidt.kraepelin@tue.nl

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

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 in the form of 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, Sreenivas, and Larson 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 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 in practice.

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, Laslier, and Skowron 2018) to recommender systems (Streviniotis and Chalkiadakis 2022a,b; 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 (Burdges et al. 2020; Cevallos and Stewart 2021).[1]

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

- 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 empirical in-depth study of different voting rules, with a particular focus on their application in Polkadot. We analyze the similarities between rules and whether their outcomes under- or overrepresent voters.

[1]Polkadot currently uses Phragmén's sequential rule but there are considerations to switch to the Phragmms rule (Cevallos and Stewart 2021).

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.

A full version of the paper is available on arxiv.org (Boehmer et al. 2023). 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 our full version, 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.

Related Work. There is a large literature developing (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). 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 proportionality axioms (Faliszewski et al. 2023b; Mehra, Sreenivas, and Larson 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) provided tools for more empirical studies of ABC rules by proposing several synthetic models for generating data and 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.

There are also some examples beyond Polkadot and Kusama using voting in blockchains. However, most of them use non-proportional voting rules. For instance, the EOS network uses *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.

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* $\alpha : V \times W \to [0, 1]$. 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 on 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. We introduce six rules that we study in this paper; all of them were originally proposed for unweighted elections, but their adaptation to weighted votes is straightforward. 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**) selects k candidates with highest approval weight. Satisfaction Approval Voting (**SAV**) assigns a score of $\sum_{v \in V_c} \frac{w(v)}{|A_v|}$ to each candidate $c \in C$ and selects k candidates with the highest score. The other four rules all aim to provide some form of proportional representation, so we refer to them as *pro-*

portional rules. The rules sequential PAV (**seq-PAV**), sequential Phragmén (**seq-Phragmén**), and Method of Equal Shares (**MES**) with seq-Phragmén completion are standard; see our full version (Boehmer et al. 2023) for their definitions. Seq-PAV is a greedy heuristic for optimizing the *PAV score* $\sum_{v \in V} w(v) \cdot (\sum_{i=1}^{|A_v \cap W|} \frac{1}{i})$ of a committee W. Our final proportional rule is less common: The **Phragmms** rule was introduced by Cevallos and Stewart (2021) for potential use in Polkadot. It combines aspects of seq-Phragmén and the maximin support method (Sánchez-Fernández et al. 2022); we do not include the maximin support method in our analysis due to its prohibitive computational complexity. The goal of Phragmms is to select a committee W together with a vote assignment α that guarantees a high backing weight for all candidates in W. The rule repeats the following 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. (ii) Add a candidate with highest score to W. (iii) Compute a new vote assignment for W that is "balanced". For details, we refer to the paper by Cevallos and Stewart (2021).

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. This is different from Proof-of-Work (PoW) blockchains such as Bitcoin (Nakamoto 2008), where everyone can propose new blocks. Thereby, PoS avoids the reliance on energy-intensive computing power that characterizes PoW systems, thus making it an environmentally friendly alternative. 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 the validators behave maliciously (Lamport, Shostak, and Pease 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, 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 are elected and diligently perform their duties. The 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.

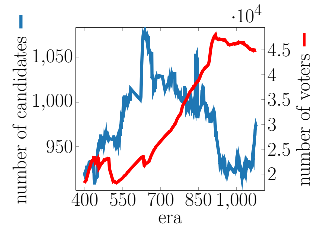


Figure 1: Number of candidates (blue) and voters (red) in our 496 elections.

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 voting rule 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.

Description of the Data. In Polkadot, every day is viewed as an *era*, and one election per era is conducted, implying that our 496 collected elections have a time-based ordering.[2] 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 $v \in V$.[3] 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 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 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

[2]This 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).

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

	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.

"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. Our full version (Boehmer et al. 2023) contains additional analyses of the data. For instance, we observe that elections typically only change little from one era to the next, yet the first election in the dataset is indeed different from the last.[4] Moreover, we analyze the structural properties of our elections. E.g., we find that the weight of voters is distributed unevenly: On average, the richest 31 voters combined have more weight than all other voters together.

Overlap Between Committees. In Table 1, we show the similarity between 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. 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 our full version (Boehmer et al. 2023), 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 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.

[4] For the sake of readability, in our following analysis, we mostly report average values, as the variance is usually quite low.

	PAV score	JR violations	EJR+ violations
AV	2.502	28.347	32.343
SAV	2.547	25.143	32.321
seq-PAV	2.585	0	0
Phragmms	2.578	0 (Guarantee)	0
seq-Phrag.	2.578	0 (Guarantee)	0
MES	2.579	0 (Guarantee)	0 (Guarantee)

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

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.

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).[5]

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+ in 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

[5] EJR+ 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).

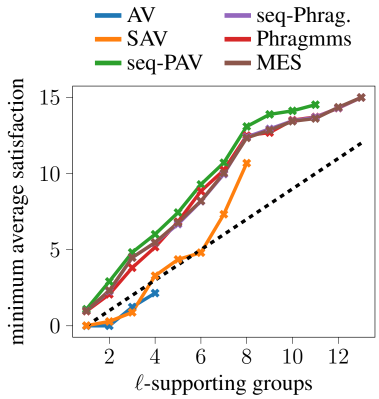


Figure 2: Minimum average satisfaction of ℓ-supporting groups. The dashed line is the function $f(\ell) = \ell - 1$. Lines stop in case no ℓ-supporting group of this size exists.

similar to each other, it follows that if a candidate violates EJR+, then they often violate JR as well.

EJR+ and JR only check for ℓ-supporting groups in which *all* voters are unsatisfied. Weakening this approach, one can also search for ℓ-supporting groups with low *average* satisfaction. This view is reflected in the notion of representativeness introduced by Brill and Peters (2023), which is related to the proportionality degree (Skowron 2021). 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 2, 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 $\ell \geq 4$.

In our full version (Boehmer et al. 2023) 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. It turns out that SAV and AV return committees far away from being priceable, highlighting that some voters had a much larger influence on the outcome. For seq-PAV, the picture is mixed, whereas all other rules are known to satisfy priceability.

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).[6] In the following, we explore three perspectives on overrepresentation.

	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.

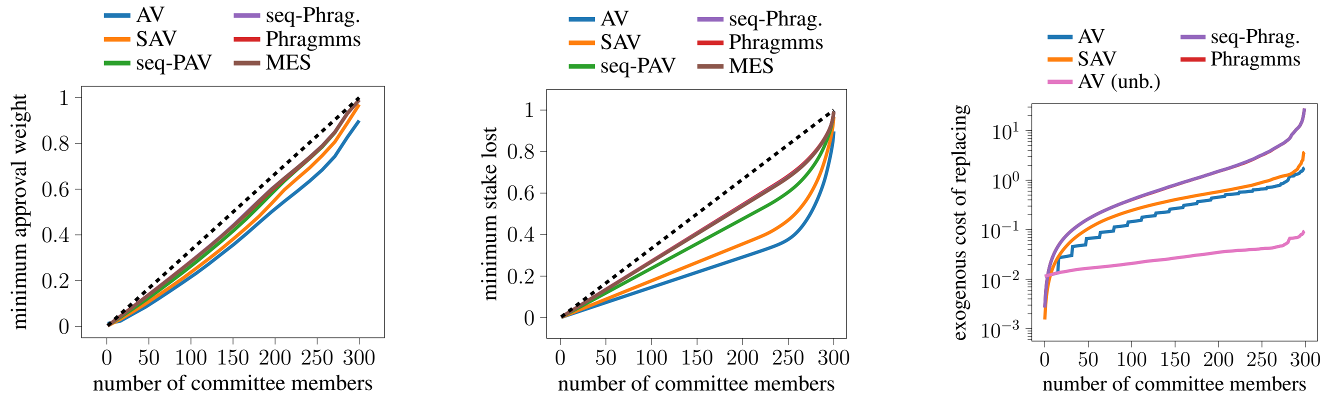
The first measure we consider was introduced by Cevallos and Stewart (2021). In order to make it as difficult as possible for an attacker to get ℓ committee members selected, they propose to maximize $\min_{W' \subseteq W : |W'| = \ell} w(W')$, i.e., the minimum approval weight of a group of ℓ committee members.[7] Figure 3a depicts these values for a varying value of ℓ.[8] 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).

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 2 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

[6] The critical threshold of elected malicious candidates depends on the type of consensus and is $\frac{k}{3}$ in Byzantine fault-tolerant consensus (Pease, Shostak, and Lamport 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).

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

[8] Computing these values is NP-hard, which is why we resorted to an ILP (see our full version (Boehmer et al. 2023) for details). As solving a single instance of the ILP took sometimes more than one day, in Figure 3a we only averaged over 15 instances, uniformly spaced over our dataset, and considered values of ℓ from $\{1, 15, 30, 45, \ldots, 300\}$.



(a) Minimum approval weight of a group of committee members.

(b) Minimum stake lost in case ℓ committee members misbehave.

(c) Stake needed to replace committee members; y-axis is logarithmic.

Figure 3: Different figures for metrics to prevent overrepresentation. The dashed line is the function $\frac{1}{300}\ell$.

four proportional rules, the MMS value is very close to the minimum approval weight of a selected candidate (also in Table 3), constituting 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 in Polkadot 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.

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 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 2. *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 (Burdges 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 3b 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 3b is closely connected to Figure 3a; the main difference is that in Figure 3b 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 3a we consider the full stake of all voters approving at least one candidate from the group. Accordingly, in Figure 3b, 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 3b, 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.

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.*

For details, we refer to our full version, where we also discuss the complexity of this problem for seq-PAV and MES. Figure 3c 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 3c). 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 our full version.

Choosing the Committee Size. Polkadot's committee size is determined by the network's governance body and is thus an adjustable design choice. This raises the question: Based on the previously formulated desiderata, would it be beneficial to increase or decrease the committee size?

We focus on $k = 200$ and $k = 400$ as two alternative committee sizes. 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.

Selecting Candidates Multiple Times. Brill et al. (2023) 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.[9]

Allowing for 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. Remarkably, this decision 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. Generally speaking, we found that the "proportional" rules (seq-PAV, seq-Phragmén, MES, and Phragmms) behave very similarly and that they outperform their axiomatic guarantees.[10] 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.

For future work, it would be interesting to analyze the connection between over- and underrepresentation in more detail, both from a theoretical and an empirical perspective. Moreover, this paper aims to provide motivation and new perspectives for further empirical work. We hope that others will (i) use the real-world election data presented here for more empirical studies and (ii) come up with additional quantitative measures to better understand and quantify the differences between voting rules in practice.

[9] While this is currently not possible in the Polkadot network, it is very easy for candidates to create copies of themselves. Indeed, entities that own a large amount of stake create multiple (up to 100) candidates in the system and label them as copies. Thus, *de facto* the same entity can already control multiple candidates.

[10] The fact that seq-PAV satisfies EJR+ on all our elections might suggest that voters' preferences in our elections are simple. However, in our full version (Boehmer et al. 2023), we argue that our elections are far away from falling into the common simple class of "party-list" elections and that our evidence rather suggests that rules like seq-PAV only violate proportionality axioms on adversarially constructed elections unlikely to occur in the real world.

Ethics Statement

This paper acknowledges the significant negative environmental impact associated with blockchain technology. However, the Polkadot network employs a Proof-of-Stake (PoS) consensus mechanism that does not involve the problematic 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. In fact, Polkadot has been acknowledged as one of the PoS networks 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 the consensus is limited in Polkadot, a restriction enabled by the election process that we study in the paper.

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 Matemtico (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).

References

Aziz, H.; Brill, M.; Conitzer, V.; Elkind, E.; Freeman, R.; and Walsh, T. 2017. Justified Representation in Approval-Based Committee Voting. *Social Choice and Welfare*, 48(2): 461–485.

Aziz, H.; and Lee, B. E. 2021. Proportionally Representative Participatory Budgeting with Ordinal Preferences. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*, 5110–5118. AAAI Press.

Boehmer, N.; Bredereck, R.; Faliszewski, P.; Niedermeier, R.; and Szufa, S. 2021. Putting a Compass on the Map of Elections. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, 59–65.

Boehmer, N.; M. Brill, A. C.; Gehrlein, J.; Sánchez Fernández, L.; and Schmidt-Kraepelin, U. 2023. Approval-Based Committee Voting in Practice: A Case Study of (Over-)Representation in the Polkadot Blockchain. Technical report, arXiv:2312.11408 [cs.GT].

Boehmer, N.; and Niedermeier, R. 2021. 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)*, 1–5. ACM.

Boehmer, N.; and Schaar, N. 2023. Collecting, Classifying, Analyzing, and Using Real-World Ranking Data. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 1706–1715. ACM.

Bredereck, R.; Faliszewski, P.; Niedermeier, R.; and Kaczmarczyk, A. 2019. An Experimental View on Committees Providing Justified Representation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 109–115.

Brill, M.; Gölz, P.; Peters, D.; Schmidt-Kraepelin, U.; and Wilker, K. 2023. Approval-Based Apportionment. *Mathematical Programming*.

Brill, M.; Laslier, J.-F.; and Skowron, P. 2018. Multiwinner Approval Rules as Apportionment Methods. *Journal of Theoretical Politics*, 30(3): 358–382.

Brill, M.; and Peters, J. 2023. Robust and Verifiable Proportionality Axioms for Multiwinner Voting. In *Proceedings of the 24th ACM Conference on Economics and Computation (ACM-EC)*, 301. ACM. Full version arXiv:2302.01989 [cs.GT].

Burdges, J.; Cevallos, A.; Czaban, P.; Habermeier, R.; Hosseini, S.; Lama, F.; Alper, H. K.; Luo, X.; Shirazi, F.; Stewart, A.; and Wood, G. 2020. Overview of Polkadot and its Design Considerations. Technical report, arXiv:2005.13456 [cs.CR].

Cevallos, A.; and Stewart, A. 2021. A verifiably secure and proportional committee election rule. In *Proceedings of the 3rd ACM Conference on Advances in Financial Technologies (AFT)*, 29–42. ACM.

Chong, N. 2019. Crypto Ratings Agency Downgrades EOS For "Serious Centralization Problems". https://www.newsbtc.com/news/eos/crypto-ratings-agency-downgrades-eos-serious-centralization-problems/. Accessed: 2024-01-22.

Crypto Carbon Ratings Institute. 2023. PoS Benchmark Study 2023. carbon-ratings.com/dl/pos-report-2023. Accessed: 2024-01-22.

Elkind, E.; Faliszewski, P.; Laslier, J.-F.; Skowron, P.; Slinko, A.; and Talmon, N. 2017. 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)*, 494–501. AAAI Press.

Faliszewski, P.; Flis, J.; Peters, D.; Pierczynski, G.; Skowron, P.; Stolicki, D.; Szufa, S.; and Talmon, N. 2023a. Participatory Budgeting: Data, Tools and Analysis. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI)*, 2667–2674. ijcai.org.

Faliszewski, P.; Lackner, M.; Sornat, K.; and Szufa, S. 2023b. An Experimental Comparison of Multiwinner Voting Rules on Approval Elections. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2675–2683.

Garg, P. 2019. EOS voting structure encourages centralization. https://cryptoslate.com/eos-voting-structure-encourages-centralization/. Accessed: 2024-01-22.

Gawron, G.; and Faliszewski, P. 2022. Using Multiwinner Voting to Search for Movies. In *Proceedings of the 19th European Conference on Multi-Agent Systems (EUMAS)*, 134–151. Springer.

Gehrlein, J.; Miebs, G.; Brunelli, M.; and Kadziński, M. 2023. An active preference learning approach to aid the selection of validators in blockchain environments. *Omega*, 118: 102869.

Grigg, I. 2017. EOS – An introduction. https://whitepaperdatabase.com/eos-whitepaper. Accessed: 2024-01-22. White Paper.

Lackner, M. 2020. Perpetual Voting: Fairness in Long-Term Decision Making. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 2103–2110. AAAI Press.

Lackner, M.; and Skowron, P. 2022. *Multi-Winner Voting with Approval Preferences*. Springer.

Lamport, L.; Shostak, R.; and Pease, M. 2019. The Byzantine generals problem. In *Concurrency: The Works of Leslie Lamport*, 203–226.

Mattei, N.; and Walsh, T. 2013. PrefLib: A Library for Preferences http://www.preflib.org. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*, 259–270. Springer.

Mattei, N.; and Walsh, T. 2017. A preflib.org Retrospective: Lessons Learned and New Directions. In Endriss, U., ed., *Trends in Computational Social Choice*, chapter 15, 289–305. AI Access.

Mehra, K.; Sreenivas, N. K.; and Larson, K. 2023. Deliberation and Voting in Approval-Based Multi-Winner Elections. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2853–2861.

Nakamoto, S. 2008. Bitcoin: A peer-to-peer electronic cash system.

Pease, M. C.; Shostak, R. E.; and Lamport, L. 1980. Reaching Agreement in the Presence of Faults. *Journal of the ACM*, 27(2): 228–234.

Peters, D.; and Skowron, P. 2020. Proportionality and the Limits of Welfarism. In *Proceedings of the 21st ACM Conference on Economics and Computation (ACM-EC)*, 793–794. ACM. Full version arXiv:1911.11747 [cs.GT].

Pommerening, A.; Brill, M.; Schmidt-Kraepelin, U.; and Haufe, J. 2020. Democratising forest management: Applying multiwinner approval voting to tree selection. *Forest Ecology and Management*, 478: 118509.

Reichert, J.; and Elkind, E. 2023. Agreement Among Proportional Approval-Based Multi-Winner Voting Rules: An Empirical Analysis. Presented at the 5th Games, Agents, and Incentives Workshop (GAIW).

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

Sánchez-Fernández, L.; Fernández, N.; Fisteus, J. A.; and Brill, M. 2022. The Maximin Support Method: An Extension of the D'Hondt Method to Approval-Based Multiwinner Elections. *Mathematical Programming*. Forthcoming.

Skowron, P. 2021. Proportionality Degree of Multiwinner Rules. In *Proceedings of the 22nd ACM Conference on Economics and Computation (ACM-EC)*, 820–840. ACM.

Skowron, P.; and Górecki, A. 2022. Proportional Public Decisions. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, 5191–5198. AAAI Press.

Stifter, N.; Judmayer, A.; Schindler, P.; Zamyatin, A.; and Weippl, E. 2018. Agreement with Satoshi On the Formalization of Nakamoto Consensus. Cryptology ePrint Archive, Paper 2018/400. https://eprint.iacr.org/2018/400. Accessed: 2024-01-22.

Streviniotis, E.; and Chalkiadakis, G. 2022a. Multiwinner Election Mechanisms for Diverse Personalized Bayesian Recommendations for the Tourism Domain. In *Proceedings of the Workshop on Recommenders in Tourism (RecTour)*, 65–82. CEUR-WS.org.

Streviniotis, E.; and Chalkiadakis, G. 2022b. 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)*, 331–346. Springer.

Szufa, S.; Faliszewski, P.; Janeczko, Ł.; Lackner, M.; Slinko, A.; Sornat, K.; and Talmon, N. 2022. How to Sample Approval Elections? In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, 496–502.

Szufa, S.; Faliszewski, P.; Skowron, P.; Slinko, A.; and Talmon, N. 2020. 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)*, 1341–1349. IFAAMAS.

Wood, G. 2016. Polkadot: Vision for a heterogeneous multichain framework. https://assets.polkadot.network/Polkadot-whitepaper.pdf. Accessed: 2024-01-22. White Paper.