Analysing Irresolute Multiwinner Voting Rules with Approval Ballots via SAT Solving

Kluiving, B.; de Vries, A.; Vrijbergen, P.; Boixel, A.; Endriss, U.

DOI
10.3233/FAIA200085

Publication date
2020

Document Version
Final published version

Published in
ECAI 2020

License
CC BY-NC

Link to publication

Citation for published version (APA):
Kluiving, B., de Vries, A., Vrijbergen, P., Boixel, A., & Endriss, U. (2020). Analysing Irresolute Multiwinner Voting Rules with Approval Ballots via SAT Solving. In G. De Giacomo, A. Catala, B. Dilkina, M. Milano, S. Barro, A. Bugarín, & J. Lang (Eds.), *ECAI 2020: 24th European Conference on Artificial Intelligence : 29 August-8 September 2020, Santiago de Compostela, Spain, including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020) : proceedings* (pp. 131-138). (Frontiers in Artificial Intelligence and Applications; Vol. 325). IOS Press. https://doi.org/10.3233/FAIA200085

ECAI 2020
G.D. Giacomo et al. (Eds.)

doi:10.3233/FAIA200085

Analysing Irresolute Multiwinner Voting Rules with Approval Ballots via SAT Solving

Boas Kluiving and **Adriaan de Vries** and **Pepijn Vrijbergen** and **Arthur Boixel** and **Ulle Endriss**[1]

Abstract. Suppose you want to design a voting rule that can be used to elect a committee or parliament by asking each voter to approve of a subset of the candidates standing. There are several properties you may want that rule to satisfy. First, voters should enjoy some form of proportional representation. Second, voters should not have an incentive to misrepresent their preferences. Third, outcomes should be Pareto efficient. We show that it is impossible to design a voting rule that satisfies all three properties. We also explore what possibilities there are when we weaken our requirements. Of special interest is the methodology we use, as a significant part of the proof is outsourced to a SAT solver. While prior work has considered similar questions for the special case of resolute voting rules, which do not allow for ties between outcomes, we focus on the fact that, in practice, most voting rules allow for the possibility of such ties.

1 INTRODUCTION

When several people need to aggregate their possibly diverging views to arrive at an acceptable compromise, then they may wish to delegate this task to a voting rule [1, 9]. Whether a given voting rule is an adequate tool for this purpose in a given context depends on the formal properties it satisfies. Properties of interest might relate to the operational simplicity of a rule, to its fairness, or to the level of protection it can provide against undesirable strategic behaviour. In this paper, we continue a recent line of research in AI [29, 8, 10, 11, 25] aimed at improving our understanding of the opportunities available when designing normatively appealing voting rules by encoding the problem of deciding whether there exists a rule that satisfies a certain combination of properties as a query to a SAT solver [7]. We specifically focus on so-called *multiwinner voting rules* with *approval ballots*. Under such a rule each voter is asked to indicate which of the candidates available for election she approves of. The rule then elects a fixed number of those candidates.

While the bulk of past research on voting has focused on the case of single-winner voting rules with ranked ballots, where voters report rankings of the candidates and we need to elect a single such candidate, and while many of the seminal results in the field relate to that particular model of voting [20, 28, 13], there recently has been much renewed interest in multiwinner voting [15] and the normative properties of multiwinner voting rules with approval ballots in particular [2, 27, 25]. Indeed, many real-world decision making scenarios are naturally modelled as multiwinner elections. Examples include electing a parliament or any other kind of committee, but also, for instance, deciding which of a number of projects to spend public money on. Asking for approvals rather than rankings is attractive due to its convenience for the voters. When designing a voting rule, we will often be interested in the following properties:

- *Proportionality:* Sufficiently large subgroups of voters with sufficiently coherent views should get adequate representation. For instance, when electing a committee of k members and one in k voters votes only for candidate c, then c should get elected.
- *Strategyproofness:* Voters should have no incentive to misrepresent their views. For instance, if you find candidates a and b acceptable, then you should not feel that voting for just one of them will increase your chances of getting at least one of them elected.
- *Efficiency:* Outcomes should be Pareto efficient. In particular, swapping an elected with an unelected candidate should never make an outcome better for some but no worse for the other voters.

In recent work—using the SAT-based approach we also adopt here—Peters [25, 26] was able to show that, for certain natural ways of making these requirements precise, it is mathematically *impossible* to design a multiwinner voting rule with approval ballots that meets all of them. This is an important finding, complementing known results for the single-winner model of voting such as the seminal Gibbard-Satterthwaite Theorem [20, 28]. It clearly maps out the limitations we are subject to when designing a voting rule. However, Peters' result only applies to *resolute* rules. These are deterministic rules that must return a *single* winning set of candidates for every possible profile. This is a demanding requirement that, in fact, is not met by any of the standard voting rules that have been proposed in the literature [15]: any reasonable rule will sometimes have to report a tie. This raises the question whether dropping the resoluteness requirement will lead to more favourable results. Unfortunately, analysing (potentially) *irresolute* rules is significantly more complex than analysing resolute rules. Indeed, for the standard model of single-winner voting the most influential attempt at addressing the same shortcoming in the Gibbard-Satterthwaite Theorem, the Duggan-Schwartz Theorem [13], was published over two decades after that original result. One of the difficulties involved is that there are multiple ways in which to formalise strategyproofness for irresolute rules and there is still ongoing research on the topic today [10, 11].

Contribution. In this paper, we prove a new impossibility theorem for multiwinner voting rules with approval ballots that applies to resolute and irresolute rules alike. A significant part of this proof has been derived automatically, with the help of a SAT solver. In broad terms, we show that it is impossible to design a rule that is proportional, immune to strategic manipulation, and Pareto efficient, i.e., we show that dropping the resoluteness requirement does not provide a way out of the impossibility established by Peters. In fact, there are a number of different ways in which one could reasonably formalise these normative principles. We prove our theorem for one

[1] Institute for Logic, Language and Computation (ILLC), University of Amsterdam, The Netherlands. Contact: ulle.endriss@uva.nl.

specific formalisation that is particularly strong and that appears particularly natural to us. However, we also acknowledge that alternative formalisations may be of interest as well. This, arguably, is where the benefits of the SAT-based approach we use are most evident: having automated a core part of the proof makes it easy to experiment with a large number of variants of the same kind of result. We explore this opportunity and report and comment on several such variants.

Related work. The approach of proving impossibility theorems in social choice theory with the help of a SAT solver was initiated by Tang and Lin [29]. The fundamental idea at the core of this approach is that, if we fix the relevant parameters of the social choice scenario we are interested in (such as the number of voters and the number of alternatives in a voting scenario), then we can represent the requirements imposed by the axioms we want to study as formulas in propositional logic. We can then use a SAT solver to check whether the conjunction of these formulas is satisfiable—if it is not, then we have found an impossibility theorem for that particular choice of parameters. Intuitively, such an impossibility will often generalise to all settings with larger choices of parameters and it is often possible (though sometimes difficult) to prove this intuition formally correct. Tang and Lin analysed social welfare functions and resolute voting rules in this manner, and the SAT-based approach has also been applied to irresolute voting rules [8, 10, 11], to resolute multiwinner voting rules [25], to preference extension schemes [18], and to matching mechanisms [14]. For an introduction to the approach, we refer to the recent survey article by Geist and Peters [19]. The broader interest of the SAT-based approach for economic theory has been discussed by Chatterjee and Sen [12].

Paper outline. The remainder of this paper is organised as follows. We first formalise the model of voting and the core axioms we are going to be working with in Section 2. We then state and prove our main theorem in Section 3, before discussing a range of interesting variants of this result in Section 4 and concluding in Section 5.

Code. The code written to generate the SAT instances we used to obtain the results discussed in this paper is available online [23].

2 THE MODEL

In this section we recall the familiar model of multiwinner voting with approval ballots [22, 15], fixing in particular the assumptions we are going to make regarding the preferences of voters. We also define three axioms encoding normative properties that arguably any multiwinner voting rule should enjoy. (In Section 4 we are going to discuss several alternatives to these assumptions and axioms.)

Throughout, for any given set S and $k \in \mathbb{N}$, we use 2^S to denote the set of all subsets of S (i.e., the powerset of S) and we use $[S]^k$ to denote the set of all subsets of S with exactly k elements each.

2.1 Multiwinner Voting with Approval Ballots

Fix a finite set $N = \{1, \ldots, n\}$ of *voters*, a finite set C of *candidates* with $m = |C|$, and a natural number $k < m$. We are asked to choose a *committee* $X \subseteq C$ of size $|X| = k$ on the basis of a *profile* $\boldsymbol{A} = (A_1, \ldots, A_n)$ submitted by the voters, where the *ballot* $A_i \subseteq C$ of a given voter $i \in N$ is the set of candidates she *approves* of. We want to delegate this choice to a *voting rule* F. As most reasonable rules will sometimes be unable to differentiate between two or more equally good committees, a voting rule formally is a function F mapping any given profile of approval ballots to a nonempty set of committees:

$$F : (2^C)^n \to 2^{[C]^k} \setminus \{\emptyset\}$$

Thus, a voting rule might be *irresolute*. We say that F is *resolute* in case $|F(\boldsymbol{A})| = 1$ for every profile $\boldsymbol{A} \in (2^C)^n$.

A basic example for a voting rule is *approval voting* (AV), which amounts to electing the candidates that receive the most approvals:

$$F_{\text{AV}}(\boldsymbol{A}) \quad = \quad \operatorname*{argmax}_{X \in [C]^k} \sum_{i \in N} |A_i \cap X|$$

Another example is a rule known as *proportional approval voting* (PAV), going back to a proposal by Thiele [31] from 1895. Under this rule, voter i assigns a score of $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{p}$ to committee X, with p being the number of candidates on which X agrees with A_i. We then elect the committees that maximise the sum of these scores:

$$F_{\text{PAV}}(\boldsymbol{A}) \quad = \quad \operatorname*{argmax}_{X \in [C]^k} \sum_{i \in N} \sum_{\ell=1}^{|A_i \cap X|} \frac{1}{\ell}$$

Observe that neither F_{AV} nor F_{PAV} are resolute, and the same is true for basically every other natural definition of a voting rule—unless we specifically combine it with a tie-breaking rule. This is why, in this paper, we explicitly focus on rules that might be irresolute.

2.2 Inducing Preferences from Approval Sets

When pondering the merits and demerits of different voting rules, we sometimes need to refer to the preferences of the voters over the committees such a rule might elect. But note that the model of multiwinner voting we use here does not come with an inherent notion of preference: voters express their views by specifying which *candidates* they *approve* of, not which *committees* they *prefer*. To be able to infer such preferences, we need to make additional assumptions on how these two concepts relate to each other.

Consider a voter $i \in N$ whose truthfully held approval set is $A_i \subseteq C$. We are going to assume that the associated (cardinality-based) *preference order* of voter i is the unique weak order $\succcurlyeq_{A_i}$ on $[C]^k$ satisfying the following condition for all committees $X, X' \in [C]^k$:

$$X \succcurlyeq_{A_i} X' \;\text{ if and only if }\; |A_i \cap X| \geqslant |A_i \cap X'|$$

Thus, we assume that a voter will rank any two committees in terms of the *number* of candidates they include that she approves of. (In Section 4.4 we are also going to consider a weaker assumption.) We write $X \succ_{A_i} X'$ in case $X \succcurlyeq_{A_i} X'$ and $X' \not\succcurlyeq_{A_i} X$, meaning that voter i (with approval set A_i) *strictly* prefers X to X'.

Recall that voting rules return *sets* of committees. Thus, when reasoning about the preferences of voters over different outcomes, we must compare different sets of committees, not just different committees. A host of different *preference extensions* have been defined in the literature to model this process [4]. Here we are going to make what is probably the weakest of assumptions one can reasonably make in this context, namely that a voter weakly prefers a nonempty set $\mathcal{X}$ of committees over another nonempty set $\mathcal{X}'$ if she weakly prefers *every* committee in $\mathcal{X}$ to *every* committee in $\mathcal{X}'$. This is the so-called *Kelly extension* [21]. We may interpret this definition as the voter in question being *cautious:* she will commit to (weakly) preferring $\mathcal{X}$ to $\mathcal{X}'$ only in this most clear-cut of cases.

Formally, if $\succcurlyeq_{A_i}$ reflects the preferences of voter i over committees, then her preference relation $\succcurlyeq^{\text{CAU}}_{A_i}$ over nonempty sets of committees is the unique preorder satisfying the following condition:

$$\mathcal{X} \succcurlyeq^{\text{CAU}}_{A_i} \mathcal{X}' \;\text{ if and only if }\; (\forall X \in \mathcal{X})\,(\forall X' \in \mathcal{X}')\; X \succcurlyeq_{A_i} X'$$

Observe that $\succcurlyeq^{\text{CAU}}_{A_i}$ is not a complete relation: it will leave many pairs of sets of committees unranked. We write $\mathcal{X} \succ^{\text{CAU}}_{A_i} \mathcal{X}'$ in case voter i strictly prefers $\mathcal{X}$ to $\mathcal{X}'$, i.e., in case $\mathcal{X} \succcurlyeq^{\text{CAU}}_{A_i} \mathcal{X}'$ and $\mathcal{X}' \not\succcurlyeq^{\text{CAU}}_{A_i} \mathcal{X}$.

2.3 Axioms: Normative Properties of Voting Rules

What makes for a good voting rule? To narrow down the field of contenders, let us define a few axioms, i.e., intuitively appealing properties of rules. The first one encodes a basic efficiency requirement: a good voting rule should not have committee X' amongst its winning committees when there is another committee X that is weakly preferred by all (and strictly preferred by at least one) of the voters.

Axiom 1. *F is **Pareto efficient** if, for any profile $\boldsymbol{A}$ and any two committees $X, X' \in [C]^k$ with $X \succcurlyeq_{A_i} X'$ for all voters $i \in N$ and $X \succ_{A_i} X'$ for at least one of them, it is the case that $X' \notin F(\boldsymbol{A})$.*

The second axiom encodes the idea that we do not want to incentivise voters to manipulate elections by misrepresenting the set of candidates they truthfully approve of. To define it, we require some additional terminology: two profiles $\boldsymbol{A}$ and $\boldsymbol{A}'$ are *i-variants* of each other (with $i \in N$ being one of the voters) if $A_j = A'_j$ for all voters $j \in N \setminus \{i\}$. We write $\boldsymbol{A} =_{-i} \boldsymbol{A}'$ in this case. Thus, in the truthful profile $\boldsymbol{A}$ voter i can move to exactly those (untruthful) profiles $\boldsymbol{A}'$ for which $\boldsymbol{A} =_{-i} \boldsymbol{A}'$.

Axiom 2. *F is **immune to manipulation by cautious voters** if, for any voter $i \in N$ and any two profiles $\boldsymbol{A}$ and $\boldsymbol{A}'$ that are i-variants of each other, it is the case that $F(\boldsymbol{A}') \not\succ^{\text{CAU}}_{A_i} F(\boldsymbol{A})$.*

We also say that F is (*Kelly*)*-strategyproof*. Note that in case $F(\boldsymbol{A})$ and $F(\boldsymbol{A}')$ are incomparable according to the relation $\succcurlyeq^{\text{CAU}}_{A_i}$, the formulation of our axiom presupposes that no manipulation will take place. Thus, given that the Kelly extension is highly incomplete, this is a very undemanding strategyproofness axiom. (We are going to consider alternative forms of preference extension, and thus alternative notions of strategyproofness, in Section 4.2.)

The third axiom encodes the idea that voters should enjoy some form of proportional representation. Several such axioms have been put forward in the literature [2, 27]. The basic idea is that any sufficiently large group of voters with sufficiently similar views regarding the candidates should get appropriate representation in the committee elected. Following Peters [25], we define a very undemanding proportionality axiom that is implied by every other formulation of proportionality in the literature we are aware of. This axiom imposes requirements on the outcome of a rule F only for the special case of so-called *party-list profiles*, which are profiles $\boldsymbol{A}$ with $A_i = A_j$ or $A_i \cap A_j = \emptyset$ for all voters $i, j \in N$. Such profiles arise when the set of candidates C can be partitioned into parties and each voter approves of all members of some party and only those.

Axiom 3. *F is **minimally proportional** if, for any candidate $c \in C$ and any party-list profile $\boldsymbol{A}$ with $|\{i \in N : A_i = \{c\}\}| \geqslant \frac{n}{k}$, it is the case that $c \in X$ for all committees $X \in F(\boldsymbol{A})$.*

Thus, if $\frac{n}{k}$ of the n voters approve of c alone and we need to elect a committee of size k, then c should get elected with certainty. In fact, the axiom is even weaker than that and requires this conclusion to be drawn only in case the profile in question is a party-list profile.[2]

It will sometimes be useful to be able to assume a slightly stronger variant of the proportionality axiom that applies not only to party-list profiles with one party with a single member c, but rather more generally to profiles in which some voters approve only of c and the others only approve of candidates other than c. The following lemma shows that we can invoke the axiom on this broader range of profiles as well—if we can assume strategyproofness. Our proof closely follows Peters [25], who establishes a similar result for resolute rules.

Lemma 1. *For any voting rule F that is minimally proportional and immune to manipulation by cautious voters, for any candidate $c \in C$, and for any profile $\boldsymbol{A}$ with $|\{i : A_i = \{c\}\}| \geqslant \frac{n}{k}$ and $\{i : c \in A_i\} = \{i : A_i = \{c\}\}$, it is the case that $c \in X$ for all committees $X \in F(\boldsymbol{A})$.*

Proof. Let F, c, and $\boldsymbol{A}$ be as specified in the claim. For the sake of contradiction, suppose that $c \notin X$ for some committee $X \in F(\boldsymbol{A})$. Given c and $\boldsymbol{A}$, let us define a new party-list profile $\boldsymbol{A}'$:

$$A'_i \quad = \quad \begin{cases} \{c\} & \text{if } A_i = \{c\} \\ C \setminus \{c\} & \text{otherwise} \end{cases}$$

Then, due to F being minimally proportional, we get that $c \in X$ for all $X \in F(\boldsymbol{A}')$. Now consider a cautious voter i with true approval set $A'_i = C \setminus \{c\}$ and corresponding preference order $\succcurlyeq^{\text{CAU}}_{A'_i}$. She agrees on exactly $k-1$ candidates (all but c) with all committees $X \in F(\boldsymbol{A}')$ and on *at least* $k-1$ candidates with all conceivable committees $X \in [C]^k$, so she cannot do worse than $F(\boldsymbol{A}')$. On the other hand, by our assumption, she agrees on all k candidates with some committee $X \in F(\boldsymbol{A})$. Thus, $F(\boldsymbol{A}) \succ^{\text{CAU}}_{A'_i} F(\boldsymbol{A}')$.

Note that this is the case for *all* voters i with $A'_i = C \setminus \{c\}$. Now imagine a sequence of profiles, starting with $\boldsymbol{A}'$, with the voters i with $A'_i = C \setminus \{c\}$ switching to A_i one by one, ending in $\boldsymbol{A}$. Then there must be a first point at which some committee X with $c \notin X$ is amongst the winners. Then, if we think of the profile just before that point as the truthful profile, we have found a situation in which there exists a cautious voter that can manipulate the rule in her favour. Thus, we have found a contradiction to one of our assumptions. □

3 THE IMPOSSIBILITY THEOREM

In this section we state and prove our main result, which shows that it is impossible to simultaneously satisfy all of the normative desiderata for a voting rule we have formulated earlier. This impossibility theorem applies under certain assumptions on n, m, and k:

Theorem 1. *Let $k \geqslant 3$, let $m > k$, and let n be a multiple of k. Then no voting rule for n voters to elect committees of size k from a pool of m candidates can be simultaneously minimally proportional, Pareto efficient, and immune to manipulation by cautious voters.*

To simplify exposition of our proof of Theorem 1, let us call F a *good voting rule for* (n, m, k) if it is a voting rule for n voters to elect committees of size k from a pool of m candidates and if that rule is minimally proportional, Pareto efficient, and immune to manipulation by cautious voters. Thus, the claim made by Theorem 1 is that there exists no good voting rule for any triple (n, m, k) with $k \geqslant 3$, $m > k$, and n being a multiple of k.

We are going to prove this claim using an inductive argument. We first provide a proof of the *base case*, which is the case of $(n, m, k) = (3, 4, 3)$. This part of the proof we have been able to generate in an automated manner with the help of a SAT solver. Section 3.1 is devoted to discussing this approach and presenting the resulting proof of the base case in a human-readable form. Then, in Section 3.2, we (manually) prove three lemmas that permit us to generalise our result for the base case to Theorem 1 as stated above. Finally, in Section 3.3 we demonstrate that each of the axioms featuring in Theorem 1 is *necessary* to obtain the impossibility stated. We also made use of a SAT solver for this task.

[2] The restriction to party-list profiles ensures that minimal proportionality is implied also by similar axioms proposed in the literature on *apportionment*, which deals with voters voting for parties rather than candidates [3].

Peters [25, 26] proved a similar result for the special case of resolute voting rules,[3] and our proof proceeds in a similar manner as his original proof. We are going to further comment on the connections between his result and our Theorem 1 in Section 4.

3.1 Automated Proof of the Base Case

For fixed parameters (n, m, k), we can model voting rules in propositional logic using variables of the form $p_{\boldsymbol{A},X}$, one for each profile $\boldsymbol{A} \in (2^C)^n$ and committee $X \in [C]^k$. The idea is that $p_{\boldsymbol{A},X}$ is true if and only if the voting rule we are modelling will elect committee X in profile $\boldsymbol{A}$. Recall that there may be more than one winning committee in any given profile. So every assignment of truth values to variables corresponds to a voting rule—except for assignments that, for some profile $\boldsymbol{A}$, make $p_{\boldsymbol{A},X}$ false for every committee X. Such a truth assignment does not correspond to a voting rule, because such a "rule" would return an empty set of winning committees for that profile. Thus, for a given choice of (n, m, k), there is a bijection between the set of all voting rules and the set of all truth assignments that are models of the following formula:

$$\varphi_{\text{AT-LEAST-ONE}} \quad = \quad \bigwedge_{\boldsymbol{A} \in (2^C)^n} \ \bigvee_{X \in [C]^k} \ p_{\boldsymbol{A},X}$$

We can now encode our axioms in a similar fashion, further narrowing down the set of voting rules. Expressing Pareto efficiency is easy:

$$\varphi_{\text{PARETO}} \quad = \quad \bigwedge_{\boldsymbol{A} \in (2^C)^n} \ \bigwedge_{\substack{X \in [C]^k \\ X \text{ dominated in } \boldsymbol{A}}} \neg p_{\boldsymbol{A},X}$$

In fact, for a fixed choice of (n, m, k), *every* possible axiom can be expressed in propositional logic. This follows from the fact that an axiom is just a way of singling out some subset of the set of all voting rules, together with the fact that for every possible set of truth assignments there exists a propositional formula that is true for exactly those assignments. Nevertheless, in practice, some axioms are more naturally and/or more compactly expressed than others.

Expressing Kelly-strategyproofness is significantly more demanding than expressing Pareto efficiency:

$$\varphi_{\text{SP}} \quad = \quad \bigwedge_{i \in N} \ \bigwedge_{\boldsymbol{A} =_{-i} \boldsymbol{A}'} \ \bigwedge_{\mathcal{X}' \succ^{\text{CAU}}_{A_i} \mathcal{X}} \text{DIFF}(\boldsymbol{A}, \mathcal{X}) \ \vee \ \text{DIFF}(\boldsymbol{A}', \mathcal{X}')$$

$$\text{where DIFF}(\boldsymbol{A}, \mathcal{X}) = \bigvee_{X \in \mathcal{X}} \neg p_{\boldsymbol{A},X} \ \vee \bigvee_{X \in [C]^k \setminus \mathcal{X}} p_{\boldsymbol{A},X}$$

Here $\boldsymbol{A}$ and $\boldsymbol{A}'$ are understood to be ranging over all profiles that are i-variants of each other, and $\mathcal{X}$ and $\mathcal{X}'$ are understood to be ranging over all nonempty sets of k-committees for which voter i would rank $\mathcal{X}'$ strictly above $\mathcal{X}$ in case she is a cautious voter. Observe how $\text{DIFF}(\boldsymbol{A}, \mathcal{X})$ encodes the requirement that the set of winning committees for profile $\boldsymbol{A}$ must be different from the set $\mathcal{X}$. So φ_{SP} says that, if voter i prefers outcome $\mathcal{X}'$ to $\mathcal{X}$, then either $\mathcal{X}$ cannot be the truthful outcome (in profile $\boldsymbol{A}$) or $\mathcal{X}'$ cannot be the outcome for any profile $\boldsymbol{A}'$ that i can reach from the truthful profile $\boldsymbol{A}$.

Finally, expressing the axiom of minimal proportionality, once again, is relatively straightforward:

$$\varphi_{\text{PROP}} \quad = \quad \bigwedge_{c \in C} \ \bigwedge_{\substack{\boldsymbol{A} \in (2^C)^n \\ \text{s.t. } \boldsymbol{A} \text{ is party-list and} \\ |\{i : A_i = \{c\}\}| \geqslant \frac{n}{k}}} \ \bigwedge_{\substack{X \in [C]^k \\ \text{s.t. } c \notin X}} \neg p_{\boldsymbol{A},X}$$

[3] The formulation of Peters' result in his original paper on the topic [25] has a small mistake (it is missing a counterpart to our Pareto efficiency axiom). For a full and correct statement, we refer to his dissertation [26].

Now, if we fix n and k as well as C (and thus m), then the conjunction of our four formulas is satisfiable if and only if there exists a good voting rule for this choice of parameters.

For $n = 3$ and $m = 4$ there are $(2^4)^3 = 4096$ possible profiles, and for $m = 4$ and $k = 3$ there are 4 possible committees. So we require $4096 \times 4 = 16384$ propositional variables to express the formulas corresponding to our base case of $(3, 4, 3)$. The conjunction of our four formulas is a formula in CNF. It actually is a conjunction of close to 11 million clauses (over 99% of them are used to express the strategyproofness axiom). We have written a Python script that, on a mid-range laptop, can generate these clauses in around 20 minutes and then save them in a text file, using a standard file format for CNF formulas (the so-called DIMACS format). We can then use a state-of-the-art SAT solver, such as LINGELING [6], to determine whether our formula is satisfiable. Running LINGELING on the formula for $(3, 4, 3)$ returns an answer in under 10 seconds: the formula is unsatisfiable. Thus, there exists no good voting rule for $(3, 4, 3)$.

Of course, some may find this rather unsatisfactory a proof. What if there is a mistake in the script generating the CNF formula? And what if the SAT solver has a bug? Fortunately, SAT solving technology also provides us with a way out here. We can use a tool for extracting a *minimal unsatisfiable subset* (MUS) of the set of clauses corresponding to our formula, which allows us to pinpoint the precise source of the unsatisfiability of the formula and thus the impossibility of designing a good voting rule. Applying the state-of-the-art tool MUSER2 [5] to our formula yields, in around 45 seconds, an MUS of 120 clauses (referring to 16 different profiles). This would, in principle, allow us to manually verify the impossibility. However, as we are going to see next, a smart combination of analytic and computer-generated insights gives rise to a much simpler proof.

Lemma 2. *There exists no good voting rule for* $(3, 4, 3)$.

Proof. By Lemma 1, minimal proportionality and immunity to manipulation by cautious voters together imply a subtly stronger form of the proportionality axiom, which we can encode using the same approach as before. If we replace φ_{PROP} by this moderately longer formula and apply MUSER2 again, we obtain an MUS of just 14 clauses that make reference to just three distinct profiles:

$$\begin{array}{rcl} \boldsymbol{A} & = & \{\,\{c,d\}, \{a,c,d\}, \{b,c,d\}\,\} \\ \boldsymbol{A}' & = & \{\,\{c,d\}, \{a\}, \qquad \{b,c,d\}\,\} \\ \boldsymbol{A}'' & = & \{\,\{c,d\}, \{a,c,d\}, \{b\} \qquad\,\} \end{array}$$

It is easy to extract a human-readable proof from the MUS found. For the sake of contradiction, suppose F is a voting rule for $N = \{1, 2, 3\}$, $C = \{a, b, c, d\}$, and $k = 3$ that is minimally proportional, Pareto efficient, and immune to manipulation by cautious voters. Due to Pareto efficiency, c and d must be part of every winning committee for each of the three profiles above. Furthermore, due to Lemma 1, a must get elected in profile $\boldsymbol{A}'$ and b must get elected in profile $\boldsymbol{A}''$. Thus, we have $F(\boldsymbol{A}') = \{\{a, c, d\}\}$, $F(\boldsymbol{A}'') = \{\{b, c, d\}\}$, and $F(\boldsymbol{A}) \subseteq \{\{a, c, d\}, \{b, c, d\}\}$. Now consider voter 2 in profile $\boldsymbol{A}$ and suppose she is cautious. Due to F being immune to manipulation by this voter in this profile, we cannot have $\{b, c, d\} \in F(\boldsymbol{A})$, as she then would have an incentive to deviate to $\boldsymbol{A}'$. By analogous reasoning about voter 3 contemplating deviating to $\boldsymbol{A}''$ from $\boldsymbol{A}$ we also cannot have $\{a, c, d\} \in F(\boldsymbol{A})$. Hence, $F(\boldsymbol{A}) = \emptyset$, in contradiction to the requirements for a well-defined voting rule. □

3.2 Inductive Lemmas

We now prove three lemmas that each demonstrate that the existence of a good voting rule for a given set of parameters implies the existence of a voting rule for a different set of parameters that, in a certain sense, are "smaller". These technical results confirm a basic intuition: designing a good voting rule tends to be easier for fewer voters, fewer candidates, and smaller target sizes of the committees to be elected. Our first lemma concerns the number of candidates.

Lemma 3. *If there exists a good voting rule for $(n, m+1, k)$ with $m > k$, then also for (n, m, k).*

Proof. Suppose we are given a good voting rule F defined for a set N of n voters and a pool $C \cup \{c^\star\}$ of $m+1$ candidates that can be used to elect committees of size k. Then we can design a new rule F' for the reduced pool C of candidates as follows:

$$F'(\boldsymbol{A}) \;=\; \begin{cases} F(\boldsymbol{A}) & \text{if } |\bigcup_{i \in N} A_i| \geqslant k \\ \{X \in [C]^k \mid \bigcup_{i \in N} A_i \subseteq X\} & \text{otherwise} \end{cases}$$

Observe that F' is well-defined: even in the first case, it will never return a committee that includes $c^\star$. This follows from the Pareto efficiency of F and the fact that $c^\star \notin A_i$ for all voters $i \in N$.

It is now easy to verify that F' is both minimally proportional and Pareto efficient. So it remains for us to show that F' is also strategyproof for cautious voters. Given that F has this property, there certainly can never be a successful manipulation taking us from one profile with k or more approved candidates to another such profile. There also can be no manipulation when there are fewer than k approved candidates in the truthful profile, given that F' makes all voters maximally happy in that case. The only remaining case is when there are at least k approved candidates in the truthful profile $\boldsymbol{A}$ and some voter i reduces her approval set from A_i to A'_i to push the overall number of approved candidates below k, thereby ensuring that all of the candidates in A'_i will get elected with certainty. But this means that in the manipulated profile all voters other than i get all of the candidates they approve of elected with certainty. Thus, even if all other voters get all of their candidates elected, there still are $|A'_i|$ or more spots that can be filled with candidates approved of by voter i. Due to the Pareto efficiency of F, under profile $\boldsymbol{A}$ *all* those spots will get filled with candidates approved of by voter i. Hence, also in this case she has no incentive to manipulate. □

Our next lemma shows that we can reduce the number of voters as well, at least in situations where $m = k+1$. We use a direct reduction from $n = q \cdot k$ down to $n = k$.

Lemma 4. *If there exists a good voting rule for (n, m, k) with $m = k+1$ and n being a multiple of k, then also for (k, m, k).*

Proof. Suppose we are given a good voting rule F for $q \cdot k$ voters to elect committees of size k from a pool of $k+1$ candidates. We need to design a voting rule F' for just k voters with the same properties. For any profile $\boldsymbol{A}$ of approval ballots submitted by k voters, let $q\boldsymbol{A}$ denote the profile for $q \cdot k$ voters we obtain when we concatenate q copies of $\boldsymbol{A}$. Let us define F' as follows:

$$F'(\boldsymbol{A}) \quad = \quad F(q\boldsymbol{A})$$

We now need to show that F' satisfies our three axioms:

- *Proportionality.* For F', with the number of voters being equal to the committee size, minimal proportionality requires that $c \in X$ for all $X \in F'(\boldsymbol{A})$ if $\boldsymbol{A}$ is a party-list profile in which at least one voter reports $\{c\}$. But for any such profile, $q\boldsymbol{A}$ is a party-list profile in which at least q voters report $\{c\}$, which is the quota required to guarantee $c \in X$ for all $X \in F(q\boldsymbol{A})$. Hence, minimal proportionality of F implies minimal proportionality of F'.
- *Efficiency.* For any two committees X and X', X Pareto-dominates X' under profile $\boldsymbol{A}$ if and only if it does under profile $q\boldsymbol{A}$. Hence, Pareto efficiency of F implies Pareto efficiency of F'.
- *Strategyproofness.* We are going to exploit the fact that when there is exactly one more candidate in the pool than we have to elect, then any given voter can only distinguish between *good* committees (including all the candidates she approves of) and *bad* committees (missing exactly one candidate she approves of). So a voting rule might return (a) a set of only good committees, (b) a set including both good and bad committees, or (c) a set of only bad committees. Under the Kelly extension, outcomes of type (a) are all equally good and strictly better than all other outcomes, outcomes of type (c) are all equally bad and strictly worse than all other outcomes, and outcomes of type (b) are all incomparable to each other. Crucially for our purposes, this means that any two outcomes belonging to the same class are either equally good or incomparable to each other, and therefore that the relation $\not\succ^{\text{CAU}}_{A_i}$ between outcomes must be a *transitive* relation.

  Suppose a cautious voter i is considering to manipulate F' in profile $\boldsymbol{A}$ by changing her ballot, resulting in a new profile $\boldsymbol{A}'$ with $\boldsymbol{A} =_{-i} \boldsymbol{A}'$. Recall that $F'(\boldsymbol{A}) = F(q\boldsymbol{A})$ and $F'(\boldsymbol{A}') = F(q\boldsymbol{A}')$ and note that $q\boldsymbol{A}$ and $q\boldsymbol{A}'$ differ on exactly q ballots. Now suppose we start in profile $q\boldsymbol{A}$ and let those q voters switch, one by one, to the ballots they report in $q\boldsymbol{A}'$. Consider the resulting sequence:

  $$q\boldsymbol{A} = \boldsymbol{A}_0,\; \boldsymbol{A}_1,\; \ldots,\; \boldsymbol{A}_q = q\boldsymbol{A}'$$

  Due to the strategyproofness of F we obtain:

  $$\begin{array}{ccc} F(\boldsymbol{A}_1) & \not\succ^{\text{CAU}}_{A_i} & F(\boldsymbol{A}_0) \\ F(\boldsymbol{A}_2) & \not\succ^{\text{CAU}}_{A_i} & F(\boldsymbol{A}_1) \\ & \vdots & \\ F(\boldsymbol{A}_q) & \not\succ^{\text{CAU}}_{A_i} & F(\boldsymbol{A}_{q-1}) \end{array}$$

  Now, due to the transitivity of $\not\succ^{\text{CAU}}_{A_i}$, we get $F(\boldsymbol{A}_q) \not\succ^{\text{CAU}}_{A_i} F(\boldsymbol{A}_0)$ and thus also $F'(\boldsymbol{A}') \not\succ^{\text{CAU}}_{A_i} F'(\boldsymbol{A})$. In other words, voter i has no incentive to manipulate F' in profile $\boldsymbol{A}$.

This completes the proof. □

Our final lemma shows that we can also reduce k, the target size for committees. This reduction works only under certain constraints on n and m (which, however, is sufficient for our purposes).

Lemma 5. *If there exists a good voting rule for $(n+1, m+1, k+1)$ with $m = k+1$ and $n = k$, then also for (n, m, k).*

Proof. Suppose we are given a good voting rule F for $k+1$ voters to elect committees of size $k+1$ from a pool of $k+2$ candidates. We need to construct a new voting rule F' for k voters to elect committees of size k from a reduced pool of just $k+1$ candidates. Let $c^\star$ be the additional candidate featuring only in the original pool. We define F' as follows, for any given profile $\boldsymbol{A} = (A_1, \ldots, A_k)$:

$$F'(\boldsymbol{A}) \quad = \quad \{\, X \setminus \{c^\star\} \mid X \in F(A_1, \ldots, A_k, \{c^\star\}) \,\}$$

Thus, we create a $k+1$st dummy voter, assume she approves of only the dummy candidate $c^\star$, invoke F to obtain a set of winning committees, and then remove $c^\star$ from each of these committees.

We first need to show that F' only returns committees of size k. But by Lemma 1, every committee X in $F(A_1, \ldots, A_k, \{c^\star\})$ has to include $c^\star$. Thus, removing $c^\star$ from any such committee of size $k+1$ indeed leaves us with a committee of size k.

It remains for us to show that F' is good, i.e., that it is minimally proportional, Pareto efficient, and strategyproof for cautious voters:

- *Proportionality.* When the number of voters equals the committee size, minimal proportionality reduces to the following requirement: for any candidate c, if at least one voter reports $\{c\}$ in a party-list profile, then all winning committees must include c. But this property is immediately seen to transfer from F to F'.
- *Efficiency.* As we have seen, every $X \in F(A_1, \ldots, A_k, \{c^\star\})$ includes $c^\star$, so the final voter is fully satisfied. Hence, Pareto efficiency of F ensures that there is no Pareto-superior way of filling the remaining k seats on the committees as far as the first k voters are concerned. Thus, Pareto efficiency of F also transfers to F'.
- *Strategyproofness.* A voter whose preferences depend on the number of elected candidates she approves of and who does not approve of $c^\star$ will prefer X to X' if and only if she prefers $X \setminus \{c^\star\}$ to $X' \setminus \{c^\star\}$. Thus, strategyproofness also transfers immediately.

This completes the proof. □

We are now ready to prove our main result.

Proof of Theorem 1. For the sake of contradiction, suppose there exists a good voting rule for (n, m, k) with $k \geqslant 3$, $m > k$, and $n = q \cdot k$. Then, by $m - k - 1$ applications of Lemma 3, the same must be true for $(q \cdot k, k+1, k)$. By a single application of Lemma 4, it must also hold for $(k, k+1, k)$ and, by a further $k-3$ applications of Lemma 5, for $(3, 4, 3)$. But the latter contradicts Lemma 2. □

3.3 Necessity of Assumptions for the Impossibility

One may ask whether Theorem 1 can be strengthened further and whether the impossibility it establishes will prevail also under less stringent assumptions. Next, we show that this is not the case, in the sense that each of the three axioms it appeals to is a necessary requirement for obtaining the impossibility result. Indeed, if we drop one of the axioms, then we instead obtain a possibility result, i.e., we are able to formulate a voting rule that satisfies all of the remaining requirements.

Proposition 1 below follows immediately from the relevant definitions. Proposition 2 follows from known properties of PAV [2].

Proposition 1. *F_{AV} is Pareto efficient and immune to manipulation by cautious voters.*

Proposition 2. *F_{PAV} is Pareto efficient and minimally proportional.*

Finding a voting rule that is proportional and strategyproof (though not Pareto efficient) is much more challenging, but SAT solving can help here as well. Applying LINGELING to the conjunction of $\varphi_{\text{AT-LEAST-ONE}}$, φ_{SP}, and φ_{PROP} for $(3, 4, 3)$ returns a satisfying model. This shows that there exists a voting rule for these parameters that satisfies the two axioms. The model returned fully describes a voting rule of the desired kind. However, it does so in an *explicit form*, by listing the set of committees returned for each of the 4096 profiles. By successively encoding and adding further axioms that impose a certain amount of structure (such as the well-known *anonymity* and *neutrality* axioms), we were able to arrive at a more interpretable representation of a rule, which allowed us to prove the following result.

Proposition 3. *There exists a voting rule for 3 voters for electing committees of size 3 from a pool of 4 candidates that is minimally proportional and immune to manipulation by cautious voters.*

Proof. Let $N = \{1, 2, 3\}$, $C = \{a, b, c, d\}$, and $k = 3$. For any given profile $\boldsymbol{A}$, let $N^\star(\boldsymbol{A}) \subseteq N$ be the largest set of voters each approving at least one candidate with $|\bigcup_{i \in N^\star(\boldsymbol{A})} A_i| \leqslant |N^\star(\boldsymbol{A})|$.[4] Define the rule $F^\star$ as follows:

$$F^\star(\boldsymbol{A}) \quad = \quad \{\, X \in [C]^k \mid \textstyle\bigcup_{i \in N^\star(\boldsymbol{A})} A_i \subseteq X \,\}$$

$F^\star$ is minimally proportional, as a voter who only approves of a single candidate c in some profile $\boldsymbol{A}$ will always belong to $N^\star(\boldsymbol{A})$.

To see that $F^\star$ is immune to manipulation by cautious voters, first observe that a member of the coalition $N^\star(\boldsymbol{A})$ never has an incentive to manipulate in profile $\boldsymbol{A}$. Any other voter can alter the outcome only by changing her ballot so as to join this coalition of voters determining the outcome. We distinguish two cases. First, a voter who is not part of the coalition cannot successfully manipulate by *adding* more candidates to her ballot, as such a move would never make her join the coalition deciding the outcome. Second, a simple case distinction shows that a voter who is not part of the coalition also cannot successfully manipulate by *dropping* candidates from her ballot. To see this, keep in mind that there are only two other agents to consider and that our voting rule must elect all but one of the four candidates. The main insight here is the fact that the set of candidates approved by the new coalition formed after the manipulator has dropped some candidates from her ballot will never include *all* of the candidates she truthfully approves of (otherwise that coalition would have been feasible in the truthful profile already). □

We note that the techniques developed in this paper do not allow us to infer anything about the possibility of designing a rule that is proportional and strategyproof for larger values of the parameters than those considered in Proposition 3. Nevertheless, we have been able to verify—again with the help of a SAT solver—that Proposition 3 also holds for $(3, 5, 3)$ and $(3, 5, 4)$. However, the rules establishing this possibility look very different from the one defined in the proof above. So these results do not suggest any obvious approach for defining rules that are proportional and strategyproof for arbitrary values of our parameters.

Finally, the condition of k dividing n featuring in the statement of Theorem 1 is necessary for at least certain values of the parameters. Indeed, for $(4, 4, 3)$ the formula encoding the voting-rule design problem turns out to be satisfiable. At an intuitive level, this may be explained by the fact that the proportionality axiom is at its most demanding when n is divisible by k and at its least demanding when n is the successor of a number divisible by k. Whether or not the impossibility persists for *all* triples (n, m, k) from certain values onwards is an open question of some technical interest. Having said this, answering this question arguably would only have a limited impact on the design of voting rules in practice, given that the designer usually cannot control the exact number of voters participating.

4 VARIANTS AND DISCUSSION

Our main result shows that it is impossible to design an approval-based multiwinner voting rule that is proportional, strategyproof, and efficient. To be precise, we have shown this to be true only for one specific combination of choices for how to formalise these axioms. These choices are reasonable and our axioms are logically weaker than other, similarly reasonable formulations, meaning that the theorem is logically stronger than most reasonable variants would be. Still, other such variants may be of interest to some readers and not all reasonable variants are logically weaker than our result.

[4] Note that $N^\star(\boldsymbol{A})$ is always uniquely defined for the case of 3 voters and 4 candidates. The same construction is not possible for arbitrary n and m.

In this section, we therefore discuss some of these variants. Our objective here is not to provide further fully-fledged theorems relying on subtly different assumptions, but rather to point out what possibilities there are and, specifically, to highlight the utility of SAT solving techniques for such an exploration, which allowed us to efficiently and conveniently check whether a given combination of axioms is a promising candidate for either a possibility or an impossibility result.

4.1 Varying the Proportionality Axiom

The axiom of minimal proportionality is a technical axiom that captures a certain feature one would expect any "proportional" voting rule to exhibit, but it may appear to have a certain air of arbitrariness about it. Other proportionality axioms formulated in the literature, such a *justified representation* and *extended justified representation* [2], may appear more natural. They indeed are, but they also *imply* minimal proportionality. Thus, Theorem 1 of course remains valid if we use one of these proportionality axioms instead.

4.2 Varying the Strategyproofness Axiom

Our strategyproofness axiom requires that no agent can deviate to another profile under which the worst committee elected is at least as good as the best committee elected under the truthful profile, and furthermore the best committee elected is strictly better than the worst committee elected under the truthful profile. We do not impose any constraints on *how* a manipulator changes her ballot. It turns out that the impossibility remains valid, at least for the case of $(3, 4, 3)$, if we assume that voters will only ever attempt to manipulate by *dropping* candidates from their approval sets. This can be easily verified using a SAT solver (and is also evident from the proof of Lemma 2).

If we assume that voters will only try to manipulate by *adding* candidates, then another query to the SAT solver reveals that PAV for $(3, 4, 3)$ is immune to manipulation by cautious voters (on top of, by Proposition 2, meeting all other requirements).

The Kelly preference extension leads to a very weak strategyproofness axiom, making Theorem 1 particularly strong in this respect.[5] Of course, if we replace Kelly-strategyproofness with a more demanding axiom, such as Fishburn-strategyproofness [16] or Gärdenfors-strategyproofness [17], then the impossibility we established will persist. Of particular interest is the notion of strategyproofness used in the Duggan-Schwartz Theorem [13]. In its formulation by Taylor [30], it requires immunity to manipulation by both *optimistic* and *pessimistic* voters.

An optimistic voter is a voter who believes that ties will always be broken in her favour; a pessimistic voter is a voter who believes that ties will always be broken in the worst possible way. The corresponding preference extensions can be defined as follows:

$$\begin{aligned} \mathcal{X} \succcurlyeq_i^{\text{OPT}} \mathcal{X}' &\text{ if and only if } (\exists X \in \mathcal{X})\, (\forall X' \in \mathcal{X}')\, X \succcurlyeq_i X' \\ \mathcal{X} \succcurlyeq_i^{\text{PES}} \mathcal{X}' &\text{ if and only if } (\exists X' \in \mathcal{X}')\, (\forall X \in \mathcal{X})\, X \succcurlyeq_i X' \end{aligned}$$

Observe that requiring immunity to manipulation by *both* optimistic and pessimistic voters is strictly more demanding a requirement than immunity to manipulation by cautious voters, so Theorem 1 remains true if we use this notion of manipulation instead.

But what if we assume that either all voters are optimists or all voters are pessimists? Consulting the SAT solver for parameters $(3, 4, 3)$ shows that rules that are minimally proportional, Pareto efficient, and immune to manipulation by *either* optimistic or pessimistic voters (but not both!) exist. In the case of manipulation by pessimistic voters, this possibility result however is very brittle: if we strengthen minimal proportionality to the proportionality axiom featuring in the statement of Lemma 1, then we obtain an impossibility.[6]

4.3 Varying the Efficiency Axiom

Peters [26] used a less demanding efficiency axiom than the standard notion of Pareto efficiency we have used here. His axiom of *weak efficiency* merely requires that, provided at least k candidates have received at least one approval each, no candidate who has not received any approvals will be part of any of the committees selected:

$$\text{if } \left|\bigcup_{i \in N} A_i\right| \geqslant k \text{ then } X \subseteq \bigcup_{i \in N} A_i \text{ for all } X \in F(\boldsymbol{A})$$

Interestingly, if we replace Pareto efficiency by weak efficiency in the statement of Theorem 1, then that theorem ceases to hold and we obtain a possibility result, at least for parameters $(3, 4, 3)$. Indeed, as another query to the SAT solver demonstrates, the rule $F^\star$ shown to be minimally proportional and immune to manipulation by cautious voters defined in the proof of Proposition 3 satisfies weak efficiency as well. Of course, in practice we certainly would hope to be able to use a rule that is Pareto efficient and not just weakly efficient.[7] Nevertheless, this observation provides some intriguing insight into how relaxing the resoluteness requirement can turn an impossibility result into a possibility result.

4.4 Varying the Assumptions on Preferences

Recall how, in Section 2.2, we defined the preferences of agent i over alternative committees in terms of the *cardinality* of the intersections of those committees with her true approval set A_i. An alternative approach would be to only commit to voter i weakly preferring committee X to committee X' in case $A_i \cap X \supseteq A_i \cap X'$. This approach has the advantage of requiring us to make less specific assumptions about voters, but it has the disadvantage of making the preference relations induced by approval sets incomplete. Recall that the manner in which we define preferences affects both our efficiency axiom and our strategyproofness axiom.[8]

If we change the encoding of Pareto efficiency and Kelly-strategyproofness from the cardinality-based to this set-based notion of preference, then the SAT solver returns a possibility for the case of $(3, 4, 3)$. Interestingly, Peters' impossibility theorem for resolute rules [25, 26] also goes through for set-based preferences. So this finding constitutes another example that highlights how imposing resoluteness increases the chances of encountering an impossibility.

5 CONCLUSION

We have proved a new impossibility theorem in social choice theory regarding the design of multiwinner voting rules with approval ballots that are proportional, strategyproof, and efficient. This result

[5] For comparison, for irresolute single-winner voting rules with ranked preferences, Kelly-strategyproofness allows for possibility results [21, 10].

[6] Thus, Lemma 1 ceases to hold if we replace immunity to manipulation by cautious voters with immunity to manipulation by pessimistic voters.

[7] Having said this, as one anonymous reviewer of this paper has pointed out, there are examples for reasonable rules being discussed in the literature (e.g., a sequential variant of PAV) that are not Pareto efficient [24].

[8] The standard approach to defining preference extensions [4] assumes that the underlying preference relation is complete. Our definitions of $\succcurlyeq_i^{\text{CAU}}$, $\succcurlyeq_i^{\text{OPT}}$, and $\succcurlyeq_i^{\text{PES}}$ are well-defined and intuitively appealing also when $\succcurlyeq_i$ is incomplete. Still, developing a theory of extensions of incomplete preferences would be a deserving topic for future research in its own right.

applies even for the arguably weakest possible formulations of the core axioms we have been interested in, namely proportionality and strategyproofness, together with the standard formulation of Pareto efficiency. At the same time, our result relies on a specific assumption on how preferences are induced by the number of elected candidates a voter approves of, and there is room for follow-up work that explores different assumptions regarding this matter.

Whether or not the reader is interested in the specific contribution to social choice theory we have been able to make here, we believe that our work convincingly demonstrates the significant benefits of the SAT-based approach when working with the axiomatic method. Maybe more so than in previous contributions of this kind, what axioms to focus on has been entirely unclear at the start of our investigation. We encoded more than 25 different variants of the axioms we eventually selected to feature in our main result and it would have been much harder, maybe impossible, to carry out this kind of investigation using purely analytic methods alone. This study thus highlights the great potential of SAT solving technology, and indeed AI methods more generally, for formal investigations into the normative adequacy of methods for collective decision making.

Acknowledgments. We would like to thank our three anonymous ECAI reviewers for the helpful feedback received.

REFERENCES

[1] Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 1. North-Holland, 2002.

[2] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017.

[3] Michel L. Balinski and H. Peyton Young. *Fair Representation: Meeting the Ideal of One Man, One Vote*. Brookings Institution Press, 2nd edition, 2001.

[4] Salvador Barberà, Walter Bossert, and Prasanta K. Pattanaik. Ranking sets of objects. In Salvador Barberà, Peter J. Hammond, and Christian Seidl, editors, *Handbook of Utility Theory*, volume 2, chapter 17, pages 893–977. Kluwer Academic Publishers, 2004.

[5] Anton Belov and Joao Marques-Silva. MUSer2: An efficient MUS extractor. *Journal on Satisfiability, Boolean Modeling and Computation*, 8:123–128, 2012.

[6] Armin Biere. CaDiCaL, Lingeling, Plingeling, Treengeling, YalSAT entering the SAT competition 2017. In Tomáš Balyo, Marijn Heule, and Matti Järvisalo, editors, *Proceedings of the SAT Competition 2017: Solver and Benchmark Descriptions*, volume B-2017-1 of *Department of Computer Science Series of Publications*, pages 14–15. University of Helsinki, 2017.

[7] Armin Biere, Marijn Heule, and Hans van Maaren, editors. *Handbook of Satisfiability*. IOS Press, 2009.

[8] Florian Brandl, Felix Brandt, Christian Geist, and Johannes Hofbauer. Strategic abstention based on preference extensions: Positive results and computer-generated impossibilities. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI-2015)*, 2015.

[9] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[10] Felix Brandt and Christian Geist. Finding strategyproof social choice functions via SAT solving. *Journal of Artificial Intelligence Research*, 55:565–602, 2016.

[11] Felix Brandt, Christian Saile, and Christian Stricker. Voting with ties: Strong impossibilities via SAT solving. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2108)*. IFAAMAS, 2018.

[12] Siddharth Chatterjee and Arunava Sen. Automated reasoning in social choice theory: Some remarks. *Mathematics in Computer Science*, 8(1):5–10, 2014.

[13] John Duggan and Thomas Schwartz. Strategic manipulability without resoluteness or shared beliefs: Gibbard-Satterthwaite generalized. *Social Choice and Welfare*, 17(1):85–93, 2000.

[14] Ulle Endriss. Analysis of one-to-one matching mechanisms via SAT solving: Impossibilities for universal axioms. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-2020)*, 2020.

[15] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Multiwinner voting: A new challenge for social choice theory. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 2, pages 27–47. AI Access, 2017.

[16] Peter C. Fishburn. Even-chance lotteries in social choice theory. *Theory and Decision*, 3(1):18–40, 1972.

[17] Peter Gärdenfors. Manipulation of social choice functions. *Journal of Economic Theory*, 13(2):217–228, 1976.

[18] Christian Geist and Ulle Endriss. Automated search for impossibility theorems in social choice theory: Ranking sets of objects. *Journal of Artificial Intelligence Research*, 40:143–174, 2011.

[19] Christian Geist and Dominik Peters. Computer-aided methods for social choice theory. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 13, pages 249–267. AI Access, 2017.

[20] Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

[21] Jerry S. Kelly. Strategy-proofness and social choice functions without single-valuedness. *Econometrica*, 45(2):439–446, 1977.

[22] D. Marc Kilgour. Approval balloting for multi-winner elections. In Jean-François Laslier and M. Remzi Sanver, editors, *Handbook on Approval Voting*, Studies in Choice and Welfare, chapter 6, pages 105–124. Springer, 2010.

[23] Boas Kluiving, Adriaan de Vries, Pepijn Vrijbergen, Arthur Boixel, and Ulle Endriss. Supplementary material for "Analysing irresolute multiwinner voting rules with approval ballots via SAT solving". Zenodo, 2020. `http://doi.org/10.5281/zenodo.3675459`.

[24] Martin Lackner and Piotr Skowron. A quantitative analysis of multiwinner rules. `arXiv:1801.01527 [cs.MA]`, 2018. Extended version of a paper published in the Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI-2019).

[25] Dominik Peters. Proportionality and strategyproofness in multiwinner elections. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2018)*. IFAAMAS, 2018.

[26] Dominik Peters. *Fair Division of the Commons*. DPhil thesis, University of Oxford, 2019.

[27] Luis Sánchez-Fernández, Edith Elkind, Martin Lackner, Norberto Fernández, Jesús A Fisteus, Pablo Basanta Val, and Piotr Skowron. Proportional justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI-2017)*, 2017.

[28] Mark Allen Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

[29] Pingzhong Tang and Fangzhen Lin. Computer-aided proofs of Arrow's and other impossibility theorems. *Artificial Intelligence*, 173(11):1041–1053, 2009.

[30] Alan D. Taylor. The manipulability of voting systems. *The American Mathematical Monthly*, 109(4):321–337, 2002.

[31] Thorvald N. Thiele. Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441. 1895.