**ABSTRACT**

We define a family of runoff rules that work as follows: voters cast approval ballots over candidates; two finalists are selected; and the winner is decided by majority. With approval-type ballots, there are various ways to select the finalists. We leverage known approval-based committee rules and study the obtained runoff rules from an axiomatic point of view. Then we analyze the outcome of these rules on single-peaked profiles, and on real data.

1 Introduction

Plurality with runoff (also known as runoff voting) is a widely used single-winner voting rule, in fact the most common rule for presidential elections throughout the world. But the social choice literature has pointed out that plurality with runoff suffers from so many drawbacks that we may wonder why it is used at all: it is highly sensitive to cloning, fails monotonicity, reinforcement, participation, Condorcet-consistency, and is very easy to manipulate. In particular, its high sensitivity to cloning has a number of derived effects before the vote (at the level of the determination of candidates) and at voting time (with massive strategic voting of a specific kind, named “useful voting”). Perhaps the main reason why it is so widely used after all is related to the fact that runoff voting is not used as a one-shot voting rule but as two-round protocol: voters are called to urns for the first round, the results are made public, and then some amount of time passes (typically one or two weeks). And in between the two rounds, many things happen.

In most variants, only two candidates are selected for the runoff. The others candidates may negotiate their support to one of the two contenders, leading to adjustments in the platforms proposed in the second round. The TV debates that take place between the two finalists at that point in time are considered as the most important moment in the whole campaign, and many voters may, during this period, review their decision to participate or not to the second vote. For all these reasons, the existence of two rounds of vote separate in time is considered to be crucial for the voters’ information.

Are the informational benefits of a runoff protocol enough to overcome its numerous theoretical drawbacks? There can be diverse opinions about this. However, instead of answering this question, we may ask another one: is it possible to keep the nice benefit of the two-round protocol without having to bear all the drawbacks of Plurality at the first round?

Clearly, if the answer to this question is positive, the format of the ballots at first round must no longer be uninominal. Several possibilities exist: ordinal ballots, cardinal ballots, or more simply, approval ballots. Approval ballots have several advantages; to start with, they are simple and easy to express.

In this paper we explore this possibility seriously. We define an approval-with-runoff election as a two-round protocol:

1. First round: voters cast approval ballots, from which the two finalists are selected.
2. Second round: voters cast votes for one of the two finalists, and the majority winner wins the overall election.

Formally, we define approval-with-runoff as a voting rule, with a one-shot input, and study its properties in a similar way as we would study the properties of plurality with runoff. Then two major questions arise:

1. What should the input of the rule consist of?

---

1See [https://en.wikipedia.org/wiki/Two-round_system](https://en.wikipedia.org/wiki/Two-round_system).

2The present paper is concerned with single-winner elections. Approval voting with a runoff is effectively used in several cantons in Switzerland for multi-winner elections. The precise rules vary from one canton to the other so that the second round is sometimes almost unused, as in the canton of Zurich ([Laslier and Van der Straeten](https://doi.org/10.1016/j.orsyb.2016.08.003), [Van der Straeten et al.](https://doi.org/10.1016/j.orsyb.2018.03.003)).
2. Which rule should be used to determine the two finalists?

For question 1, the answer becomes clear once we remark that we need the approval data for computing the finalists, and the pairwise comparisons between candidates for computing the final winner. Of course, we will not need all comparisons between arbitrary pairs of candidates; but just as plurality with runoff, seen as a voting rule, takes full rankings as input although most of this information will not be asked, here too, we need more information in the input than we will ask voters, and the normative properties of the rule will be evaluated with respect to this (mostly private) information. Now, requiring pairwise comparisons between all pairs of candidates just means that we need each voter’s ranking of candidates, and requiring her approval set means that this ranking comes with a threshold that separates approved candidates from disapproved candidates. This data structure is called an approval-preference (AP) profile [Brams and Sanver 2009].

Notice that, with respect to the points mentioned in the introduction, the framework that we use does not allow taking into account the evolution of voters and candidates in between the two rounds. We leave these problems to further research and, in this paper as it is often the case in social choice theory, we concentrate on the counting of sincere ballots cast by a fixed electorate.

For question 2, things are more complex because there is not a unique way to select two candidates from approval ballots. The general setting in which we select \( k \) candidates (here, \( k = 2 \)) from an approval profile is called an approval-based committee rule (ABC rule). A recent and extensive survey is on ABC rules is Lackner and Skowron [2020], and we have now series of results that tell us which properties these various rules satisfy and for which contexts they are suitable. Most importantly, the choice of the rule used for the first round has strong implications about the very nature of the two-stage rule, both from a normative point of view and from a political science point of view: should we send to the second round the most two approved candidates? Or should we offer the voters two candidates that are diverse enough? Should we pay attention to proportionality issues? Should we guarantee the most approved candidate is among the two finalists?

Our primary aim is to define approval with runoff not just as one rule but as a family of rules, and to explore the reasons that may guide us towards the choice of one of the rules in the family.

The paper is organized as follows. We start by related work (Section 2). We define and study the family of Approval-based committee rules (Section 3) together with a selection of meaningful rules. Then, we define Approval with runoff rules and study these rules form an axiomatic point of view (Section 4). We analyse the outcome of these rules on one-dimensional Euclidean profiles (Section 5), and move on to applying the rules on real data (Section 6). We conclude in Section 7.

## 2 Related work

Approval with Runoff was first introduced in Sanver [2010] and compared to other rules based on approval-preference profiles. Green-Armytage and Tideman [2020] consider plurality with runoff together with eight other runoff rules for selecting the finalists, with varying input formats (ordinal, approval, numerical), including Approval with Runoff. Voters are supposed to vote sincerely and, for Approval voting, to approve a candidate if and only if the utility they give to this candidate is larger than the average utility of all candidates running. They evaluate these rules along four numerical criteria (expected utility of winner, of the runoff loser, representativeness, resistance to strategy); numerical results come both from using real data and from simulations. Among other conclusions, plurality with runoff scores particularly bad, and approval with runoff, slightly better, although it is beaten by plurality with runoff on two criteria: representativeness and resistance to strategy.

A runoff can also be seen as an extreme case of shortlisting (with at most two selected candidates). Using approval for shortlisting candidates was studied recently in Lackner and Maly [2021]; a crucial difference with runoff rules is that shortlisting does not impose constraints on the number of selected candidates, which leads to very different rules (such as dichotomy rules, or rules based on large gaps).

Approval-based multiwinner election rules have received enormous attention these last ten years: see Lackner and Skowron [2020] for a review. They are clustered in several groups according to the objective of the selection: excellence (select the individually best \( k \) candidates), proportional representation (ensure that each coherent group of voters is represented in the selection, proportionally to its size), or diversity (output a diverse set of candidates, avoid similar candidates in the selection). Which of these three clusters of rules suits the selection of runoff candidates better is not clear at this point; our paper aims at answering (at least partly) this question.

Defining voting rules that take as input approval-preference profiles has been initiated in Brams and Sanver [2009], who propose and study two such rules (preference approval voting and fallback voting) that have been studied in a
number of subsequent works, from the point of view of axiomatization, computation, resistance to strategic behaviour; as far as we know, they have not been studied in the context of runoffs.

3 Approval-based committee rules

Let $C = \{c_1, \ldots, c_m\}$ be a set of $m$ candidates and $V = \{v_1, \ldots, v_n\}$ a set of $n$ voters. An approval profile is a collection of approval ballots $V = \{A_1, \ldots, A_n\}$ with $A_i \subseteq C$ for all $i$. An Approval-based committee rule (ABC rule) $F$ is a rule that takes as input an approval profile $V$ and a committee size $k \geq 1$ and return a set of winning committees $W \subseteq C$ of size $|W| = k$.

For an approval profile $V$ we denote $S_V(c) = |\{i | c \in A_i\}|$ the approval score of a candidate $c \in C$. By extension, the approval score of a set of candidates $J \subseteq C$ is the number of approval ballots that contains all candidates from the set $J$, $S_V(J) = |\{i | J \subseteq A_i\}|$. For simplicity, we write sets on a simpler form, e.g. $S_V(abc)$ instead of $S_V(\{a, b, c\})$. We call approval winners the candidates that maximize $S_V$, that is the winners of standard (single-winner) approval voting.

3.1 Rules

As we explained in the introduction, an approval with runoff (AVR) rule use an ABC rule to select two finalists and then return the majority winner between the two finalists. [Lackner and Skowron, 2020] did an extensive study of these ABC rules. Because it is the case that interests us, in this section we define rules in the case of a committee size of $k = 2$. We will see that most of these rules can be defined with a simple formula for this particular case.

The most intuitive rule is probably the one that selects the candidates with the highest approval scores:

Multi-Winner Approval Voting (MAV):

$$MAV(V) = \arg \max_{x_1, x_2 \in C} S_V(x_1) + S_V(x_2)$$

Some rules discount the satisfaction of voters who are already satisfied by one of the two finalists. This is the case of Proportional Approval Voting (PAV) and Approval Chamberlin Courant (CAV). In PAV, a voter approving $j$ candidates of the committee $W$ gives a score $s = \sum_{i=1}^{j} \frac{1}{2}$ to the committee. For a committee of size $k = 2$, this means that a voter approving one candidate give a score of 1, and a voter approving both candidates a score of $\frac{3}{2}$. If we do $S_V(x_1) + S_V(x_2)$ as in MAV, voters approving both candidate give a score of 2 instead of $\frac{3}{2}$, so we have to discard a score of $\frac{1}{2}$ for each of these voters. In CAV, a voter approving one or more candidates of the committee gives a score of 1 to the committee. Therefore, in comparison to MAV, we have to discard 1 on the score given by voters approving both candidates. This is why in the case $k = 2$, we can write these rules like this:

Proportional Approval Voting (PAV):

$$PAV(V) = \arg \max_{x_1, x_2 \in C} S_V(x_1) + S_V(x_2) - \frac{1}{2} S_V(x_1 x_2)$$

Approval Chamberlin Courant (CAV):

$$CAV(V) = \arg \max_{x_1, x_2 \in C} S_V(x_1) + S_V(x_2) - S_V(x_1 x_2)$$

These rules select the pairs of candidates $\{x_1, x_2\}$ maximizing $S_V(x_1) + S_V(x_2) - \alpha S_V(x_1 x_2)$ for some $\alpha \in [0, 1]$. This $\alpha$ is equal to 0 for MAV, to 1/2 for PAV and to 1 for CCAV. We call these rules $\alpha$-AV rules. There also exists sequential versions of these rules. In these sequential versions, the first finalist is always an approval winner.

Sequential Proportional Approval Voting (S-PAV): The rule chooses the pairs $\{x_1, x_2\}$ such that $x_1$ maximizes $S_V(x_1)$ and $x_2$ maximizes $S_V(x_2) - \frac{1}{2} S_V(x_1 x_2)$.

Sequential Approval Chamberlin Courant (S-CAV): The rule chooses the pairs $\{x_1, x_2\}$ such that $x_1$ maximizes $S_V(x_1)$, and $x_2$ maximizes $S_V(x_2) - S_V(x_1 x_2)$.

As before, these definitions of S-PAV and S-CAV for the case $k = 2$ are equivalent to the definitions in the general case. In S-PAV, every voter that approves the first finalist $x_1$ now has a weight of $\frac{1}{2}$, and in S-CAV, it has a weight of 0. Note that sequential MAV would be equivalent to standard MAV. For these sequential rules, the first finalist is an
We also need the rule that returns all pairs of candidates:

A rule that almost falls into this family is the Eneström Phragmen rule. This rule is also sequential, as the first finalist \( x_1 \) is an approval winner. Then, the weight of voters approving \( x_1 \) is reduced to \( \max(0, 1 - \frac{Q}{S_V(x_1)}) \) where \( Q \) corresponds to some quota \( Q \in [0, n] \). Most of the time, we use Droop quota \( Q = n/(k+1) \) which is \( n/3 \) in our case, or Hare quota \( Q = n/k \), equal to \( n/2 \) in our case. Thus, in the case \( k = 2 \), we can define this rule as follows:

**Eneström Phragmen (EnePhr):** The rule chooses the pairs \( \{x_1, x_2\} \) such that \( x_1 \) maximizes \( S_V(x_1) \) and \( x_2 \) maximizes \( S_V(x_2) - \alpha S_V(x_1) \) for some quota \( Q \in [0, n] \).

This gives us a \( \alpha \)-seqAV rule dependent on \( Q \), with \( \alpha = \min(1, \frac{Q}{S_V(x_1)}) \).

Another popular rule to obtain proportionality in approval based committee selection is the sequential Phragmen rule.

**Sequential Phragmen (S-Phr):** The rule chooses the pairs \( \{x_1, x_2\} \) such that \( x_1 \) maximizes \( S_V(x_1) \) and \( x_2 \) minimizes

\[
1 + \frac{S_V(x_1,x_2)}{S_V(x_1)} - \alpha \frac{S_V(x_1,x_2)}{S_V(x_1)} = \max\left(0, 1 - \frac{Q}{S_V(x_1)}\right)
\]

To see how we obtain this simple formulation, we have to use the discrete formulation of S-Phr presented in [Lackner and Skowron, 2020](#). The load of every voter is initialized at \( l_i = 0 \). The first finalist selected \( x_1 \) is an approval winner. Then every voter that approves \( x_1 \) get a load of \( l_i = \frac{1}{S_V(x_1)} \). The second finalist is the candidate \( x_2 \) minimizing:

\[
1 + \sum_{x_2 \in A_i} l_i \frac{1}{S_V(x_2)} = 1 + \sum_{x_1,x_2 \in A_i} \frac{1}{S_V(x_1)} \frac{1}{S_V(x_2)} = \frac{1 + \sum_{x_1 \in A_i} \frac{1}{S_V(x_1)}}{S_V(x_2)}
\]

In the next rule, every voter has a weight of 1 and splits its weight between the candidates he approves:

**Splitted Approval Voting (SAV):** The rule chooses the pairs \( \{x_1, x_2\} \) such that \( x_1 \) and \( x_2 \) maximize the splited approval score \( S_{pv}(x) \) with

\[
S_{pv}(x) = \sum_{i:x \in A_i} \frac{1}{|A_i|}
\]

That means that each voter gives the same fraction of vote to each candidate he supports.

**Example 1.** Let \( V = \{2 \times a, 6 \times ab, 4 \times abc, 4 \times cd, 1 \times d\} \), i.e., two ballots \( \{a\}, \{a, b\} \) etc. Table 1 summarizes the score of the three main committees \( \{a, b\}, \{a, c\} \) and \( \{a, d\} \) for the different voting rules. For sequential rule S-PAV and S-CCAV they are here equivalent to their non-sequential versions, because the approval winner is \( a \). For Eneström Phragmen rule, with the Droop quota \( Q = n/(k+1) = 17/3 \), we have \( \alpha = \frac{Q}{S_V(a)} = \frac{17}{3} = 17/36 \) which is very close to 1/2, thus the results will be similar to S-PAV. The interesting case is with Hare quota \( Q = n/2 = 17/2 \), we have \( \alpha = \frac{Q}{S_V(a)} = \frac{17}{2} = 17/24 \). This gives the score of \( 12 + 8 - 4 \frac{17}{24} = 17.2 \) for \( \{a, c\} \) and \( 12 + 5 - 0 = 17 \) for \( \{a, d\} \). Thus, \( \{a, c\} \) is also the pair of finalists with Hare quota.

We also need the rule that returns all pairs of candidates:

**Trivial Approval Voting (TRIV):**

\[
TRIV(V) = \{\{x, x'\} | x, x' \in C, x \neq x'\}
\]

Note that the trivial approval rule with runoff is actually not completely trivial: it outputs all candidates except the Condorcet loser whenever there is one.

| Rule      | MAV       | (S-)PAV        | (S-)CCAV     | S-Phr               | SAV   |
|-----------|-----------|----------------|--------------|---------------------|-------|
| \{a, b\} | 12 + 10 = 22 | 12 + 10 − \frac{17}{2} = 17    | 12 + 10 − 10 = 10 | \frac{1+10/12}{2} = \frac{22}{120} | 2 + 6 + 4.7 = \textbf{10.7} |
| \{a, c\} | 12 + 8 = 20  | 12 + 8 − \frac{12}{2} = 18     | 12 + 8 − 4 = 16   | \frac{1+4/12}{2} = \textbf{16/96} | 2 + 6.2 + 4.5 + 4.2 = 9.7 |
| \{a, d\} | 12 + 5 = 17  | 12 + 5 − \frac{7}{2} = 17      | 12 + 5 − 0 = 17    | \frac{1.5/12}{2} = \frac{1}{5}   | 2 + 10.5 + 4.2 + 1 = 10   |

Table 1: Score of the different committees in Example 1
3.2 Axioms

[2020] Lackner and Skowron already did an extensive study of these approval-based committee rules. We only complete it here by adding one property that is very interesting for us: favorite-consistency. It says that at least one of the candidates in the committee should be an approval winner. This property is important because it is hard for voters to accept a voting rule in which the candidate with the highest amount of approval is not sure to go to the second round. This property and all the properties of this section are defined for every committee size $k \geq 1$.

Definition 1. An ABC rule $F$ is said to be favorite-consistent if every winning committee contains an approval winner, i.e. for all $W \in F(V)$, we have $W \cap \text{arg max}_{c \in C} S'_V(c) \neq \emptyset$.

Among the voting rules considered in Section 3.1, only sequential rules satisfy this property.

4 Approval with Runoff

4.1 The model

We already define approval profiles as collections of approval ballots. An ordinal preference profile is a collection of rankings $\succ_i = (\succ_{i1}, \ldots, \succ_{in})$, where $\succ_i$ is the preference ranking of voter $i$ over $C$. An approval-preference profile is a collection of pairs of rankings $P = \langle (A_1, \succ_{11}), \ldots, (A_n, \succ_{nn}) \rangle$ where $\succ_P = \langle A_1, \ldots, A_n \rangle$ is an approval profile and $\succ = (\succ_{11}, \ldots, \succ_{nn})$ an approval preference profile. We also note $P = (V_P, \succ)$.

In this section, we assume ballot consistency: voter $v_i$ has a threshold in her ranking $\succ_i$ such that every candidate above the threshold is approved and every candidate below is not; formally, $a \succ_i b$ holds for all $a \in A_i$ and $b \not\in A_i$. Ballot consistency allows us to use the following notation [Brams and Sanver, 2009]: $x_1 x_2 \ldots x_j | x_j+1 \ldots x_m$ represents $(\succ_i, A_i)$ with $x_1 \succ_i x_2 \succ_i \ldots \succ_i x_m$ and $A_i = \{ x_1, \ldots, x_j \}$.

Given an ordinal preference profile $\succ$, $\text{maj}(\succ, \{a, b\})$ is defined as the set of winners of the majority vote between $a$ and $b$ (which is a singleton except in the case of a tie).

We now define the family of approval with runoff (AVR) rules. The idea is that we use the approval ballots in the first round to select two finalists, and the second round consists in a majority vote between the two selected candidates. Let $F$ be an (irresolute) approval-based 2-committee rule, i.e. a function that takes as input an approval profile and returns a nonempty set of pairs of candidates. Then, $F^R$ is the (irresolute) AVR rule such that we conduct the majority rule on every pair of finalists selected by $F$. Formally:

$$ F^R(V, \succ) = \bigcup_{\{x, y\} \in F(V)} \text{maj}(\succ, \{x, y\}) $$

Example 2. We continue Example 1 where $V = (2 \times a, 6 \times ab, 4 \times abc, 4 \times cd, 1 \times d)$. We can define the approval-preference profile $P = (V, \succ) = (2 \times a|bc, 3 \times ba|dc, 3 \times ab|dc, 4 \times bac|d, 2 \times cd|ba|2 \times dc|ba, 1 \times d|bac)$, i.e. 2 voters approving $\{a\}$ with ranking $a \succ b \succ c \succ d$, 3 voters approving $\{b, a\}$ with ranking $b \succ a \succ d \succ c$, etc. In this profile $\text{maj}(\succ, \{a, b\}) = b$, $\text{maj}(\succ, \{a, c\}) = a$ and $\text{maj}(\succ, \{a, d\}) = a$. Thus, for $F$ equal to MAV and SAV, $F^R(P) = \text{maj}(\succ, \{a, b\}) = b$, for $F$ equal to (S-)IPV, S-Phr and EnePhr, $F^R(P) = \text{maj}(\succ, \{a, c\}) = a$ and for $F$ equal to (S-)CCAV, $F^R(P) = \text{maj}(\succ, \{a, d\}) = a$.

4.2 Axiomatic analysis

In this section, we study the axiomatic properties of AVR rules.

As for ABC rules, an AVR rule $F^R$ is anonymous if it is invariant by any permutation of the voters, and neutral if for any permutation of the candidates $\pi$ and every profile $P$, $F^R(\pi(P)) = \pi(F^R(P))$.

We will use the following unanimity condition, that is a strengthening of strict Pareto, adapted to the approval-preference case. We say that candidate $a$ unanimously preference-approval dominates candidate $b$ if

1. for every voter $v_i$, $a \succ_i b$
2. for some voter $v_i$, $a \in A_i$ and $b \not\in A_i$

3Ballot consistency does not necessarily hold if voters are strategic and cast insincere approval ballots. Most results in the paper still hold without assuming ballot consistency.
Together with ballot consistency, it implies that every voter who approves $b$ also approve $a$, and at least one voter who approves $a$ does not approve $b$. For simplicity we refer to this condition as our Pareto condition, and say that $a$ dominates $b$ when $a$ unanimously preference-approval dominates $b$.

**Definition 2.** An AVR rule $F^R$ is *Pareto-efficient* if for all approval-preference profile $P$ in which there exists $a, b \in C$ such that $a$ dominates $b$, we have $b \notin F^R(P)$.

**Proposition 1.** $AV^R$, $S$-PAV$^R$, $S$-Phr$^R$, EnePhr$^R$, PAV$^R$ and SAV$^R$ are Pareto-efficiency, but not $S$-CCAV$^R$, CCAV$^R$ and TRIV$^R$.

**Proof.** Consider a profile $P$ in which a candidate $a$ dominates another candidate $b$. Let $F$ be an approval based committee rule and $F^R$ the approval with runoff rule associated to it. We want to show that $b \notin F^R(P)$. We assume by contradiction that $b \in F^R(P)$. Thus, there exists $x \in C$ such that $\{b, x\} \in F(V)$.

$x \neq a$, because we know that $maj(\{a, b\}) = a$ since $a$ dominates $b$. We are going to show that for every rule $F \in \{MAV, S$-PAV, $S$-Phr, PAV, SAV$\}$, we have a contradiction.

- Let $F$ be an $\alpha$-AV rule with $\alpha < 1$ (like MAV and PAV). Define $f(y, y')$ the $\alpha$-AV score of a pair of candidates $\{y, y'\}$:

\[
f(y, y') = S_V(y) + S_V(y') - \alpha S_V(yy')
= S_V(y) + (1 - \alpha)S_V(y') + \alpha(S_V(y') - S_V(yy'))
\]

Since $a$ dominates $b$, we have $S_V(a) > S_V(b)$ and for all $x$, $S_V(a) - S_V(ax) \geq S_V(b) - S_V(bx)$. We have $f(x, a) = S_V(a) + (1 - \alpha)S_V(x) + \alpha(S_V(a) - S_V(ax)) > S_V(b) + (1 - \alpha)S_V(x) + \alpha(S_V(b) - S_V(bx)) = f(x, b)$, thus $\{x, b\}$ cannot be a pair of finalists because the rule selects the pairs of candidates maximizing $f$.

- Le $F$ be an $\alpha$-seqAV rule with $\alpha < 1$ (like MAV and S-PAV). $b$ is not the first finalist selected because $S_V(a) > S_V(b)$. Let $x$ be the first finalist selected and

\[
f_x(y) = S_V(y) - \alpha S_V(xy)
= (1 - \alpha)S_V(y) + \alpha(S_V(y) - S_V(xy))
\]

Again, $a$ dominates $b$, so we have $S_V(a) > S_V(b)$ and for all $x$, $S_V(a) - S_V(ax) \geq S_V(b) - S_V(bx)$. We have $f_x(a) = (1 - \alpha)S_V(a) + \alpha(S_V(x) - S_V(ax)) > (1 - \alpha)S_V(b) + \alpha(S_V(x) - S_V(bx)) = f_x(b)$, thus $\{x, b\}$ cannot be a pair of finalists because the rule selects the candidates maximizing $f_x$.

- Let $F = S$-Phr. Since $S_V(a) > S_V(b)$, $b$ is the second finalist selected and $x$ is the first finalist. For $y \in C$, let

\[
f_x(y) = \frac{1 + S_V(xy)}{S_V(x)}
= \frac{1}{S_V(x)} S_V(x) + S_V(xy)
\]

We want to prove that $\frac{S_V(x) + S_V(xa)}{S_V(a)} < \frac{S_V(x) + S_V(xb)}{S_V(b)}$. Now, observe that

\[
\frac{S_V(x) + S_V(xb)}{S_V(b)} \geq \frac{S_V(x)}{S_V(b)} > 1 \geq \frac{S_V(xa) - S_V(xb)}{S_V(a) - S_V(b)}
\]

which gives

\[
(S_V(x) + S_V(xb))(S_V(a) - S_V(b)) > (S_V(xa) - S_V(xb))S_V(b)
\]

$\iff (S_V(x) + S_V(xb))(S_V(a) + S_V(b))(S_V(x) - S_V(xb)) > (S_V(xa) + S_V(xb))(S_V(b) + S_V(b))(S_V(x) - S_V(xb))$

$\iff (S_V(x) + S_V(xb))(S_V(a)) > (S_V(xa) + S_V(xb))(S_V(b))$

$\iff S_V(x) + S_V(xb) > S_V(xa)
\]

Therefore, $f_x(a) < f_x(b)$, thus $\{b, x\}$ cannot be a pair of finalists because S-Phr selects the candidates minimizing $f_x$.

- If $F = SAV$, we define the Splitted Approval score $Sp_V(y) = \sum_{y \in A_i} \frac{1}{|A_i|}$. The two finalists are $x$ and $b$. However, $Sp_V(a) = Sp_V(b) + \sum_{i: a \in A_i \text{ and } b \notin A_i} \frac{1}{|A_i|} > Sp_V(b)$. This contradicts that $\{b, x\}$ is a possible pair of finalists.
In all those cases, we showed that \( \{b, x\} \notin F(V) \) for all \( x \). Therefore, \( b \notin F_R(P) \).

To show that the other rules are not Pareto-efficient, consider the profile \( P = (V, \succ) = (abc, ab|c, a|bc, k \times \{bca\}) \) with \( k > 3 \). For \( P \in \{\text{CCA V}, \text{S-CCA V}\} \), \( F(V) = \{\{a, b\}, \{a, c\}\} \) and \( \text{maj}(\succ, \{a, c\}) = c \), so \( c \in F_R(V) \). However, \( b \) dominates \( c \). Thus, \( F \) is not Pareto-efficient.

For EnePhr with quota \( Q > 0 \), observe that it is equivalent to S-CCA V if \( \frac{Q}{S_V(a)} \geq 1 \). Then, if \( Q = \beta n = \beta(k + 3) \) with \( \beta \in [0, 1] \), we set \( k \) in \( P \) such that \( \frac{Q}{S_V(a)} \geq 1 \), i.e. \( k \geq 3(\frac{1}{\beta} - 1) \).

A \( i \)-deviation of a profile \( P \) is a profile \( P' \) such that for all \( j \neq i \), \( A_j = A_j' \) and \( \succ_j = \succ_j' \). We define strong strategy-proofness as the impossibility for a voter to deviate from a profile to another in which there is a winner that she prefers to all the winners of the first profile. We also define weak strategy-proofness as the impossibility for a voter to deviate from a profile where she does not approve any winning candidate to one where she approves at least one winning candidate.

**Definition 3.** An approval with runoff rule \( F_R \) is strongly strategy-proof if for every profile \( P \), there is no \( i \)-deviation \( P' \) of \( P \) such that \( \exists x \in F(R(P')) \forall y \in F(R(P), x \succ_i y) \).

**Definition 4.** An AVR rule \( F_R \) is weakly strategy-proof if for every profile \( P \), there is no \( i \)-deviation \( P' \) of \( P \) such that \( |F^R(P') \cap A_i| = 0 \) and \( |F^R(P') \cap A_i| \geq 1 \).

It is hard to get strategy-proofness in approval-based committee voting, as it is incompatible with proportionality [Peters 2021]. For AVR rules, this is no better: weak strategy-proofness is incompatible with Pareto-efficiency. Among the rules defined in Section 3.1, the only strategy-proof rule is TRIV^R.

**Theorem 1.** No AVR rule is weakly strategy-proof and Pareto-efficient.

**Proof.** Assume \( F_R \) is Pareto-efficient and weakly strategy-proof. Let \( P = (V, \succ) = (a|bc, a|bc, c|ab, 10 \times abc) \). Then, \( V \) contains 2 approval ballots \( \{a\} \), 1 \( \{c\} \) and 10 \( \{a, b, c\} \), and \( \succ \) contains 12 rankings \( a \succ b \succ c \) and \( 1 \succ a \succ b \). In \( P \), \( a \) dominates \( b \) and \( \text{maj}(\succ, \{a, b\}) = b \), which implies \( \{a, b\} \notin F(V) \). Let \( P' = (V', \succ) = (a|cb, a|cb, c|ab, 10 \times cba) \). In \( P' \), \( c \) dominates \( b \) and \( \text{maj}(\succ, \{a, b\}) = b \), therefore \( \{a, b\} \notin F(V') \). Because \( V = V' \), \( F(V) = F(V') = \{a, c\} \).

Let \( P'' = (V'', \succ) = (a|bc, a|bc, c|ab, 10 \times abc) \). \( a \) dominates \( b \) in \( P'' \), and \( \text{maj}(\succ, \{a, b\}) = b \). By Pareto-efficiency, \( b \notin F_R(P'') \) and thus \( \{b, c\} \notin F(V'') \). Assume \( \{a, b\} \in F(V'') \). Let \( \mathcal{P} = (V, \succ') = (a|bc, a|bc, c|ab, 10 \times cab) \) and a deviation \( \mathcal{P}' = (V', \succ) = (a|bc, a|bc, c|ab, 10 \times cab) \), where voter \( v_1 \) deviates from \( \{a\} \) to \( \{a, b\} \). We know \( F(V) = \{a, c\} \) so \( F_R(\mathcal{P}) = \text{maj}(\succ', \{a, c\}) = c \), and \( a \in F^R(\mathcal{P}) \) because we assumed \( \{a, b\} \in F(V'') \). This is a successful manipulation, which contradicts strategy-proofness. Thus, \( \{a, b\} \notin F(V') \) and \( F(V'') = \{a, c\} \).

Let \( P^* = (V^*, \succ^*) = (a|bc, a|bc, c|ab, 10 \times cab) \). In \( P^* \), \( b \) dominates \( c \) and \( \text{maj}(\succ^*, \{a, c\}) = c \). Therefore, \( \{a, c\} \notin F(V^*) \). Let \( \mathcal{P} = (V, \succ) = (a|bc, a|bc, c|ab, 10 \times cab) \) and \( \mathcal{P}^* = (V^*, \succ) = (a|bc, a|bc, c|ab, 10 \times cab) \) a deviation of voter \( v_2 \). We have \( F(R(\mathcal{P})) = \text{maj}(\succ, \{a, c\}) = a \). However, if we assume \( \{b, c\} \in F(V^*) \), then \( c \in F^R(\mathcal{P}^*) \) because \( \text{maj}(\succ, \{a, c\}) = c \). Therefore, the deviating voter makes \( c \) win. This contradicts strategy-proofness, therefore, \( \{b, c\} \notin F^R(\mathcal{P}^*) \) and \( F(V^*) = \{a, b\} \).

Finally, let \( \mathcal{P}^* = (V^*, \succ) = (a|bc, a|bc, c|ab, 10 \times cab) \) and the deviation \( \mathcal{P} = (V, \succ) = (a|bc, a|bc, c|ab, 10 \times cab) \), where the voter \( v_3 \) deviates. We have \( F(R(\mathcal{P})) = \text{maj}(\succ, \{a, b\}) = a \) and \( F(R(\mathcal{P})) = \text{maj}(\succ, \{a, c\}) = c \). This deviation is a manipulation. This contradicts strategy-proofness, and proves the theorem.

This set of properties is minimal: TRIV^R is weakly strategy-proof but not Pareto-efficient, MAV^R is Pareto-efficient but not weakly strategy-proof.

**Proposition 2.** TRIV^R is strategy-proof but not AV^R, S-PAV^R, S-Phr^R, EnePhr^R, S-CCA V^R, PA^R, CCA^R and SAV^R.
Proof. We know by Theorem 1 that no rule is Pareto-efficient and Strategy-Proof. Thus, AVR, (S-)PAV, S-Phr, EnePhr, and SAV are not weakly strategy-proof. The trivial rule is clearly strategy-proof. For CCAV and S-CCA, consider the profile \( P = (V, \succ) = \{c[a]b, c[a], b, b(c[a], 10 \times abc)\} \). In this profile, all these two rules the final winner is \( c \). Let \( \alpha \) be an \( \alpha \)-improvement of \( P = (V, \succ) \), and \( a \succ b \). If the voter \( v_1 \) approving \( \{c, a\} \) now approves \( \{a\} \) only, we obtain the profile \( P' = (V, \succ') = (a[c]ba, c) \). All pairs of finalists are possible. Thus, \( a \) is a possible winner as \( \text{maj}(\succ', \{a, b\}) = a \). This is a manipulation. \( \square \)

We now focus on monotonicity. Given a profile \( P = (V, \succ) \) and \( a \in C \), a profile \( P' = (V', \succ') \neq P \) is an \( \alpha \)-improvement of \( P \) if for some \( i \in N \) we have

1. \( A'_i = A_i \cup \{a\} \) or \( A'_i = A_i \)
2. For all \( x, y \) with \( y \neq a \), if \( x \succ_i y \), then \( x \succ'_i y \)
3. For all \( j \neq i \), \( A'_j = A_j \) and \( \succ'_j = \succ_j \)

Definition 5. An AVR rule \( R \) is \( \alpha \)-improvement if for every \( a \in F^R(P) \) and for every \( \alpha \)-improvement \( P' \) of \( P \), we have \( a \in F^R(P') \).

Proposition 3. AVR and TRIVR is monotonic but not S-PAV, S-Phr, EnePhr, S-CCA, PAV, CCA, SAV are not.

Proof. Let's show that MAVR is monotonic. Let \( P = (V, \succ) \) be a profile in which \( a \in MAVR(P) \). Let \( x \in C \) such that \( \{a, x\} \in F(V) \). Let \( P' = (V', \succ') \) be an \( \alpha \)-improvement of \( P \). Since \( \succ'_i = \succ_i \) and for all \( y \neq a \), \( \succ'_i(y) = \succ_i(y) \), it is obvious that \( \{a, x\} \in F(V') \). Since \( a \in \text{maj}(\succ', \{a, x\}) \), it is also obvious that \( a \in \text{maj}(\succ', \{a, x\}) \) and thus \( a \in MAVR(P) \).

To show that the other rules are not monotonic, we can use a counterexample. Consider the profile \( P = (V, \succ) = (a[b]c, a[bc], b, ca, 10 \times cba) \). In this profile, for all rule \( F \in \{C, P, S-CCA, S-PAV, MAV, MAV, SAV, \} \), \( F(V) = \{\{a, b\}, \{a, c\}\} \) and \( F^R(P) = \{a, c\} \). Let \( a \in \text{maj}(\succ, \{a, x\}) \). Now, for all the rules considered before, \( F(V) = \{a, c\} \) and \( a \in F^R(P) \). Thus, none of these rules is monotonic. \( \square \)

MAVR satisfies both monotonicity and Pareto-efficiency. The only other rule that we considered in Section 3.1 that is monotonic is the trivial rule, which is however not Pareto-efficient. However, MAVR is not the only rule satisfying these two properties. This is also the case for \( F^R \), where \( F \) returns all pairs of candidates \( \{a, b\} \) such that either (i) neither \( a \) nor \( b \) is Pareto-dominated in \( V \) or (ii) \( a \) is the only candidate that dominates \( b \) in \( V \). Indeed, let \( b \) be a candidate dominated by \( a \) in some profile \( P = (V, \succ) \). Then, if there exists \( x \) such that \( \{x, b\} \in F(V) \), we must have \( x = a \) by definition of \( F \). Since \( a \) dominates \( b \), it is clear that \( \text{maj}(\succ, \{a, b\}) = a \), and \( b \notin F^R(P) \). This proves Pareto-efficiency.

Let's now show monotonicity. For a profile \( P = (V, \succ) \), let \( a \in F^R(P) \) a winning candidate, and \( P' \) be an \( \alpha \)-improvement of \( P \). Let \( \{x, a\} \in F(V) \) be the pair of finalists such that \( \text{maj}(\succ, \{x, a\}) = a \). \( x \) is either not dominated or dominated only by \( a \). If one new voter approves \( a \), then \( x \) is still either not dominated or dominated only by \( a \), therefore we still have \( \{x, a\} \in F(V) \). If the new voter approves \( a \), then we obviously have \( \{x, a\} \in F(V) \). If \( a \) wins the majority vote against \( x \) in \( P \), then it clearly wins it in \( P' \). Therefore, \( F \) is also monotonic and Pareto-efficient (but also neutral and anonymous), which means we cannot characterize MAVR with only these properties.

Finally, we focus on clone-proofness. Informally, this property means that adding a clone of a candidate does not change significantly the outcome of the election. Formally, let \( a \in C \), and a profile \( P' \) over set of candidates \( C' = C \cup \{a'\} \). \( P' \) is an \( \alpha \)-cloning extension of \( P \) if

1. For every \( v_i \) if and only if \( a' \in A_i \)
2. For every \( v_i \) if and only if \( x \in A_i \)

Definition 6. An AVR rule \( R \) is \( \alpha \)-cloning-proof if for any profile \( P \), candidate \( a \in C \), and an \( \alpha \)-cloning extension \( P' \) of \( P \), the two following conditions hold:

1. For every \( c \neq a \), \( c \in F^R(P) \) if and only if \( c \in F^R(P') \)
2. \( a \in F^R(P) \) if and only if \( |F^R(P') \cap \{a', a\}| \geq 1 \)

MAVR is not clone-proof. Let \( P = (\succ, V) = (a[b], b[a], b[ab]) \); MAVR \( (P) = \{b\} \). \( P' = (V', \succ') = (a[ba], b[a], ba[a]) \) is an \( \alpha \)-cloning extension of \( P \) and yet MAVR \( (P) = \text{maj}(\succ', \{a, a'\}) = \{a\} \). At first sight, CCAV and S-CCA
We now define a weaker version of clone-proofness, with a domain restriction that eliminates pathological profiles:

**Theorem 3.** No AVR rule is monotonic, weakly clone-proof.

**Theorem 2.** No AVR rule is clone-proof and Pareto-efficient.

**Definition 7.** Let \( \hat{R} \) be an AVR rule for two candidates \( \{a, b\} \). We define the \( \hat{R} \)-improvement of \( \hat{R} \) as:

\[
R' = R(\hat{R}) = (a|b, ba|, ba|, 10 \times cab). 
\]

Unfortunately, we have the following impossibility:

**Proposition 4.** CCAV\textsuperscript{R} and S-CCA\textsuperscript{R} are not clone-proof: CCAV\textsuperscript{R}(P) = S-CCA\textsuperscript{R}(P) = \{b\}; \ P' = (V', >') = (aa'|b, ba', baa') is an a-cloning extension of P; and yet CCAV(V') = S-CCA\textsuperscript{R}(V') = \{(b, a), (b, a'), (a, a')\}, so CCAV\textsuperscript{R}(P') = S-CCA\textsuperscript{R}(P') contains a, breaking clone-proofness.

Among the rules considered in Section 3.1 none is actually clone-proof. However, there exist AVR clone-proof rules: such a rule is defined by the AUC rule that selects the pairs of candidates maximizing \( f(x_1, x_2) = S_{V}(x_1) + S_{V}(x_2) - 2S_{V}(x_1x_2) \). However, this rule is not Pareto-efficient. More generally, clone-proofness and Pareto-efficiency are incompatible:

**Theorem 2.** No AVR rule is clone-proof and Pareto-efficient.

**Proof.** Assume \( F_R \) is clone-proof and Pareto-efficient. Take the profile \( P = (V, >) = (a|bc, b|ca, c|ab, 10 \times cab) \) and let \( P' = (V', >') = (aa'|b, ba', baa') \) be an a-cloning extension of \( P \). Because \( F_R(P) = \{b\} \), by clone-proofness we have \( F_R(P') = \{b\} \), therefore, \( \{a, b\} \in F(V') \) or \( \{a', b\} \in F(V') \). Without loss of generality, assume \( \{a, b\} \in F(V') \). Let \( P^* = (V^*, >^*) = (aa'|b, ba|, a'ba|) \) in which \( a' \) dominates \( b \). Since \( V^* = V' \), \( \{a, b\} \in F(V^*) \) and thus \( b \in F(R(P^*)) \), which contradicts Pareto-efficiency.

We now define a weaker version of clone-proofness, with a domain restriction that eliminates pathological profiles:

**Definition 7.** An AVR rule \( F_R \) is weakly clone-proof if it is clone-proof on every profile \( P \) such that no candidate \( c \in C \) is approved in every non-empty ballot: for every \( c \in C \), there exists a voter \( v \) such that \( A_v \neq \emptyset \) and \( c \notin A_v \).

CCAV\textsuperscript{R} and S-CCA\textsuperscript{R} are weakly clone-proof (Proposition 4). Recall that they are not Pareto-efficient. A rule weakly clone-proof and Pareto-efficient is \( F_R \), where \( F \) selects the CCAV finalists, and uses MAV as a tie-breaking if there are several pairs of finalists.

Unfortunately, we have the following impossibility:

**Theorem 3.** No AVR rule is monotonic, weakly clone-proof.

**Proof.** Assume \( F_R \) is monotonic and weakly clone-proof. Let \( P = (V, >) = (a|bc, b|ca, c|ab, 10 \times cab) \). We can assume without loss of generality \( \{a, b\} \in F(V) \), so \( a \in F_R(P) \). \( P' = (V', >') = (a|bc, b|ca, c|ab, 10 \times cab) \) is an \( a \)-improvement of \( P \). By monotonicity, \( a \in F_R(P') \). Because \( \text{maj}(>, \{a, c\}) = c \), this implies \( \{a, b\} \in F(V') \).

Let \( P'' = (V'', >'') = (a|bc, b|ca, c|ab, 10 \times cab) \). Since \( V'' = V' \) we have \( \{a, b\} \in F(V'') \), and \( b \in F_R(P'') \). Let \( P^* = (V^*, >^*) = (a|bc, b|ca, c|ab, 10 \times cab) \) a b-improvement of \( P'' \). By monotonicity, \( b \in F(R(P^*)) \). Since \( \text{maj}(>, \{b, c\}) = c \), we have \( \{a, b\} \in F(V^*) \).

Now consider \( \hat{P} = (V, >) = (a|c, a|c, c|a, 10 \times ca) \). We have \( F_R(\hat{P}) = \{c\} \). If we clone a into another candidate \( b \), we can define the a-cloning extension \( \hat{P}^* = (V^*, >^*) = (a|bc, ab|c, c|ab, 10 \times cab) \). By weak clone-proofness, neither \( a \) nor \( b \) is in \( F_R(\hat{P}^*) \), so \( \{a, b\} \notin F(\hat{V}^*) \). We obtain a contradiction and this concludes the proof.

This set of properties is minimal: MAV\textsuperscript{R} is monotonic and neutral; CCAV\textsuperscript{R} is weakly clone-proof and neutral; a rule with a constant pair of finalists is weakly clone-proof and monotonic.

**Proposition 4.** CCAV\textsuperscript{R} and S-CCA\textsuperscript{R} are weakly clone-proof but not AVR\textsuperscript{R}, S-PAV\textsuperscript{R}, S-Phr\textsuperscript{R}, EnePhr\textsuperscript{R}, PAV\textsuperscript{R} and SAV\textsuperscript{R} are not.

**Proof.** Let’s show that S-CCA\textsuperscript{R} and CCAV\textsuperscript{R} are weakly clone-proof. Let \( P = (V, >) \) be a profile in which there is no candidate \( c \) that is approved in every non-empty ballot. Let \( a \in C \) and \( P' = (V', >') \) be an a-cloning extension of \( P \) with a clone \( a' \) of \( a \). Let \( f_V(x, y) \) be the Chamberlin-Courant score for two candidates \( x \) and \( y \) in \( V \), i.e. \( f_V(x, y) = S_V(x) + S_V(y) - S_V(xy) \). For CCAV, we have for all \( x, y \notin \{a, a'\}, f_V(x, y) = f_V(x, y) \) and
We now want to explore the spectrum between rules that select the most popular candidates, typically MA V , and rules that favour diversity in the set of finalists, typically CCA V and S-CA V or, to a lesser extent, PAV and S-PAV.

|             | AV R | S-PA V R | S-Phr R | EnePhr R | S-CA V R | PAV R | CCA V R | SA V R | TRIV R |
|-------------|------|----------|---------|----------|---------|-------|---------|-------|--------|
| Pareto-efficient | ✓    | ✓        | ✓       | ✓        | ✓       | ✓     | ✓       | ✓     | ✓      |
| Monotonic    | ✓    | ✓        | ✓       | ✓        | ✓       | ✓     | ✓       | ✓     | ✓      |
| Strategy-proof |      |          |         |          |         |       |         |       |        |
| Weakly Clone-proof |      |          |         |          |         |       |         |       |        |

Table 4: Approval with Runoff rules and their properties

|                    | Strategy-proof | Monotonic | Clone-proof | Weakly clone-proof |
|--------------------|----------------|-----------|-------------|--------------------|
| Pareto-efficient   | \{\}           | MAV R     | \{\}        | CCA V R            |
| Strategy-proof     |                | TRIV R    | \{\}        | \{\}               |
| Monotonic          |                |           | MAV R       | MAV R              |
| Clone-Proof        |                |           | MAV R       | MAV R              |

Table 5: Summary of impossibilities between properties (assuming neutrality and anonymity). \{\} means there is no rule satisfying both properties. Otherwise, we propose a rule satisfying both properties — note that this is not a characterization.

\[f_{V'}(x, a) = f_{V'}(x, a') = f_V(x, a).\] Finally, \[f_V(x, a, a') = S_V(a) + S_V(a') - S_V(aa') = S_V(a).\] Let \(w\) be an approval winner in \(V\). Let’s prove that \(\{a, a'\} \notin (S-)CCA(V')\)

- If \(a\) is not an approval winner in \(V\), then \(\{a, a'\} \notin S-CCA(V')\) for sure. For CCAV, we have \(f_{V'}(w, a) = f_V(w, a) \geq S_V(w) > S_V(\alpha) = f_{V'}(\alpha, a')\), thus \(\{a, a'\} \notin CCA(V')\).

- If \(a\) is the approval winner in \(P\), since there is no candidate \(c \in C\) that is approved in every non-empty ballot, there exists a candidate \(x\) and a \(A_1 \in V\) such that \(x \in A_1\) and \(a \notin A_1\). Thus, \(f_{V'}(x, a) = f_V(x, a) = S_V(x) + S_V(a) - S_V(xa) \geq S_V(a) + 1 > S_V(x) = f_{V'}(\alpha, a')\), thus \(\{a, a'\} \notin CCA(V')\).

In all cases, \(\{a, a'\} \notin (S-)CCA(V')\) and the Chamberlain-Courant score of all other pairs is unchanged, so it is easy to observe that for all \(x \notin \{a, a'\}\), \(\{x, a', a\} \notin F(V')\) and if only if \(x, y \in F(V)\) if and only if \(x, y \in F(V')\) if and only if \(x, y \in F(V)\). Therefore, for all \(x \neq a', x \in F^R(P')\) if and only if \(x \in F(R(P'))\). Thus, CCAV is weakly clone-proof.

Let’s now show a counter-example for other rules. Consider the profile \(P = (V, \succ) = (a|b, a|b, |b, 10 \times ba|)\). Note that \(P\) satisfies the condition of the definition of weak clone-proofness. Since there are only two candidates, for every rule \(F\) we have \(F(V) = \{a, b\}\) and \(F(R(P)) = max(\succ, \{a, b\}) = \{b\}\). Let’s now clone \(a\) into another candidate \(a'\). We have \(a\)-cloning extension \(P' = (V', \succ') = (aa|b, aa|b, ba|, 10 \times ba|)\). For all rules \(F \in \{MAV, PAV, S-PAV, S-Phr, SA\}\), we have \((a, a') \in F(V')\) and \(a \in F^R(P')\), thus breaking clone-proofness. For EnePhr with a quota \(Q\) such that \(2Q < a\), we use the profile \(P_k = (V_k, \succ_k) = (b[a, k \times a|b, k \times ba])\) with a sufficiently high \(k\).

Finally, Table 5 summarizes the impossibilities between the different properties of this section and proposes a voting rule when two properties are possible together.

In this table, CCAV R is the rule that select the CCAV winners and solve ties using MAV. 2-AV is the \(\alpha\)-AV rule with \(\alpha = 2\), i.e. selecting the pairs of finalists \(\{x, y\}\) maximizing \(f(x, y) = S_V(x) + S_V(y) - 2S_V(xy)\). This way, the score of two clones \(a\) and \(a'\) is \(f(a, a') = 0\). This is the reason why this rule is always clone-proof.

Let’s show that CCAV R is weakly clone-proof and pareto-efficient. For weak clone-proofness, we can directly adapt the proof for CCAV R. For pareto-efficiency, consider a profile \(P = (V, \succ)\) in which \(a\) dominates \(b\). It is clear that for all \(x \neq a, b\) if \((x, b) \in CCA(V)\), then \((x, a) \in CCA(V)\). However, \((x, a)\) has a better approval score than \((x, b)\), so \((x, b)\) will never be returned by CCAV R and \(b\) cannot be a winner. Thus, the rule is Pareto-efficient.

5  Statistical Analysis with one-dimensional Euclidean preferences

We now want to explore the spectrum between rules that select the most popular candidates, typically MAV, and rules that favour diversity in the set of finalists, typically CCAV and S-CAV or, to a lesser extent, PAV and S-PAV.
In this section we focus on one-dimensional Euclidean preferences: we assume that there is a function $\phi : V \cup C \to \mathbb{R}$ such that $c \succ_i c'$ if $|\phi(c) - \phi(v_i)| < |\phi(c') - \phi(v_i)|$. Such uni-dimensional preferences are extremely common in the literature and, indeed, capture an important political structure often met in actual politics.

We also need to define the voters approval behavior. Unfortunately, the literature (empirical or theoretical) on this issue is sparse. [Green-Armytage and Tideman, 2020] use a threshold for approval that is the average utility but do not offer any justification for that assumption. Classical political science [Chapman, 1955; Cox, 1984; 1997] argues that in the first round of a run-off, like in a two-member district, the voter concentrates on which candidate is going to arrive second and third (this is in practice the more pressing question) and will vote on this basis. More precisely, [Laslier and Van der Straeten, 2016] put forward theoretical arguments that imply putting the threshold of approval at the utility level of the candidate supposed to arrive second in the race. From an empirical study, [Van der Straeten et al., 2018] conclude that the strategic model behaves well except with respect to candidates who are ex post ranked very low. For models that set individual thresholds of approval, the ones that impose a fixed number of approved candidates work well if this number is of the order of magnitude of the number of seats to be allocated, and the average utility model is the least satisfactory.

For the simulations in this paper, we will assume that there exists $d > 0$ such that every voter $v_i$ approves all candidates $c$ such that $|\phi(c) - \phi(v_i)| < d$. This simple one-parameter family of rules will allow to discuss in a clean manner the collective consequences of voters being more or less flexible in their approvals.

Given a distribution of voters, we want to know in which position a candidate can maximize his score with $\alpha$-seqAV rules.
5.1 Triangular distribution

We first assume that the distribution of voters follows the following triangular density function \( f \), all voters and candidates are located on \([-1, 1]\), and for all \( x \in [-1, 1] \), \( f(x) = \frac{1-|x|}{2} \). The distribution is showed in Figure 1a.

Let us consider \( \alpha \)-seqAV rules, for \( \alpha \in [0, 1] \): for all pairs of finalists, one finalist \( x_1 \) is an approval winner and the other finalist \( x_2 \) maximizes \( S_V(x_2) - \alpha S_V(x_1 x_2) \). Recall that \( \alpha = 0 \) leads to MAV and \( \alpha = 1 \) to S-CCA. The first finalist will always be the closest to the middle point of the interval (here 0). It can be shown that the position in which the second finalist will get the maximum score for specific \( \alpha \) and \( d \) is

\[
|x^*_2| = \begin{cases}
\frac{\alpha(1-d)}{2-\alpha} & \text{if } \alpha \leq 2d \\
1 + d - \frac{2d}{\alpha} & \text{if } 2d \leq \alpha \leq \frac{2d}{1-d} \\
2d & \text{if } \alpha > \frac{2d}{1-d}
\end{cases}
\]

**Proof.** We know that the distribution of voters follow the function \( f(x) = \frac{1-|x|}{2} \) and the first finalist \( w \) is at position \( \phi(w) = 0 \). Let \( c \) be a candidate at position \( \phi(c) = x \).

First, consider the case \( x \leq d \), then \(-d \leq x - d \leq 0 \) and the score of \( c \) is given by

\[
S(c) = S_V(c) - \alpha S_V(cw) = \int_{x-d}^{x+d} f(t)dt - \alpha \int_{x-d}^{d} f(t)dt = \frac{1}{2}((1-\alpha)(0 - (x-d)^2) + (x+d - (x-d)^2 - 0) + C = \frac{1}{2}(x^2(\alpha - 1) - \frac{1}{2}) + x((1-\alpha)(-1 + d) + 1 + d)) + C = \frac{1}{2}(x^2\alpha - \frac{1}{2} + x\alpha(1-d)) + C
\]

where \( C \) is a constant independent of \( x \). The maximum of this function is reached when \( x^* = \frac{\alpha(1-d)}{2-\alpha} \), which is the optimal position. Note that \( x^* \leq d \) if and only if \( \alpha(1-d) \leq (2-\alpha)d \), and therefore \( \alpha \leq 2d \). This gives us the first formula.

Let’s now assume that \( 2d \geq x \geq d \). The score of \( c \) is now given by

\[
S(c) = S_V(c) - \alpha S_V(cw) = \int_{x-d}^{x+d} f(t)dt - \alpha \int_{x-d}^{d} f(t)dt = (1-\alpha)(-x - \frac{x^2-2xd + d^2}{2} + d) + (x+d - \frac{x^2+2dx + d^2}{2}) + C = \frac{1}{2}(x^2\alpha - \frac{1}{2} + x\alpha(1-d) - 2d)) + C
\]

where \( C \) is a constant independent of \( x \). The maximum of this function is reached when \( x^* = \frac{\alpha(1+d) - 2d}{1-\alpha} = 1 + d - \frac{2d}{\alpha} \). Note that \( d \leq x^* \leq 2d \) if and only if \( 0 \leq 1 - \frac{2d}{\alpha} \leq d \) so \( \frac{2d}{1-d} \geq \alpha \geq 2d \). This gives the second formula.

Finally, if \( x^* \geq 2d \) then \( S(x) = \int_{x-d}^{x+d} f(t)dt \) and since \( f(t) \) is decreasing, this is maximum when \( x^* = 2d \). Therefore, for \( \alpha > \frac{2d}{1-d} \) we have \( x^* = 2d \). This gives the third formula and concludes the proof.
Table 6: The datasets we used. For 2017-Crolles, half of the voters responded thinking there will be no runoff (2017-Crolles-1), and the other half thinking there will be a runoff (2017-Crolles-2).

This optimal distance to the first finalist is depicted on Figure 2a. This figure clearly shows that for rules close to MAV, the second finalist is quite centrist, and the closer we are to S-CCAV, the more extreme is the second finalist. This being said, \( \alpha \)-seqAV rules starts to be equivalent to S-CCAV before reaching \( \alpha = 1 \) when \( d < \frac{4}{3} \).

### 5.2 Gaussian distribution

Now, we assume that we have a Gaussian distribution of voters, centered at 0 and with standard deviation 1/2, as shown in Figure 1b. We cannot compute the exact optimum for each \( d \) and \( \alpha \) so we used simulations on synthetic data instead. We sampled 20,000 voters and 1,000 candidates and as before, every voter approves candidates that are at distance \( \leq d \). For each \( d \) and \( \alpha \) we compute the two finalists and observed their positions on the line. The first selected finalist is always the closest to the center, and Figure 2b shows the distance between the two finalists for \( \alpha \)-seqAV rules with \( \alpha \in [0, 1] \) and various \( d \). We observe that Figure 2b is very similar to Figure 2a.

### 6 Experiments

Finally, we want to compare the different rules on real data. We used approval ballot datasets from different sources:

- Datasets from the poll "Voter Autrement" conducted in several cities during the 2017 French presidential election [Bouveret et al., 2019]. Each dataset contains approval ballots of around 1000 voters for the 11 running candidates.
- A dataset from the 2002 French presidential election, which had 16 candidates.
- Two datasets of the San Sebastian Poster Competition held during The Summer School on Computational Social Choice (17 candidates, around 60 voters per dataset).

Table 6 summarizes our different datasets.

The first thing we did was to debiaised the results of the 2017 datasets. Indeed, the pollsters asked respondents who they actually voted in the election, and the distribution of votes is very different between the poll respondents and the global results. For instance, in the Strasbourg dataset, 35% of the voters indicated that they voted JLM in the election, against 20% nationwide. Similarly, 3% of the voters in this poll indicated they voted MLP, against 21% nationwide.

After this debiasing step, we can look at the approval score of candidates in the election. For instance, is there a candidate with more than 50% approvals? Table 7 and 8 summarize the approval scores of the main candidates in the different datasets of 2002 and 2017 elections. We can see that it is actually tight between the different candidates.

We also played with the distribution of the approval ballots to understand the affinity network of the candidates. Two candidates are closer if voters that approve one of them tend to approve the second one. We compute the affinity between two candidates with the Jaccard index, i.e. we divide the number of voters who approved both candidates by the number of voters who approved at least one of them. Figure 3 shows the affinity network for the 2017 presidential election in Grenoble. The size of a node is proportional to its approval score. We can clearly see an group with candidates from the left (in red), and another one with candidates from the right (in blue), EM being somewhat in the middle of the two groups. The affinity networks of all datasets can be found in the Appendix

---

4The last two datasets are available on [www.preflib.org](http://www.preflib.org)
### Approval with Runoff

|        | JLM | BH | EM | FF | MLP |
|--------|-----|----|----|----|-----|
| Real   | 19.6% | 6.4% | 24% | 20% | 21.3% |
| Strasbourg | 45.5% | 37.3% | 46.3% | 26.7% | 25.8% |
| HSC    | 41.2% | 36.6% | 45.6% | 29.4% | 28.7% |
| Grenoble | 41.8% | 43.2% | 42.3% | 33.2% | 29.5% |
| Crolles-1 | 40.5% | 36.3% | 54.7% | 34% | 30.3% |
| Crolles-2 | 43.6% | 40.7% | 55.3% | 38.7% | 31.4% |

Table 7: Approval scores of candidates in the 2017 election datasets

|        | LJ | NM | FB | JC | MLP |
|--------|----|----|----|----|-----|
| Real   | 16.2% | 5.2% | 6.8% | 19.9% | 16.9% |
| Approvals | 40.5% | 28.8% | 33.4% | 36.4% | 14.6% |

Table 8: Approval scores of candidates in the 2002 election dataset

![Affinity network](image)

Figure 3: Affinity network for the dataset 2017-Grenoble with a threshold at 10%, i.e. the Jacquard index must be > 0.1 for the edge to be visible

| Rule         | SAV   | CCAV  | S-CCAV | PAV   | S-PAV | S-Phr | SAV   |
|--------------|-------|-------|--------|-------|-------|-------|-------|
| 2017-Strasbourg | EM/JLM | EM/JLM | EM/JLM | EM/JLM | EM/JLM | EM/JLM | BH/JLM |
| 2017-Grenoble   | BH/JLM | BH/FF | BH/FF | EM/JLM | EM/JLM | EM/JLM | FF/BH |
| 2017-HSC        | EM/JLM | EM/JLM | EM/JLM | EM/JLM | EM/JLM | EM/JLM | EM/JLM |
| 2017-Crolles-1  | EM/JLM | EM/MLP | EM/MLP | EM/JLM | EM/JLM | EM/JLM | EM/JLM |
| 2017-Crolles-2  | EM/JLM | EM/MLP | EM/MLP | EM/JLM | EM/JLM | EM/JLM | EM/JLM |
| 2002-Presidential | LJ/JC | LJ/JC | LJ/JC | LJ/JC | LJ/JC | LJ/JC | LJ/JC |
| Best-Poster-A   | #1/#2 | #1/#6 | #1/#6 | #1/#4 | #1/#4 | #1/#4 | #1/#2 |
| Best-Poster-B   | #1/#2 | #1/#2 | #1/#2 | #1/#2 | #1/#2 | #1/#2 | #1/#2 |

Table 9: Finalists with different approval rules on all our datasets

We ran the different rules presented in Section 3.1 on our datasets. Table 9 summarizes the finalists obtained for each dataset and each rule. For some datasets, the choice of the rule has a strong impact on the finalists, which suggests that the choice of the ABC rule should be made with care.

We can also look at the evolution of the finalists with $\alpha$-AV and $\alpha$-seqAV rules when $\alpha$ varies from 0 to 1. This gives us a spectrum of rules from MAV to (S-)CCAV and we can see how the results evolve between the extremes. For MAV, the finalists are the two candidates from the strongest group, and for CCAV, they are from two very different groups.
Figure 4: The two finalists in 2017-Grenoble Dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.

Figure 4a depicts for instance the evolution of the pair of finalists for $\alpha$-AV rules for the 2017-Grenoble Dataset. We can see that the pairs of finalists change twice and involve 4 different candidates (Liberal, Left, Socialist and Conservative). Pairs that are never selected are represented by grey lines.

Figure 4b depicts the evolution of the $\alpha$-seqAV score of candidates for second finalist spot in the 2017-Grenoble dataset. The approval winner is the Socialist candidate; the second finalist is either the Liberal or the Conservative.

More results and figures for these datasets, and for other datasets, are in the Appendix.

7 Conclusion

Our main message is that approval with runoff is not one rule but a family of rules, parameterized by the ABC rule chosen for determining the finalists. Our axiomatic and experimental results in Sections 4.2, 5 and 6 show that this choice does make a big difference. If such rules have to be used in political elections, the choice of the ABC rule will be crucial, and is far from easy, but our results already give some useful elements: on a political election on a single-peaked profile on a classical left-right spectrum, we conclude from our theoretical and experimental results that: multiwinner approval voter tends to select two centrist finalists (possibly clones of each other); Chamberlin-Courant tends to select a left-wing and a right-wing candidate; proportional approval voting is inbetween and tends to select a center-left and/or a center-right candidate; greedy versions of the latter two rules select the candidate with the largest number of approvals, and a left-wing or right-wing candidate.

An important question is, will citizens understand and accept such rules especially in comparison with plurality with runoff and standard (single-winner) approval voting? Will there be a difference between citizens used to runoff voting in their country and those who are not?

References

Jean-François Laslier and Karine Van der Straeten. Strategic voting in multi-winners elections with approval balloting: a theory for large electorates. Social Choice and Welfare, 47(3):559–587, 2016.

Karine Van der Straeten, Romain Lachat, and Jean-François Laslier. Strategic voting in multi-winner elections with approval balloting: An application to the 2011 regional government election in Zurich. In Laura B. Stephenson, John H. Aldrich, and André Blais, editors, The Many Faces of Strategic Voting. Tactical Behavior in Electoral Systems Around the World, pages 178–202. The University of Michigan Press, 2018.

Steven J. Brams and M. Remzi Sanver. Voting systems that combine approval and preference. In The Mathematics of Preference, Choice and Order, Studies in Choice and Welfare, pages 215–237. Springer, 2009.

Martin Lackner and Piotr Skowron. Approval-based committee voting: Axioms, algorithms, and applications. CoRR, abs/2007.01795, 2020. URL https://arxiv.org/abs/2007.01795.

M. Remzi Sanver. Approval as an intrinsic part of preference. In Handbook on Approval Voting, pages 469–481. Springer, 04 2010. doi[10.1007/978-3-642-02839-7_20]

James Green-Armytage and T. Nicolaus Tideman. Selecting the runoff pair. Public Choice, 182(1):119–137, 2020.
Martin Lackner and Jan Maly. Approval-based shortlisting. In AAMAS, pages 737–745. ACM, 2021.
Dominik Peters. Proportionality and strategyproofness in multiwinner elections. CoRR, abs/2104.08594, 2021. URL https://arxiv.org/abs/2104.08594
Brian Chapman. Political Parties: Their Organization and Activity in the Modern State. International Affairs, 31(2): 208–208, 04 1955. ISSN 0020-5850. doi:10.2307/2604342 URL https://doi.org/10.2307/2604342
Gary W. Cox. Strategic electoral choice in multi-member districts: Approval voting in practice? American Journal of Political Science, 28:722, 1984.
Gary W. Cox. Making Votes Count: Strategic Coordination in the World’s Electoral Systems. Political Economy of Institutions and Decisions. Cambridge University Press, 1997. doi:10.1017/CBO9781139174954
Sylvain Bouveret, Renaud Blanch, Antoinette Baujard, François Durand, Herrade Igersheim, Jérôme Lang, Annick Laruelle, Jean-François Laslier, Isabelle Lebon, and Vincent Merlin. Voter Autrement 2017 for the French Presidential Election. Technical Report, HAL halshs-02379941, November 2019. URL https://halshs.archives-ouvertes.fr/halshs-02379941
A  Affinity networks

Figure 5: Affinity network for the dataset 2017-Strasbourg with a threshold at 10%
Approval with Runoff

Figure 6: Affinity network for the dataset 2017-HSC with a threshold at 10%

Figure 7: Affinity network for the dataset 2017-Grenoble with a threshold at 10%
Figure 8: Affinity network for the dataset 2017-Crolles-1 with a threshold at 10%
Figure 9: Affinity network for the dataset 2017-Crolles-2 with a threshold at 10%

Figure 10: Affinity network for the dataset 2002-Presidential with a threshold at 12%
Figure 11: Affinity network for the dataset Poster-1 with a threshold at 30%
Figure 12: Affinity network for the dataset Poster-2 with a threshold at 30%
B Results for all datasets

Figure 13: The two finalists in 2017-Grenoble Dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.

Figure 14: The two finalists in 2017-Crolles-1 Dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.
Finalists in Crolles-2

(a) $\alpha$-AV rules

Finalists in HSC

(a) $\alpha$-AV rules

Second finalist in Crolles-2 (Liberal is first)

(b) $\alpha$-seqAV rules

Figure 15: The two finalists in 2017-Crolles-2 Dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.

Finalists in HSC (Liberal is first)

(b) $\alpha$-seqAV rules

Figure 16: The two finalists in 2017-HSC Dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.
Figure 17: The two finalists in 2017-Strasbourg Dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.

Figure 18: The two finalists in the 2002-Presidential dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.
Figure 19: The two finalists in the Best-Poster-A dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.

Figure 20: The two finalists in the Best-Poster-B dataset for $\alpha$-AV and $\alpha$-seqAV rules with $\alpha \in [0, 1]$. For every $\alpha$, the upper line indicates the finalists (resp. the second finalist) for the $\alpha$-AV (resp. $\alpha$-seqAV) rule.