On the Robustness of Preference Aggregation in Noisy Environments

Ariel D. Procaccia, Jeffrey S. Rosenschein, and Gal A. Kaminka

Abstract

In an election held in a noisy environment, agents may unintentionally perturb the outcome by communicating faulty preferences. We investigate this setting by introducing a theoretical model of noisy preference aggregation and formally defining the (worst-case) robustness of a voting rule. We use our model to analytically bound the robustness of various prominent rules. Our results essentially specify the voting rules that allow for reasonable preference aggregation in the face of noise.

1 Introduction

Preference aggregation mechanisms, and voting rules in particular, have been the object of scientific study for many years. Such mechanisms are used to aggregate the preferences of human or synthetic agents, over alternatives (or candidates). The alternatives in question may be entities such as joint plans for execution, schedules [7], movie choices [5], etc. A voting rule generates an outcome that reflects the individual preferences over candidates, while striving to satisfy different desiderata. Indeed, much of the research in social choice theory has focused on formally analyzing the properties of social choice mechanisms, with respect to these desiderata.

One important feature of study in preference aggregation mechanisms is their resistance to manipulation. Such manipulations are instances of adversarial worst-cases in the context of mechanisms: they consider self-interested voters that intentionally cast untruthful ballots in order to manipulate the outcome in their favor. An important theorem asserts that every voting rule (under certain minimal assumptions) is manipulable [6, 9]. More recent work in computer science suggests that computational complexity may help circumvent this impossibility result [1, 3].

However, little attention has been paid to a simpler—and arguably more common—form of voting manipulation, where the truthful votes are *unintentionally* changed, as a result of uncertainty in the actions or perception of the agents casting the votes. For instance, agents may misunderstand the choices laid out for them, and may thus inadvertently cast a vote that is inconsistent with their true choice. Or, in the case of robots operating where communication is unreliable, true choices may be miscommunicated, resulting again in unintentional manipulation.

This paper takes first steps towards a formal analysis of the impact of errors in the preferences of voters. We define the k-robustness of a voting rule to be the resistance of the rule to k faults. In more detail, it is the probability that the outcome changes as a result of the faults, when each fault is chosen independently at random. We analyze the connection between 1-robustness (i.e., resistance to a single fault) and k-robustness, and conclude that it is sufficient to examine the 1-robustness of different rules. Most importantly, we use our definitions and tools to give tight upper and lower bounds on the robustness of several prominent voting rules. In fact, we show that the robustness of voting rules is extremely diverse, with some rules positioned at both ends of the spectrum.

Given that voters rank the candidates (voters express *ordinal preferences*), we analyze a theoretical model where a fault is a switch in the rankings of two adjacent candidates (e.g., the fifth-ranked candidate is accidentally ranked sixth, and the sixth is ranked fifth). Such faults may easily be caused by confusion on the part of voters, or even by a single bit-flip when communicating the votes (see Section 3). Our goal is to understand the robustness of different voting rules to such faults; this understanding would aid system designers in selecting voting rules that can faithfully aggregate the preferences of agents in the system.

Previous work by Kalai [8] has investigated the issue of *noise-sensitivity* of social welfare functions in simple games; such functions give an entire social ranking of the candidates, instead of simply designating the winner of the election. The author engages in an asymptotic average-case analysis, where the basic assumption is that the voters' votes are distributed uniformly at random. Kalai presents a family of "chaotic" social welfare functions: a change in the preferences of a small fraction of the voters leads to social preferences that are asymptotically uncorrelated with the original preferences. In contrast, our model in this paper is quite different; in addition, we are interested in examining the robustness of prominent voting rules, as opposed to investigating extreme asymptotic phenomena.

This paper is organized as follows. In Section 2 we give an introduction to voting, and describe the voting rules we examine thereafter. In Section 3 we outline our model of preference profile errors, and give some general results regarding robustness. In Section 4 we bound the 1-robustness of some prominent voting rules, and in Section 5 we discuss our results and directions for future work.

2 Preliminaries

In this section we give a brief introduction to classic social choice theory. The information here is sufficient to understand the paper, but readers who are interested in more details can consult [2].

Let $V = \{v^1, v^2, \dots, v^n\}$ be the set of voters, and let $C = \{c_1, c_2, \dots, c_m\}$ be the set of candidates, $|C| = m$. We usually use the index i (in superscript) to refer to voters, and the index j (in subscript) to refer to candidates.

Let $\mathcal{L} = \mathcal{L}(C)$ be the set of all linear orders[1] on C. Each voter has ordinal preferences $\succ^i \in \mathcal{L}$, i.e., each voter v^i ranks the candidates: $c_{j_1} \succ^i c_{j_2} \succ^i \cdots \succ^i c_{j_m}$. We refer to $\succ^V = \langle \succ^1, \ldots, \succ^n \rangle \in \mathcal{L}^N$ as a *preference profile*.

Given $\succ^i$, let $j_1, \ldots, j_m$ be indices of candidates such that $c_{j_1} \succ^i c_{j_2} \succ^i \cdots \succ^i c_{j_m}$; we denote by $\pi_l(\succ^i)$ the candidate that voter i ranks in the l'th place, i.e., $\pi_l(\succ^i) = c_{j_l}$. We also denote by l^i_j the ranking of c_j in $\succ^i$; it holds that $\pi_{l^i_j}(\succ^i) = c_j$.

2.1 Voting rules

A *voting rule* is a function $F : \mathcal{L}^V \to C$, i.e., a mapping from preferences of voters to candidates, which designates the winning candidate. We shall consider the following voting rules:

- *Scoring rules* are defined by a vector $\vec{\alpha} = \langle \alpha_1, \ldots, \alpha_m \rangle$.[2] Given $\succ \in \mathcal{L}^N$, the score of candidate j is $s_j = \sum_i \alpha_{l^i_j}$. The candidate that wins the election is $F(\succ) = \text{argmax}_j s_j$. Some of the well-known scoring rules are:
  - *Borda*: $\vec{\alpha} = \langle m-1, m-2, \ldots, 0 \rangle$.
  - *Plurality*: $\vec{\alpha} = \langle 1, 0, \ldots, 0 \rangle$.
  - *Veto*: $\vec{\alpha} = \langle 1, \ldots, 1, 0 \rangle$.
- *Copeland*: we say that candidate j beats j' in a pairwise election if $|\{i : l^i_j < l^i_{j'}\}| > n/2$. The score s_j of candidate j is the number of candidates that j beats in pairwise elections, and $\text{Copeland}(\succ) = \text{argmax}_j s_j$.
- *Maximin*: the Maximin score of candidate j is the candidate's worst performance in a pairwise election: $s_j = \min_{j'} |\{i : l^i_j < l^i_{j'}\}|$, and $\text{Maximin}(\succ) = \text{argmax}_j s_j$.
- *Bucklin*: for any candidate c_j and $l \in \{1, \ldots, m\}$, let $B_{j,l} = \{i : l^i_j \leq l\}$. It holds that $\text{Bucklin}(\succ) = \text{argmin}_j(\min\{l : |B_{j,l}| > n/2\})$.
- *Plurality with Runoff*: The election proceeds in two rounds. After the first round, only the two candidates that maximize $|\{i \in N : l^i_j = 1\}|$ survive. In the second round, a pairwise election is held between these two candidates.

3 Our Model of Faults and Robustness

We consider situations where (for example) noisy communication leads to changes in voters' rankings of candidates. The exact manifestation of these

[1] Binary relations that satisfy antisymmetry, transitivity, and totality.

[2] More formally, a scoring rule is defined by a sequence of such vectors, one for each value of m, but we abandon this formulation for clarity's sake.

faults largely depends on the representation of preferences. In order to obtain results that are as general as possible, we here simply regard a fault as an alteration of one voter's ordering of candidates, which nevertheless maintains the integrity of the voter's preferences as a linear ordering (other types of faults remain for future work).

Definition 1. A preference profile $\succ_1^V$ is obtained from a preference profile $\succ^V$ by an *elementary transposition* (write: $\succ^V \rightsquigarrow \succ_1^V$) if there exists a voter v^i and $l \in \{2, \ldots, m\}$ such that:

1. for all $i' \neq i$, $\succ^{i'} = \succ_1^{i'}$.
2. $\pi_l(\succ^i) = c = \pi_{l-1}(\succ_1^i)$.
3. $\pi_{l-1}(\succ^i) = c' = \pi_l(\succ_1^i)$.
4. $\succ^i\downarrow_{C\setminus\{c,c'\}} = \succ_1^i\downarrow_{C\setminus\{c,c'\}}$.

We say that $\pi_{l-1}(\succ^i)$ was *demoted* and $\pi_l(\succ^i)$ was *promoted*.

Example 1. The preference profile

$\succ^1$	$\succ^2$
c_1	c_2
c_3	c_1
c_2	c_3

is obtained from the preference profile

$\succ^1$	$\succ^2$
c_1	c_1
c_3	c_2
c_2	c_3

by an elementary transposition that promotes c_2 and demotes c_1 (in the notations of the definition, $i = 2$ and $l = 2$).

In other words, we focus here on faults where a switch has occurred between two adjacent candidates in a voter's ranking of candidates. Such faults are interesting from a theoretical perspective, but may also occur in practice. For instance, if voters build their preferences incrementally, they may easily be confused by spatial proximity of alternative candidates. In other cases, depending on the representation and communication protocol, communication errors may cause a switch to occur. Below we describe a representation for preferences, in which a flip of a single bit either causes a switch between two adjacent candidates, or can easily be detected.

3.1 The Pairwise Representation

We here describe a representation of preferences that is compatible with our fault model, and argue that it has some nice advantages. One can represent preferences using a bit for each ordered pair of candidates (with $\binom{m}{2}$ ordered pairs): the bit is 1 if the first candidate is preferred to the second, and 0 otherwise. We shall refer to this representation as the *pairwise representation.* In this representation, a flip of a single bit corresponding to a pair of adjacent candidates in the ordering entails an elementary transposition. However, flipping a bit that does not correspond to adjacent candidates would create an ordering that is not transitive, and therefore not linear. Indeed, if (w.l.o.g.) $c_1 \succ c_2 \succ c_3$, and the bit corresponding to c_1 and c_3 is flipped, then we obtain the preferences $c_1 \succ c_2$, $c_2 \succ c_3$, and $c_3 \succ c_1$ — transitivity is not satisfied. It follows that faults which switch the ranking of two non-adjacent candidates can always be detected. So, when considering bit flips that may change the outcome without being detected, we can restrict our attention to faults that manifest themselves as elementary transpositions.

The pairwise representation is not the most compact possible. Consider the following elementary representation: each voter specifies the location of each candidate in their ranking; this requires $m \log m$ bits. In the pairwise representation, each voter requires $\binom{m}{2} = \frac{m(m-1)}{2}$ bits to express its preferences.

On the other hand, the pairwise representation allows us to test properties in constant time using bitmasks. For instance, say we want to test if a voter has ranked candidate c_1 highest. We construct the ordered pairs in a way that candidate c_1 is always first; we then examine the conjunction (a bitwise AND) of all pairs in which c_1 participates — c_1 is ranked first iff this conjunction is 1. This can be done in constant time, while in the elementary representation, one would have to examine all $\log m$ bits that represent c_1's ranking in order to answer this question. Similarly, given that one knows (from polls, for example) which candidates are placed in the first k places, one can test in constant time whether a candidate is ranked in place $k+1$, using a bitmask on the pairwise representation.

3.2 The Definition of Robustness

So far, we have described our model of faults, and have argued that it has practical justification. The switches in preferences that we consider may seem harmless, but in fact, for essentially any voting rule, there exist instances where even one switch changes the outcome of the election.

Theorem 1. *Let $F : \mathcal{L}^V \to C$ be a voting rule such that $Ran(F) > 1$. Then there exists a preference profile $\succ^V$ and a profile $\succ_1^V$ which is obtained from $\succ^V$ by an elementary transposition, such that $F(\succ^V) \neq F(\succ_1^V)$.*

Proof. Assume that for every preference profile $\succ^V$ and any elementary transposition, the outcome does not change. Let $\succ^V$ and $\succ_1^V$ be any two preference profiles; we will derive a contradiction to the assumption on F's range by showing that they necessarily have the same value under F.

Indeed, a preference profile is essentially a series of permutations on C (one for each voter); a basic result regarding permutation groups implies that $\succ_1^V$ can be obtained from $\succ^V$ by iterative elementary transpositions [4]. In other words, there are $\succ^V_{i_1}, \ldots, \succ^V_{i_t}$ such that $\succ^V_{i_1} = \succ^V$, $\succ^V_{i_t} = \succ^V_1$, and each $\succ^V_{i_{j+1}}$ can be obtained from $\succ^V_{i_j}$ by an elementary transposition, for $j = 1, \ldots, t-1$. By our assumption, all $\succ^V_{i_j}$ have the same value under F, and in particular $F(\succ^V) = F(\succ^V_1)$ — a contradiction. □

Given a voting rule, we wish to consider the implications of faults *in the worst-case*, i.e., in the worst instance. Theorem 1 motivates a probabilistic analysis: we will calculate the probability of the faults affecting the outcome in the worst-case.

Given a preference profile $\succ^V$, we define the probability distribution $D_k(\succ^V)$ over preference profiles as follows: the probability of the preference profile $\succ^V_1$ is the probability of obtaining $\succ^V_1$ from $\succ^V$ by k elementary transpositions chosen independently and randomly. In other words, in order to draw a profile $\succ^V_1$ according to this distribution, we independently choose k values $\{l_1, l_2, \ldots, l_k\}$ and k values $\{i_1, \ldots, i_k\}$, where each l_j is chosen according to the uniform distribution over $\{2, \ldots, m\}$, and each i_j is chosen according to the uniform distribution over $\{1, \ldots, n\}$. Now, starting with $\succ^V$, we perform k successive elementary transpositions — the j'th transposition promotes candidate $\pi_l(\succ^i)$ and demotes $\pi_{l-1}(\succ^i)$.

Definition 2. The *k-robustness* of a preference profile $\succ^V$ is:

$$\rho(F, \succ^V) = \Pr_{\succ^V_1 \sim D_k(\succ^V)} [F(\succ^V) = F(\succ^V_1)].$$

The k-robustness of a profile reflects its immunity to k independent faults. As our analysis is worst-case, in order to define the robustness of a voting rule we take the minimum over all instances:

Definition 3.

The *k-robustness* of a voting rule F with n voters and m candidates is:

$$\rho_k^{n,m}(F) = \min_{\succ^V \in \mathcal{L}(C)^n} \rho(F, \succ^V).$$

Example 2. Consider the Plurality rule with 3 voters and 2 candidates, and consider the preference profile $\succ^V$ given by:

$\succ^1$	$\succ^2$	$\succ^3$
c_1	c_1	c_2
c_2	c_2	c_1

The outcome of this election is c_1. There are three possible profiles resulting from an elementary transposition:

$\succ^1_1$	$\succ^2_1$	$\succ^3_1$
c_2	c_1	c_2
c_1	c_2	c_1

$\succ^1_2$	$\succ^2_2$	$\succ^3_2$
c_1	c_2	c_2
c_2	c_1	c_1

$\succ^1_3$	$\succ^2_3$	$\succ^3_3$
c_1	c_1	c_1
c_2	c_2	c_2

In two of these profiles, the outcome is c_2. Therefore, $\rho(F, \succ^V) = 1/3$. Repeating the same calculation for all preference profiles $\succ^V \in \mathcal{L}(C)^n$ and taking the minimum, it is possible to conclude that $\rho_1^{3,2}(\text{Plurality}) = 1/3$.

3.3 Bounding k-robustness with 1-robustness

The definition of $D_k(\succ^V)$ as sampling k independent elementary transpositions allows a very strong link between 1-robustness and k-robustness: a lower bound on the former entails a lower bound on the latter.

Proposition 2. $\rho_k^{n,m}(F) \geq (\rho_1^{n,m}(F))^k$.

Proof. Consider the preference profile $\succ_1^V$, and the preference profile $\succ_2^V$ obtained by k independent and random elementary transpositions — we claim that the probability that $F(\succ_1^V) = F(\succ_2^V)$ is at least $(\rho_1^{n,m})^k$.

Indeed, let $\succ_{i_1}^V, \ldots, \succ_{i_{k+1}}^V$ be the intermediate preference profiles obtained by the elementary transpositions, i.e., $\succ_{i_1}^V = \succ_1^V$, $\succ_{i_{k+1}}^V = \succ_2^V$, and each $\succ_{i_{j+1}}^V$ is obtained from $\succ_{i_j}^V$ by an independently and randomly chosen elementary transposition, for $j = 1, \ldots, k$. By the definition of 1-robustness, we have that for every preference profile $\succ^V$, the probability that one randomly chosen elementary transposition does not change the outcome of the election under F is at least $\rho_1^{n,m}(F)$. Therefore, we have that for $j = 1, \ldots, k$,

$$\Pr[F(\succ_{i_j}^V) = F(\succ_{i_{j+1}}^V) \mid \succ_{i_j}^V] \geq \rho_1^{n,m}(F).$$

By analyzing the conditional probabilities we have that:

$$\begin{aligned}\Pr\left[F(\succ_1^V) = F(\succ_2^V)\right] &= \Pr\left[\forall j = 1, \ldots, k,\ F(\succ_{i_j}^V) = F(\succ_{i_{j+1}}^V)\right] \\ &= \prod_{j=1}^{k} \Pr\left[F(\succ_{i_j}^V) = F(\succ_{i_{j+1}}^V) \mid \succ_{i_j}^V\right] \\ &\geq (\rho_1^{n,m})^k .\end{aligned}$$

□

The above proposition is very useful when the number of errors is constant. Otherwise, the bound on k-robustness which the proposition yields may not be very good, even if the voting rule seems 1-robust. Nevertheless, we have the following immediate corollary regarding $k = m$:

Corollary 3. *Let F be a voting rule such that $\rho_1^{n,m}(F) \geq 1 - x/m$ for some constant x, and let $\epsilon > 0$. Then $\rho_m^{n,m}(F) \geq \frac{1}{e^x} - \epsilon$ for a large enough m.*

4 Results on 1-Robustness

Proposition 2 dictates the direction of the bulk of our results: we are satisfied with calculating the 1-robustness of voting rules. If we achieve a high lower bound, this also implies high k-robustness (at least for a constant k). However, in case 1-robustness is low, there is no point in considering the rule's k-robustness.

Remark 1. Given the number of voters and candidates, and a preference profile $\succ^V$, there are exactly $n(m-1)$ possible elementary transpositions. Therefore:

$$\rho_1^{n,m}(F) = \frac{|\{\succ_1^V \in \mathcal{L}(C)^n : \succ^V \rightsquigarrow \succ_1^V \ \wedge F(\succ^V) = F(\succ_1^V)\}|}{n(m-1)}.$$

Before we deal with specific voting rules, we note that we cannot expect a rule's 1-robustness to be exactly 1.

Proposition 4. *Let $F : \mathcal{L}(C)^n \to C$ be a voting rule such that $Ran(F) > 1$. Then $\mu_1^{n,m}(F) < 1$.*

Proof. Follows directly from Proposition 1 and the definition of 1-robustness. □

4.1 Scoring rules

In this subsection we fully characterize the robustness of scoring rules as a function of their parameters. Our results imply that some common scoring rules are very robust, while others are extremely susceptible to faults.

Given a scoring rule F with parameters $\vec{\alpha}$, let $A_F = |\{l \in \{2, \ldots, m\} : \alpha_{l-1} > \alpha_l\}|$; denote $|A_F| = a_F$.

Proposition 5. *Let n and m be the number of voters and candidates, let F be a scoring rule. Then $\rho_1^{n,m}(F) \geq \frac{m-1-a_F}{m-1}$.*

Proof. For any preference profile $\succ^V$, the outcome can only be affected by elementary transpositions that promote $\pi_l(\succ^i)$, for some $l \in A_F$ and i, and demote $\pi_{l-1}(\succ^i)$. For each voter v^i, there are exactly a_F such values of l, out of $m-1$ possible elementary switches. Therefore, the number of elementary transpositions that are guaranteed *not* to change the outcome is at least $n(m-1) - a_F n$, and the 1-robustness of F is at least $\frac{n(m-1)-a_F n}{n(m-1)} = \frac{m-1-a_F}{m-1}$. □

We match this lower bound with a pretty tight upper bound. In this example, we require that the number of candidates divide the number of voters. However, such a special case is sufficient, as it implies that the lower bound cannot be improved in general.

Proposition 6. *Let n and m be the number of voters and candidates such that m divides n, and let F a scoring rule. Then $\rho_1^{n,m}(F) \leq \frac{m-a_F}{m}$.*

Proof. By the assumption on n and m, it is possible to group the voters in m subsets of size d, $T_1, \ldots, T_m$. Consider the preference profile $\succ^V$ where the subsets of voters vote cyclically:

$\succ^{T_1}$	$\succ^{T_2}$	.	.	$\succ^{T_m}$
c_1	c_2	.	.	c_m
c_2	c_3	.	.	c_1
.	.	.	.	.
.	.	.	.	.
c_{m-1}	c_m	.	.	c_{m-2}
c_m	c_1	.	.	c_{m-1}

Notice that under any scoring rule, all candidates have the same score; without loss of generality candidate c_1 is the winner of this election. How many profiles obtained by a single transposition necessarily have a different outcome? An elementary transposition between places $l-1$ and l, where $l \in A_F$, strictly increases a candidate's score, and changes the outcome — given that the promoted candidate is not candidate 1. For every $l \in A_F$, exactly d voters rank candidate c_1 in place l. Hence, there are da_F voters with $a_F - 1$ possible elementary transpositions that change the outcome of the election (the voters that rank candidate c_1 in place $l \in A_F$), and $n - da_F$ voters with a_F such transpositions. It follows that the probability that the outcome changes, under the uniform distribution over instances such that $\succ^V \rightsquigarrow \succ_1^V$, is at least (substituting $n = md$):

$$\frac{da_F(a_F-1) + (dm - da_F)a_F}{dm(m-1)} = \frac{a_F}{m}.$$

In other words, the probability that the outcome *does not change* is at most $\frac{m-a_F}{m}$. As the robustness is defined to be the minimum over all instances, we obtain the desired result. $\square$

We conclude that the Veto and Plurality rules, where $a_F = 1$, are extremely robust. On the other hand, the Borda rule, for which $a_F = m-1$, is very susceptible to failures.

4.2 Copeland

We give an upper bound that relies on an example where the number of voters is even. However, since the number of candidates is not restricted, this example implies that it is not possible to establish a good general lower bound. In addition, as the upper bound is very small, an exact lower bound is of no consequence.

Proposition 7. *Let m be the number of candidates, and let the number of voters n be even. Then $\rho_1^{n,m}$(Copeland) $\leq 1/(m-1)$.*

Proof. Consider the preference profile where for $i = 1, 3, 5, \ldots, n-1$, voters v^i and v^{i+1} vote as follows:

$\succ^i$	$\succ^{i+1}$
c_1	c_m
c_2	c_{m-1}
.	.
.	.
c_m	c_1

Under the above profile, for every two candidates c and c', exactly $n/2$ voters prefer c over c'. Thus, the Copeland score of all candidates is 0, and the winner is some candidate $c \in C$. Any elementary transposition that promotes candidate $c' \neq c$ would raise the score of c' to 1, making c' the new winner. This implies that for every voter, there are at least $m-2$ elementary transpositions that change the outcome of the election, and thus the probability that the outcome does not change is at most $1 - \frac{n(m-2)}{n(m-1)} = \frac{1}{m-1}$. □

4.3 Maximin

Proposition 8. *Let n and m be the number of voters and candidates such that m divides n. Then $\rho_1^{n,m}(\text{Maximin}) \leq 1/(m-1)$.*

Proof. Our adversarial preference profile is identical to the one in the proof of Proposition 6. However, we are going to construct the profile algorithmically, as this is going to aid us in establishing some of the profile's properties. We iteratively expand the list of candidates; initially, it contains only c_1, so each voter's linear preferences are in fact the empty set. In the second stage, we add to the slate the candidate c_2; for $\frac{1}{m}n$ voters, candidate c_2 is ranked at the top (above c_1), but the other $\frac{m-1}{m}n$ voters rank c_2 below c_1. Now, c_3 is added as follows: $\frac{1}{m}n$ voters that ranked c_2 last (i.e., previously voted $c_1 \succ c_2$), now rank c_3 first (i.e., vote $c_3 \succ c_1 \succ c_2$); the other $\frac{m-1}{m}n$ voters rank c_3 immediately below c_2 (e.g., if the ranking was $c_2 \succ c_1$, it is now $c_2 \succ c_3 \succ c_1$). In general, when adding candidate c_j, $\frac{1}{m}n$ voters that ranked c_{j-1} last now rank c_j first,[3] and the rest rank c_j just below c_{j-1}.

For example, for 8 voters and 4 candidates, initially we have: (in each stage j, the $\frac{1}{m}n = 2$ grayed voters are the ones that rank candidate c_j first instead of just under c_{j-1})

$\succ^1$	$\succ^2$	$\succ^3$	$\succ^4$	$\succ^5$	$\succ^6$	$\succ^7$	$\succ^8$
c_1	c_1	c_1	c_1	c_1	c_1	c_1	c_1

In the second stage we have:

$\succ^1$	$\succ^2$	$\succ^3$	$\succ^4$	$\succ^5$	$\succ^6$	$\succ^7$	$\succ^8$
c_2	c_2	c_1	c_1	c_1	c_1	c_1	c_1
c_1	c_1	c_2	c_2	c_2	c_2	c_2	c_2

In the third stage we have:

[3] It is easy to verify that there always are $\frac{1}{m}n$ such voters.

$\succ^1$	$\succ^2$	$\succ^3$	$\succ^4$	$\succ^5$	$\succ^6$	$\succ^7$	$\succ^8$
c_2	c_2	c_3	c_3	c_1	c_1	c_1	c_1
c_3	c_3	c_1	c_1	c_2	c_2	c_2	c_2
c_1	c_1	c_2	c_2	c_3	c_3	c_3	c_3

Ultimately, the preference profile that the algorithm constructs is:

$\succ^1$	$\succ^2$	$\succ^3$	$\succ^4$	$\succ^5$	$\succ^6$	$\succ^7$	$\succ^8$
c_2	c_2	c_3	c_3	c_4	c_4	c_1	c_1
c_3	c_3	c_4	c_4	c_1	c_1	c_2	c_2
c_4	c_4	c_1	c_1	c_2	c_2	c_3	c_3
c_1	c_1	c_2	c_2	c_3	c_3	c_4	c_4

Lemma 9. *In stage j (after candidate c_j is added to the slate), it holds that for every $i < j$, the number of voters that prefer c_i to c_j is $\frac{m-(j-i)}{m}n$.*

Proof. By induction on j. The basis of the induction ($j = 1$) is trivial. Now, assume the claim holds for $j - 1$; we shall prove it for j. Let $i < j$; if $i = j - 1$, notice that c_j is ranked under c_i, except in $\frac{1}{m}n$ cases. In other words, the number of voters that prefer $c_i = c_{j-1}$ to c_j is $\frac{m-1}{m}n$, as desired.

It remains to deal with the case where $i < j - 1$. Recall that c_j is always ranked directly under c_{j-1}, except for $\frac{1}{m}n$ voters that rank c_j first. As for the rest of the voters, c_i is ranked above c_j iff c_i was ranked above c_{j-1} in stage $j-1$. By the induction assumption, we had $\frac{m-((j-1)-i)}{m}n$ ranking c_i above c_{j-1} in stage $j - 1$, and thus the number of voters ranking c_i above c_j is:

$$\frac{m-((j-1)-i)}{m}n - \frac{1}{m}n = \frac{m-(j-i)}{m}n,$$

as desired. □

Lemma 9 implies that candidate c_j's *unique* worst pairwise election is against c_{j-1} for $j > 1$: the number of voters that prefer c_j to c_{j-1} is exactly $\frac{1}{m}n$; notice that this is also true for c_1 versus c_m: only $\frac{1}{m}n$ rank c_1 above c_m. In addition, for $j > 1$, clearly c_j is ranked just under c_{j-1} by $\frac{m-1}{m}n$ voters — but this, too, is also true for c_1 versus c_m; indeed, c_1 is ranked just under c_m by all $\frac{m-1}{m}n$ voters that do not rank c_1 first.

So, the candidates are all tied with respect to their maximin scores, and each candidate c_j is ranked just below its "worst pairwise" candidate by all voters that do not rank c_j first. Therefore, any elementary transposition that promotes a candidate that is not the current winner of the election must change the outcome of the election. As before, we have that the probability of the outcome changing as a result of a single transposition, under our adversarial preference profile, is at least $\frac{m-2}{m-1}$, and thus robustness of this preference profile is at most $\frac{1}{m-1}$. □

4.4 Bucklin

Proposition 10. *$\rho_1^{n,m}$(Bucklin) $\geq \frac{m-2}{m-1}$ for any values of the number of voters n and the number of candidates m.*

Proof. Consider a preference profile $\succ^V$, and assume that the winner c_j of the election satisfies: $l_0 = \min_l B(j,l) > n/2$. We argue that any elementary transposition that switches the candidates in places l and $l-1$, for $l \neq l_0, l_0+1$, cannot change the outcome of the election. Indeed, we consider two cases:

Case 1: $l > l_0 + 1$. In this case, if some candidate $c_k \neq c_j$ is promoted, the switch increases $B(k, l-1)$ — but this is irrelevant to the outcome of the election, since $B(k, l_1)$ remains unchanged for $l_1 \leq l_0$.

Case 2: $l < l_0$. If candidate c_k is promoted, this might increase $B(k, l_0 - 2)$. However, the value of $B(k, l_0 - 2)$ after the switch took place is bounded from above by the value of $B(k, l_0 - 1)$ before the switch. We know that $B(k, l_0 - 1) \leq n/2$ before the switch — so this transposition is not going to affect the value of $\min_k(B(k,l) > n/2)$.

If so, it remains to consider the case where $l = l_0$ or $l = l_0 + 1$. When $l = l_0$, promoting $\pi_{l_0}(\succ^i)$ may affect the outcome only if $\pi_{l_0}(\succ^i) \neq c_j$, where c_j is the winner of the election. However, when $l = l_0 + 1$, promoting $\pi_{l_0+1}(\succ^i)$ and demoting $\pi_{l_0}(\succ^i)$ might affect the outcome only if $\pi_{l_0}(\succ^i) = c_j$. Otherwise, if $\pi_{l_0+1}(\succ^i) = c_k \neq c_j$, then $B(k, l_0)$ might be affected, but since c_j already has a majority of voters ranking it in the top l_0 places, the outcome of the election is indifferent to this perturbation.

As these two last subcases are mutually exclusive, it follows that for every voter there is at most one transposition that may affect the outcome of the election. Thus $\rho_1^{n,m}(\text{Bucklin}) \geq \frac{m-2}{m-1}$. □

4.5 Plurality with Runoff

The rules we have discussed in the previous subsections all have in common some concept of score. Since Plurality with Runoff is a bit different, we require an additional assumption regarding tie-breaking. Consider a situation where, say, c_{j_1} and c_{j_2} survive the first round, and exactly half the voters prefer c_{j_1} to c_{j_2}, but c_{j_1} is the winner of the election. We assume that if a fault makes c_{j_1} and c_{j_3} survive the first round, and again these two candidates are tied in the second round, then c_{j_1} *loses* the election. This assumption is consistent with our worst-case analysis throughout.

Proposition 11. *For all values of n and m, $\rho_1^{n,m}$(Plurality with Runoff) $\geq \frac{m-5/2}{m-1}$.*

Proof. Consider some preference profile $\succ^V$, and assume w.l.o.g. that candidates c_1 and c_2 survive the first round, and c_1 wins the election. Only two types of elementary transposition can potentially affect the outcome of the election. The first is promoting the candidate $\pi_2(\succ^i)$ for some i, i.e., making this candidate voter v^i's favorite — this might affect the list of candidates that

Function	Lower Bound	Upper Bound
Scoring	$\frac{m-1-a_F}{m-1}$	$\frac{m-a_F}{m}$
Copeland	0	$\frac{1}{m-1}$
Maximin	0	$\frac{1}{m-1}$
Bucklin	$\frac{m-2}{m-1}$	1
Plurality w. Runoff	$\frac{m-5/2}{m-1}$	$\frac{m-5/2}{m-1} + \frac{5/2}{m(m-1)}$

Table 1: Upper and lower bounds on the 1-robustness of several prominent voting rules.

are eliminated in the first round. A second transposition which might have an effect is one that promotes candidate c_2 and demotes c_1 — this might change the outcome of the second round, but only if exactly half the voters prefer c_1 to c_2 in $\succ^V$ (it cannot be the case that more voters prefer c_2, as then c_2 would have prevailed in the second round). To conclude, at most $n/2$ voters have two transpositions that may affect the outcome, and at least $n/2$ voters have only one. We have that

$$\rho(F, \succ^V) \geq \frac{n(m-1) - (n/2 \cdot 1 + n/2 \cdot 2)}{n(m-1)} = \frac{m-5/2}{m-1}.$$

□

Proposition 12. $\rho_1^{2m,m}$*(Plurality with Runoff)* $\leq \frac{m-5/2}{m-1} + \frac{5/2}{m(m-1)}$.

Proof. Omitted due to space constraints. □

5 Discussion

We have defined the k-robustness of a voting rule as the worst-case probability that k independent switches in the preferences of voters change the outcome of the election. We have shown that high 1-robustness implies high k-robustness, at least for a constant k. Inversely, low 1-robustness clearly suggests that the rule is not robust in general. Accordingly, we have presented bounds on the 1-robustness of different voting rules; these bounds are summarized in Table 5.

We intend our results to be used as a tool for designers of multiagent systems. When dealing with noisy environments, successful aggregation of preferences can only be expected when a robust voting rule is applied. In particular, among the prominent voting rules, our results imply that Plurality, Plurality with Runoff, Veto, and Bucklin are robust to faults, whereas Borda, Copeland, and Maximin are susceptible to faults.

The model of errors we have introduced is a theoretical one, but we have also shown it is grounded in a reasonable representation of preferences. Nevertheless, future work should include an investigation of different error models.

In addition, our analysis was worst-case — an approach which leads to the conclusion that when the number of errors is large, voting rules are bound to fail. It would be interesting to complement our results with an asymptotic *average-case* analysis.

References

[1] J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241, 1989.

[2] S. J. Brams and P. C. Fishburn. Voting procedures. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, chapter 4. North-Holland, 2002.

[3] V. Conitzer and T. Sandholm. Complexity of manipulating elections with few candidates. In *Proceedings of the National Conference on Artificial Intelligence*, pages 314–319, 2002.

[4] J. D. Dixon and B. Mortimer. *Permutation Groups*. Springer, 1996.

[5] S. Ghosh, M. Mundhe, K. Hernandez, and S. Sen. Voting for movies: the anatomy of a recommender system. In *Proceedings of the Third Annual Conference on Autonomous Agents*, pages 434–435, 1999.

[6] A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587–602, 1973.

[7] T. Haynes, S. Sen, N. Arora, and R. Nadella. An automated meeting scheduling system that utilizes user preferences. In *Proceedings of the First International Conference on Autonomous Agents*, pages 308–315, 1997.

[8] G. Kalai. Noise sensitivity and chaos in social choice theory. Preprint, 2005.

[9] M. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.

Ariel D. Procaccia and Jeffrey S. Rosenschein
School of Engineering and Computer Science
The Hebrew University of Jerusalem
Givat Ram, Jerusalem 91904, Israel
Email: `{arielpro,jeff}@cs.huji.ac.il`

Gal A. Kaminka
Computer Science Department
Bar Ilan University
Ramat Gan 52900, Israel
Email: `galk@cs.biu.ac.il`