Some Notes on Justified Representation*

Luis Sánchez-Fernández
Universidad Carlos III de Madrid, Spain
luiss@it.uc3m.es

Norberto Fernández
Escuela Naval Militar (CUD), Spain
norberto@cud.uvigo.es

Jesús A. Fisteus
Universidad Carlos III de Madrid, Spain
jaf@it.uc3m.es

Pablo Basanta Val
Universidad Carlos III de Madrid, Spain
pbasanta@it.uc3m.es

Abstract

Multi-winner voting systems are often applied to scenarios in which it is desirable that the set of winners represents the different opinions or preferences of the agents involved in the election. Because of that, the development of axioms that capture the idea of representation and the study of multi-winner voting rules with such axioms is of great interest. In the context of approval-based committee voting, Aziz *et al.* proposed in 2015 at the AAAI Conference two axioms related to the concept of representation. These axioms are called *justified representation* (JR) and *extended justified representation* (EJR). In this paper we present new results related to these axioms. First of all, we close an issue that was left open by Aziz *et al.* regarding the maximum number of seats for which the Reweighted Approval Voting satisfies JR. Second, we discuss a problem in the definition of EJR: a set of candidates can provide perfect representation for a given election and fail to provide EJR. We propose an alternative axiom which we have called *proportional justified representation* (PJR). We prove that PJR remedies that problem, while providing precisely the same results as EJR for all the voting systems that Aziz *et al.* analyzed in their paper.

1 Introduction

Decision making based on the aggregation of possibly conflicting preferences is a central problem in the field of social choice that has received a considerable amount of attention from the artificial intelligence research community [Conitzer, 2010; Sandholm, 1999]. A voting system is the usual way of making collective decisions.

The selection of a single candidate out of several is the most common scenario in which voting systems are studied. However, the scenario in which a winning set of candidates is selected (multi-winner elections) is also frequent. The most typical situation is the election of a parliament or a committee. Multi-winner elections can also be used by software agents in scenarios such as deciding on a set of plans [Elkind *et al.*, 2011] or resource allocation [Skowron *et al.*, 2013]. Recently, their complexity [Betzler *et al.*, 2013] and social choice properties [Elkind *et al.*, 2011; 2014; Aziz *et al.*, 2015] have been the subjects of active research by the artificial intelligence community.

Multi-winner voting systems are applied often in scenarios in which the set of winners needs to represent the different opinions or preferences of the agents involved in the election. Due to this, the development of axioms that capture the idea of representation and the study of multi-winner voting rules with such axioms is of great interest. In the context of approval-based committee voting, in 2015 an interesting paper by Aziz *et al.* [Aziz *et al.*, 2015] presented at the AAAI Conference proposed two axioms related to the concept of representation. These axioms are called *justified representation* (JR) and *extended justified representation* (EJR).

Roughly speaking, JR establishes requirements on when a large enough group of agents deserves to have at least one of the candidates they approve elected. Similarly, EJR establishes requirements on when a large enough group of agents deserves to have several of the candidates approved by them elected. The formal definitions can be found in [Aziz *et al.*, 2015] and will be reviewed later in this paper. Both axioms are interesting, because voting rules that satisfy them guarantee that a large enough group of agents (even if it is a minority of the total agents) will receive at least one (for JR) or at least x (for EJR) representatives that they approve regardless of any strategic vote followed by the reminder agents. Similar axioms have been proposed for multi-winner voting rules with ranked ballots [Dummet, 1984; Elkind *et al.*, 2014].

In this paper we present new results related to these axioms. It was proved in [Aziz *et al.*, 2015] that the Reweighted Approval Voting (RAV)[1] satisfies JR if the number of seats is 2 but it fails JR if the number of seats is greater than or equal to 10. Whether RAV satisfies JR if the number of seats is between 3 and 9 was left open. We close this issue here and prove that RAV satisfies JR if the number of seats is smaller

*This work was supported in part by the Spanish Ministerio de Economa y Competitividad (project HERMES-SMARTDRIVER TIN2013-46801-C4-2-R) and by the Autonomous Community of Madrid (project e-Madrid S2013/ICE-2715).

[1] The definition of the Reweighted Approval Voting will be reviewed later in this paper.

than or equal to 5 and fails JR if the number of seats is greater than or equal to 6. This is discussed in Section 3.

Second, in Section 4 we discuss a problem in the definition of EJR: a set of candidates can provide perfect representation for a given election and fail to provide EJR. We propose an alternative axiom which we have called *proportional justified representation* (PJR). We prove that PJR remedies that problem, while providing precisely the same results than EJR for all the voting systems analyzed in the paper of Aziz *et al.* Section 5 deals with some complexity issues related to the topics discussed in Section 4.

The other Sections in which the paper is organized are Section 2, where we introduce some useful notation and briefly describe the voting systems that we will analyze, and Section 6, devoted to discuss the results we have presented along the paper and to propose future lines of work.

2 Preliminaries

We will use the same notation used in [Aziz *et al.*, 2015], that we repeat here.

We consider a social choice setting (an election) with a set of agents (voters) $N = \{1, \ldots, n\}$ and a set of candidates $C = \{c_1, \ldots, c_q\}$. Each agent $i \in N$ submits an approval ballot $A_i \subseteq C$, which represents the subset of candidates that she approves of. We refer to the list $\mathcal{A} = (A_1, \ldots, A_n)$ of approval ballots as the *ballot profile*. We will consider approval-based multi-winner voting rules that take as input $(N, C, \mathcal{A}, k)$, where k is a positive integer that satisfies $k \leq |C|$, and return a subset $W \subseteq C$ of size k, which we call the *winning set*. We omit N and C from the notation when they are clear from the context.

We assume that voting rules can output several winning sets of candidates. We will use the expression "(tied) winning sets of candidates" to express that the voting rule may return more than one winning set. We do not make any particular assumption about how ties are broken.

We adopt the following criterion with respect to axioms and ties. If a voting rule outputs several (tied) winning sets of candidates for a certain election and one of such sets would make the voting rule fail a certain axiom we assume that such voting rule fails that axiom.

Along this paper we consider the following voting rules, also discussed in [Aziz *et al.*, 2015] and in [Kilgour, 2010]:

Proportional Approval Voting (PAV) Under PAV, an agent is assumed to derive an utility of $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{j}$ from a committee that contains exactly j of her approved candidates and the goal is to maximize the sum of the agents' utilities. Formally, the PAV-score of a set $W \subseteq C$ is defined as $\sum_{i \in N} r(|W \cap A_i|)$, where $r(p) = \sum_{j=1}^{p} \frac{1}{j}$, and PAV outputs the set $W \subseteq C$ of size k with the highest PAV-score.

We can generalize the definition of PAV by using an arbitrary non-increasing score vector in place of $(1, \frac{1}{2}, \frac{1}{3}, \cdots)$: for every vector $\mathbf{w} = (w_1, w_2, \ldots)$, where $w_1, w_2, \ldots$ are non-negative reals, $w_1 = 1$ and $w_1 \geq w_2 \geq \ldots$, we define a voting rule $\mathbf{w}$-PAV that, given a ballot profile $(A_1, \ldots, A_n)$ and a target number of winners k, returns a set W of size k with the highest $\mathbf{w}$-PAV score, defined by $\sum_{i \in N} r_{\mathbf{w}}(|W \cap A_i|)$, where $r_{\mathbf{w}}(p) = \sum_{j=1}^{p} w_j$.

Reweighted Approval Voting (RAV) RAV is a multi-round rule that in each round selects a candidate and then reweights the approvals for the subsequent rounds. Specifically, it starts by setting $W = \emptyset$. Then in round $j, j = 1, \ldots, k$, it computes the *approval-weight* of each candidate c as:

$$\sum_{i | c \in A_i} \frac{1}{1 + |W \cap A_i|},$$

selects a candidate with the highest approval weight, and adds her to W. Just as for PAV, we can extend the definition of RAV to score vectors other than $(1, \frac{1}{2}, \frac{1}{3}, \cdots)$: every vector $\mathbf{w} = (w_1, w_2, \ldots)$, where $w_1, w_2, \ldots$ are non-negative reals, $w_1 = 1$ and $w_1 \geq w_2 \geq \ldots$ defines a sequential voting rule $\mathbf{w}$-RAV, which proceeds as RAV except that it computes the approval weight of a candidate c in round j as $\sum_{i | c \in A_i} w_{|W \cap A_i|+1}$, where W is the winning set after the first $j - 1$ rounds.

3 JR and RAV

First of all, we repeat here the definition of justified representation (JR), given in [Aziz *et al.*, 2015]:

Justified representation (JR) *Given a ballot profile $\mathcal{A} = (A_1, \ldots, A_n)$ over a candidate set C and a target committee size k, $k \leq |C|$, we say that a set of candidates W of size $|W| = k$ provides justified representation for $(\mathcal{A}, k)$ if there does not exist a set of voters (agents) $N^* \subseteq N$ with $|N^*| \geq n/k$ such that $\bigcap_{i \in N^*} A_i \neq \emptyset$ and $A_i \cap W = \emptyset$ for all $i \in N^*$. We say that an approval-based voting rule satisfies justified representation (JR) if for every profile $\mathcal{A} = (A_1, \ldots, A_n)$ and every target committee size k it outputs a winning set that provides justified representation for $(\mathcal{A}, k)$.*

It was proved in [Aziz *et al.*, 2015] that RAV satisfies JR for $k = 2$, but fails it for $k \geq 10$. Whether RAV satisfies JR for $k = 3, \ldots, 9$ was left an open problem. The following theorem closes this issue.

THEOREM 1 *RAV satisfies JR for $k \leq 5$ but fails it for $k \geq 6$.*

Proof: First of all, for a fixed committee size k, we compute the maximum approval weight of a candidate c_k after $k - 1$ candidates have already been elected. This is done by solving a linear programming problem, shown below.

Because non-elected candidates do not have any influence in the approval weight under RAV, we may restrict to the case in which $C = \{c_1, \ldots, c_k\}$. We also assume, without loss of generality, that the order in which the candidates are elected is: $c_1, c_2, \ldots, c_k$. In the linear programming problem we will define constraints that impose such order.

The linear programming problem has a variable x_A for each nonempty candidate subset $A \subseteq C$. x_A is equal to n_A/n, where n_A is the number of agents that submit the approval ballot A. The objective function of the linear programming problem is the approval weight of candidate c_k (normalized by the total number of agents n) after candidates $\{c_1, \ldots, c_{k-1}\}$ have already been elected:

$$\text{maximize} \sum_{c_k \in A} \frac{x_A}{1 + |\{c_1, \ldots, c_{k-1}\} \cap A|}.$$

The constraints of the linear programming problem are as follows.

1. For any nonempty candidate subset $A \subseteq C$, x_A must be nonnegative:

$$x_A \geq 0 \text{ for all } x_A.$$

2. The total number of votes should be equal to n and therefore the sum of all x_A must be equal to 1 ($\sum_A x_A = \sum_A n_A/n = (\sum_A n_A)/n = n/n = 1$):

$$\sum_A x_A = 1.$$

3. For $i = 1, \ldots, k-1$, the candidate elected at iteration i must be c_i. That is, the approval weight of candidate c_i at iteration i must be greater than or equal to the approval weight of any other not yet elected candidate c_j, with $j = i+1, \ldots, k$:

$$\sum_{A: c_i \in A} \frac{x_A}{1 + |\{c_1, \ldots, c_{i-1}\} \cap A|} \geq \sum_{B: c_j \in B} \frac{x_B}{1 + |\{c_1, \ldots, c_{i-1}\} \cap B|}$$
$$\text{for } i = 1, \ldots, k-1 \text{ and for } j = i+1, \ldots, k.$$

The number of variables grows exponentially with k, but this is not a problem because we have only to run the linear programming problem for small values of k. In particular, for $k = 6$, the value of the objective function obtained when running the linear programming problem is 0.204, slightly greater than $1/(k-1) = 1/5$. Now, we modify the election obtained with the linear programming problem as follows. We add a new candidate c_7. We also add $n/5$ agents that submit $\{c_7\}$. The total number of agents is now $n' = 6n/5$ and since $n'/6 = n/5$ agents approve c_7, according to JR c_7 must be in the set of winners. But this does not happen[2] because the approval weight of c_6 in round 6 is $0.204n > n/5$. Finally, we adjust the values of x_A so that the values of n_A are all nonnegative integers and to avoid ties. We get a counterexample that proves that RAV does not satisfy JR for $k = 6$ (table 1).

The total number of agents in this example is 5992. Since 1000 agents approve only c_7, according to JR c_7 should be in the winning set. But this does not happen: in round 1 c_1 is elected with an approval weight of 2000; in round 2 c_2 is elected with an approval weight of 1499; in round 3 c_3 is elected with an approval weight of 1220.5; in round 4 c_4 is elected with an approval weight of 1060.33; in round 5 c_5 is elected with an approval weight of 1017.67; and finally, in round 6 c_6 is elected with an approval weight of 1017.17. Ties do not happen.

For $k = 7, 8, 9$ we simply quote the argument used in [Aziz *et al.*, 2015] to prove that RAV does not satisfy JR for $k > 10$ (adapted to the election in table 1):

> "We simply add to the election shown in table 1 $k-6$ additional candidates and $1000(k-6)$ additional agents such that for each new candidate there are 1000 agents who approve that candidate only. Note that we still have $1000 > n/k$. RAV will proceed to select $c_1, \ldots, c_6$, followed by $k-6$ additional candidates, and c_7 or one of the new candidates will remain unselected."

For $k = 3, 4, 5$ the objective function is always smaller than $1/(k-1)$. In particular, for $k = 5$, the objective function is 0.2389. Therefore, the approach used for $k = 6$ of adding an additional candidate would fail. For instance, in the case of $k = 5$ if we add a candidate c_6 and we add $n/4$ agents that submit $\{c_6\}$ (this is the minimum value such that if c_6 is not elected the JR rule would be violated), then at round 5 the approval weight of c_5 would be $0.2389n$ which is less that $n/4$. Therefore c_6 would be elected and the JR rule would not be violated.

Now, for $k = 3, 4, 5$, consider an arbitrary ballot profile $\mathcal{A}$. Let W_{k-1} be the set of the first $k-1$ elected candidates and let c_{k-1} the candidate elected at round $k-1$. First of all, we are going to prove that it is not possible that after $k-1$ iterations two disjoint sets of agents N_1^* and N_2^* exist such that for $j = 1, 2$ it holds that $N_j^* \subseteq N$, $|N_j^*| \geq n/k$, $\bigcap_{i \in N_j^*} A_i \neq \emptyset$ and $A_i \cap W_{k-1} = \emptyset$ for all $i \in N_j^*$ (if such two sets of agents could exist at iteration k one candidate approved by one of the sets could be elected and still the other set would violate JR). The linear programming problem says that the approval weight of c_{k-1} at round $k-1$ is strictly smaller than $1/(k-2)$ multiplied by the total number of agents that approve any of the candidates in W_{k-1} (agents in N_1^* and N_2^* do not approve any candidate in W_{k-1}, and therefore they cannot contribute to the approval weight of c_{k-1} by any means). That is, the approval weight of c_{k-1} if N_1^* and N_2^* would exist would be strictly smaller than $\frac{1}{k-2}(n - |N_1^*| - |N_2^*|) \leq \frac{1}{k-2}(n - \frac{n}{k} - \frac{n}{k}) = n\frac{1}{k-2}\frac{k-2}{k} = \frac{n}{k}$. But any candidate in $\bigcap_{i \in N_j^*} A_i$ would have an approval weight of at least $\frac{n}{k}$, and this contradicts the assumption that c_{k-1} is the candidate elected at round $k-1$.

Suppose now that after $k-1$ candidates have been elected, there exists a (unique) set of agents $N^* \subseteq N$ with $|N^*| \geq n/k$ such that $\bigcap_{i \in N^*} A_i \neq \emptyset$ and $A_i \cap W_{k-1} = \emptyset$ for all $i \in N^*$. A candidate c in $\bigcap_{i \in N^*} A_i$ would have an approval weight of at least n/k.

JR could be violated only if a candidate c' that is not approved by at least $\frac{n}{k}$ of the agents in N^* is elected at iteration k. But once again, the linear programming problem says that such candidate c' would have an approval weight strictly smaller than $\frac{1}{k-1}$ multiplied by the total number of agents that approve c' or any candidate in W_{k-1}. Therefore, the approval weight of c' has to be strictly smaller than $\frac{1}{k-1}(n - \frac{n}{k}) = n\frac{1}{k-1}\frac{k-1}{k} = \frac{n}{k}$. $\square$

4 A criticism to EJR and an alternative proposal

As discussed in [Aziz *et al.*, 2015], the JR axiom establishes requirements regarding when a large enough group of

[2]Observe that in RAV the approval weights of the candidates elected at each round are monotonically non-increasing, and therefore c_7 cannot also be elected in rounds $1, \ldots, 5$.

A	$\{c_1, c_2, c_4, c_5\}$	$\{c_1, c_2, c_4, c_6\}$	$\{c_1, c_3, c_5\}$	$\{c_1, c_3, c_6\}$	$\{c_2, c_3, c_4\}$	$\{c_2, c_3, c_5\}$	$\{c_2, c_5\}$	$\{c_2, c_6\}$
n_A	500	500	500	500	222	333	55	389
A	$\{c_3, c_4\}$	$\{c_3, c_5\}$	$\{c_3, c_6\}$	$\{c_4\}$	$\{c_5\}$	$\{c_6\}$	$\{c_7\}$	
n_A	246	43	154	530	566	454	1000	

Table 1: Example of RAV failing JR for $k = 6$

agents deserves at least *one* representative. However, JR does not capture the intuition that large enough groups of agents should be allocated several representatives. To remedy this, the authors of [Aziz *et al.*, 2015] propose a new axiom, called extended justified representation (EJR):

Extended justified representation (EJR) *Consider a ballot profile $\mathcal{A} = (A_1, \ldots, A_n)$ over a candidate set C, a target committee size k, $k \leq |C|$, and a positive integer ℓ, $\ell \leq k$. We say that a set of candidates W, $|W| = k$, provides ℓ-justified representation for $(\mathcal{A}, k)$ if there does not exist a set of voters (agents) $N^* \subseteq N$ with $|N^*| \geq \ell\frac{n}{k}$ such that $|\bigcap_{i \in N^*} A_i| \geq \ell$, but $|A_i \cap W| < \ell$ for each $i \in N^*$; we say that W provides extended justified representation (EJR) for $(\mathcal{A}, k)$ if it provides ℓ-JR for $(\mathcal{A}, k)$ for all $\ell, 1 \leq \ell \leq k$. We say that an approval-based voting rule satisfies ℓ-justified representation (ℓ-JR) if for every profile $\mathcal{A} = (A_1, \ldots, A_n)$ and every target committee size k it outputs a committee that provides ℓ-JR for $(\mathcal{A}, k)$. Finally, we say that a rule satisfies extended justified representation (EJR) if it satisfies ℓ-JR for all $\ell, 1 \leq \ell \leq k$.*

We are going to show that according to a widely accepted view of representation, EJR fails as a requirement for proportional/representative voting rules.

The idea of representation has been addressed in several research studies [Monroe, 1995; Dummet, 1984; Black, 1958]. Although different nuances can be found in these studies there exists a common view that representation aims at selecting a committee that reflects as fairly as possible the different opinions or preferences of the agents involved in the election. In the context of approval-based voting this idea of 'fairness' can be expressed as "as most agents as possible must be represented by a winner that they approve and each winner must represent approximately the same number of agents".

Viewed from this perspective, the "ideal" of "perfect" situation occurs when all the agents involved in an election are represented by a winner that they approve and each winner represents exactly the same number of agents. We call this perfect representation.

DEFINITION 1 **Perfect representation** *(PR) Consider a ballot profile $\mathcal{A} = (A_1, \ldots, A_n)$ over a candidate set C, and a target committee size k, $k \leq |C|$. We consider only ballot profiles for which a positive integer m exists such that $n = mk$. We say that a set of candidates W, $|W| = k$, provides perfect representation (PR) if it is possible to partition the set of agents in k equal size disjoint subsets $N_1, \ldots, N_k$ ($N = N_1 \cup \cdots \cup N_k$, $N_i \cap N_j = \emptyset$ for $i, j = 1, \ldots, k, i \neq j$, and $|N_i| = m$ for $i = 1, \ldots, k$), such that it is possible to assign each candidate w in W to one (and only one) subset N_i so that for all pairs (w, N_i) all the agents in N_i approve their assigned candidate w. We say that an approval-based voting rule is a PR voting rule if for every profile $\mathcal{A} = (A_1, \ldots, A_n)$ and every target committee size k it does not output any winning set of candidates W that does not provide PR for $(\mathcal{A}, k)$ if at least one set of candidates W' that provides PR for $(\mathcal{A}, k)$ exists.*

Consider the following example. Let $k = 4$. We have 6 candidates, $C = \{c_1, \ldots, c_6\}$. 8 agents submit the following approval ballots. $A_1 = \{c_1\}$, $A_2 = \{c_2\}$, $A_3 = \{c_3\}$, $A_4 = \{c_4\}$, $A_5 = \{c_1, c_5, c_6\}$, $A_6 = \{c_2, c_5, c_6\}$, $A_7 = \{c_3, c_5, c_6\}$ and $A_8 = \{c_4, c_5, c_6\}$. Now consider the following possible set of winners $W = \{c_1, c_2, c_3, c_4\}$. Clearly W provides PR because we can partition the set of agents in 4 disjoint subsets of size 2 ($N_1 = \{1, 5\}$, $N_2 = \{2, 6\}$, $N_3 = \{3, 7\}$ and $N_4 = \{4, 8\}$) such that it is possible to assign each candidate in W to one (and only one) subset so that all the agents approve their assigned candidate (assign c_1 to N_1, c_2 to N_2, c_3 to N_3 and c_4 to N_4). However, W does not provide EJR. In fact, it does not even provide 2-JR, because for agents 5, 6, 7 and 8 we have $|\bigcap_{i=5}^{8} A_i| = |\{c_5, c_6\}| = 2$ and $|\{5, 6, 7, 8\}| = 4 \geq 2\frac{n}{k} = 2\frac{8}{4} = 4$, but $|A_i \cap W| = 1$, for $i = 5, \ldots, 8$. However, observe that $|W \cap \bigcup_{i=5}^{8} A_i| = 4$, that is, agents 5, 6, 7 and 8 achieve *together* the number of winners they deserve. Observe also that for this election W is the only possible set of winners that provides PR because for any other possible set of winners there always exists at least one agent that does not approve any of the winners. The following theorem and corollary follow immediately.

THEOREM 2 *There exist ballot profiles $\mathcal{A}$ and target committee sizes k for which sets of candidates that provide PR exist but all of them fail to provide EJR (even 2-JR).*

COROLLARY 1 *PR voting rules fail EJR (even 2-JR).*

Voting rules that aim at achieving representativeness should be expected to search for sets of winners that are close to PR. Such voting rules can fail EJR because they output a set of winners for certain election that do not provide EJR despite such set of winners provide PR. To remedy this problem we propose the following alternative axiom:

DEFINITION 2 **Proportional justified representation** *(PJR) Consider a ballot profile $\mathcal{A} = (A_1, \ldots, A_n)$ over a candidate set C, a target committee size k, $k \leq |C|$. We say that a set of candidates W, $|W| = k$, provides proportional justified representation (PJR) for $(\mathcal{A}, k)$ if there does not exist a set of agents $N^* \subseteq N$ and a positive integer ℓ with $|N^*| \geq \ell\frac{n}{k}$ such that $|\bigcap_{i \in N^*} A_i| \geq \ell$, but $|W \cap (\bigcup_{i \in N^*} A_i)| < \ell$. We say that an approval-based voting rule satisfies proportional justified representation (PJR) if for every profile $\mathcal{A} = (A_1, \ldots, A_n)$ and every target committee size k it outputs a committee that provides PJR for $(\mathcal{A}, k)$.*

Sets of candidates that provide PR also provide PJR as proved by the following Theorem:

THEOREM 3 *For every profile $\mathcal{A} = (A_1, \dots, A_n)$ and every target committee size k, if a set of candidates W, $|W| = k$ provides PR, then W provides also PJR.*

Proof: Observe that because W provides PR, $n/k = m$ where m is a positive integer. Consider any set of agents $N^* \subseteq N$ and any positive integer ℓ such that $|N^*| \geq \ell\frac{n}{k} = \ell m$. Because W provides PR, k disjoint subsets $N_1, \dots, N_k$ of size m exist such that $N = N_1 \cup \dots \cup N_k$ and that it is possible to assign each candidate w in W to one (and only one) subset N_i so that for all pairs (w, N_i) all the agents in N_i approve their assigned candidate w. Because the size of N^* is greater than or equal to ℓm and the size of all subsets N_i is equal to m, the number of subsets N_i that have a common agent with N^* is at least ℓ. But since for each N_i there is a different candidate in W that is approved by all agents in N_i, the number of candidates in W approved by some agent in N^* must be greater than or equal to ℓ. □

We now study the relation of PJR with JR and EJR, and whether RAV and PAV (and their variations) satisfy PJR.

A voting rule that satisfies PJR has to satisfy JR because for any set of agents $N^* \subseteq N$ such that $|N^*| \geq \frac{n}{k}$ and $\bigcap_{i\in N^*} A_i \neq \emptyset$ PJR imposes that the set of winners W holds that $|W \cap (\bigcup_{i\in N^*} A_i)| \geq 1$ and therefore, $W \cap A_i \neq \emptyset$ for at least one A_i. Therefore:

LEMMA 1 *PJR implies JR.*

A voting rule that satisfies EJR has to satisfy PJR because for any set of agents $N^* \subseteq N$ and any positive integer ℓ such that $|N^*| \geq \ell\frac{n}{k}$ and $|\bigcap_{i\in N^*} A_i| \geq \ell$, EJR imposes that certain $j \in N^*$ has to exist such that $|A_j \cap W| \geq \ell$ and therefore it has to be $|W \cap (\bigcup_{i\in N^*} A_i)| \geq \ell$. Therefore:

LEMMA 2 *EJR implies PJR.*

As discussed in [Aziz *et al.*, 2015] and also in the previous Section RAV does not satisfy JR, and therefore the following corollary follows immediately from Lemma 1:

COROLLARY 2 *RAV does not satisfy PJR.*

Similarly, in [Aziz *et al.*, 2015] it is proved that PAV satisfies EJR, and therefore, from Lemma 2 it follows:

COROLLARY 3 *PAV satisfies PJR.*

For $\mathbf{w}$-RAV we have the following result.

LEMMA 3 $\mathbf{w}$*-RAV does not satisfy PJR.*

Proof: For every vector $\mathbf{w} = (w_1, w_2, \dots)$ with $w_2 > 0$, a Theorem in [Aziz *et al.*, 2015] proves that $\mathbf{w}$-RAV does not satisfy JR. Therefore, in such cases $\mathbf{w}$-RAV does not satisfy PJR either.

The only case that remains is for $\mathbf{w_0} = (1, 0, \dots, 0)$. For $\mathbf{w_0}$ consider the following election. Let $k = 3$, $C = \{c_1, c_2, c_3, c_4\}$. 6 agents submit the following ballots: 4 agents submit $\{c_1, c_2\}$, one agent submits $\{c_3\}$ and the last agent submits $\{c_4\}$. The set of winners W that $\mathbf{w_0}$-RAV outputs for this election is: one of c_1 or c_2 plus c_3 and c_4. This violates PJR because according to PJR both $\{c_1\}$ and $\{c_2\}$ would have to be elected (let N^* be the four agents that submitted $\{c_1, c_2\}$, and $\ell = 2$, we have $|N^*| = 4 = 2\frac{n}{k}$ and $|\bigcap_{i\in N^*} A_i| = |\{c_1, c_2\}| = 2$). □

Finally, for $\mathbf{w}$-PAV we have the following result.

LEMMA 4 *For every weight vector $\mathbf{w} \neq (1, \frac{1}{2}, \frac{1}{3}, \dots)$, the rule $\mathbf{w}$-PAV does not satisfy PJR.*

For the sake of brevity we omit the details of the proof of this Lemma, but it can be derived directly from the equivalent proof for EJR in [Aziz *et al.*, 2015]. First, a Lemma in [Aziz *et al.*, 2015] proves that for a weight vector $\mathbf{w}$ such that $w_j > \frac{1}{j}$ for some $j > 1$, $\mathbf{w}$-PAV fails JR. Therefore, in such cases $\mathbf{w}$-PAV fails also PJR.

Second, for a weight vector $\mathbf{w}$ such that $w_j < \frac{1}{j}$ for some $j > 1$, another Lemma in [Aziz *et al.*, 2015] shows an election in which a given set of candidates C_0 with $|C_0| = j$ is approved by $j\frac{n}{k}$ agents but $\mathbf{w}$-PAV elects only $j - 1$ candidates of C_0. Such an election proves that for a weight vector $\mathbf{w}$ such that $w_j < \frac{1}{j}$ for some $j > 1$, $\mathbf{w}$-PAV does not satisfy PJR.

In summary, we have shown that while PJR has the advantages discussed at the beginning of this Section (Theorems 2 and 3 and corollary 1), it provides the same results for RAV, PAV, $\mathbf{w}$-RAV and $\mathbf{w}$-PAV as EJR: PAV satisfies both axioms and all the other rules fail to satisfy both PJR and EJR.

We have shown that any voting rule that satisfies EJR satisfies also PJR. It may be wondered whether voting rules that satisfy PJR but not EJR exist. The following (naive) voting rule solves this issue.

DEFINITION 3 **PR-PAV** *For every profile $\mathcal{A} = (A_1, \dots, A_n)$ and every target committee size k, if sets of candidates W that provide PR exist, then PR-PAV outputs (tied) all the sets of candidates W that provide PR. Otherwise, PR-PAV outputs the same sets of candidates as PAV.*

LEMMA 5 *PR-PAV satisfies PJR but fails EJR.*

Proof: By corollary 1, PR-PAV fails EJR. By Theorem 3, when sets of candidates W that provide PR exist for $(\mathcal{A}, k)$, PR-PAV outputs sets of candidates that provide PJR. By corollary 3 when sets of candidates W that provide PR do not exist, PR-PAV outputs sets of candidates that provide PJR. Finally, because PR-PAV always outputs sets of candidates that provide PJR, PR-PAV satisfies PJR. □

5 Complexity issues

For PR voting rules we have the following negative result.

THEOREM 4 *If P ≠ NP, then PR voting rules cannot be computed in polynomial time.*

Proof: We borrow a proof by Procaccia *et al.* [Procaccia *et al.*, 2008]. Following the ideas exposed in that proof we show a polynomial-time reduction from the EXACT 3-COVER (X3C) problem to the problem of determining a winning set of candidates with any PR voting rule. The X3C problem is known to be NP-Complete [Garey and Johnson, 1979]. Therefore such polynomial-time reduction proves the Theorem.

First, we repeat here the definition of the X3C problem given in [Procaccia *et al.*, 2008].

In the X3C problem we are given a set U of n points such that n is divisible by 3, and a collection of r subsets of U, $\mathcal{F} = \{F_1, \dots, F_r\}$, each of cardinality 3, i.e., for all j, $|F_j| =$

3. We are asked whether it is possible to find $n/3$ disjoint subsets in $\mathcal{F}$ such that their union covers the entire set U.

Given an instance of the X3C problem, we map it to a ballot profile, a set of candidates, and a target committee size as follows. U is the set of agents, $\mathcal{F}$ is the set of candidates, and $k = n/3$. Each agent $u \in U$ approves all the candidates F_i such that $u \in F_i$ in the original instance of the X3C problem. Observe that according to the definition of X3C n/k is the positive integer 3.

For any PR voting rule pick any winning set of candidates W that the voting rule outputs for the election described in the previous paragraph. If a set of candidates that provides PR exist, then by the definition of PR voting rule the winning set of candidates has to provide PR. If the winning set of candidates provides PR, it is possible to partition the set of agents in $n/3$ disjoint subsets of 3 agents, such that it is possible to assign each candidate w in W to one (and only one) subset so that all agents approve their assigned candidate w. This assignment solves the original X3C problem. If the winning set does not provide PR then the corresponding instance of the X3C problem does not have a solution.

For additional details we refer to [Procaccia *et al.*, 2008]. □

COROLLARY **4** *If P $\neq$ NP, then PR-PAV cannot be computed in polynomial time.*

PAV can neither be computed in polynomial time as shown by Aziz *et al.* [Aziz *et al.*, 2014].

Observe that Theorem 4 does not at all imply that any approval-based voting rule that satisfies PJR has to be non-polynomial. Observe also that Theorem 4 does not even imply that it would be more difficult to find a polynomial voting rule that satisfies PJR than to find a polynomial voting rule that satisfies EJR. Since EJR implies PJR (Lemma 2) any approval-based voting rule that satisfies EJR and can be computed in polynomial time (if such rule exists) would be also an approval-based voting rule that satisfies PJR and can be computed in polynomial time. All that we know is that any voting rule that can be computed in polynomial time and that satisfies PJR (or EJR) will output for some elections winning sets of candidates that do not provide PR even if sets of candidates that provide PR exist.

6 Discussion, conclusions and future lines of work

The development of axioms that capture the idea of representation and the study of multi-winner voting rules with such axioms is of great interest. We consider that justified representation is a very interesting axiom, and that it must be a necessary requirement (maybe not sufficient) for considering the use of an approval-based multi-winner voting rule when it is desired that the winning set represents the different opinions or preferences of the agents involved in the election.

As discussed in [Aziz *et al.*, 2015] it is useful to know for which values of k a multi-winner voting rule satisfies JR. For that reason, we have proved that RAV satisfies JR for $k \leq 5$ and fails it for $k \geq 6$. Therefore, for elections with a relatively small number of seats to allocate (≥ 6) RAV should not be used when JR is considered a desirable property.

However, we believe that the definition of extended justified representation is much more questionable. We have shown that a set of candidates can provide perfect representation for a given profile $\mathcal{A}$ and target committee size k, and despite that, fail to provide EJR. We have proposed an alternative axiom, which we have called proportional justified representation and we have shown that it does not suffer from this problem and that it provides the same results than EJR for all the multi-winner voting rules that were analyzed in [Aziz *et al.*, 2015][3].

It can be wondered which is the point in proposing a new axiom that provides the same results as EJR for almost all the multi-winner voting rules that have been analyzed (except PR-PAV). We insist on our argument: we believe that it is not reasonable to disqualify a multi-winner voting rule (from the point of view of representation) because it outputs in certain cases sets of candidates that do not provide EJR, if in such cases the winning sets provide PR.

We believe that another argument for PJR is that it captures better the idea of when a large enough group of agents should be allocated several representatives, because for PJR it is enough that the required number of representatives is achieved counting all the candidates that are approved at least by one agent in the group, while EJR requires one agent in the group to approve the required number of representatives in the winning set. Moreover, we believe that this requirement that "one agent in the group approves the required number of representatives in the winning set" is unjustified: one should expect that each candidate in the winning set should represent many agents, and not that an agent is represented by several candidates. Of course, when a group of agents agree in approving the same set of candidates, and the group is large enough to deserve several representatives, both PJR and EJR require that *each* of the agents in such group is represented by several (and the same) representatives.

As a final remark, we agree with the authors of [Aziz *et al.*, 2015] in that it would be very interesting to identify a voting rule that satisfies PJR (or EJR) but can be computed efficiently (in polynomial time). Unfortunately, the only two rules that have been identified so far that satisfy PJR are PAV and PR-PAV, and none of them can be computed in polynomial time. Voting rules that cannot be computed in polynomial time are interesting from a theoretical point of view. For instance, the existence of PAV and PR-PAV is a proof that it is possible to develop multi-winner voting rules that satisfy PJR. However, as discussed in [Bartholdi III *et al.*, 1989], voting rules that cannot be computed in polynomial time are mostly useless in practice: they can be used only in very small elections. We remark that the results that we have presented in this paper do not at all exclude the possibility to find a voting rule that satisfies PJR and can be computed in polynomial time. We plan to address this issue in our future work.

[3] It should be noted that in [Aziz *et al.*, 2015] other approval-based multi-winner voting rules (surveyed in [Kilgour, 2010]) are analyzed, in addition to PAV, RAV and their variations. All these other voting rules fail JR, and therefore, they also fail EJR and PJR.

References

[Aziz *et al.*, 2014] Haris Aziz, Serge Gaspers, Joachim Gudmundsson, Simon Mackenzie, Nicholas Mattei, and Toby Walsh. Computational aspects of multi-winner approval voting. In *Eighth Multidisciplinary Workshop on Advances in Preference Handling (MPREF 2014)*, Quebec City, Canada, July 2014. arXiv preprint arXiv:1407.3247.

[Aziz *et al.*, 2015] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. Justified representation in approval-based committee voting. In *Twenty-Ninth AAAI Conference*. AAAI Press, January 2015.

[Bartholdi III *et al.*, 1989] John Bartholdi III, Craig A Tovey, and Michael A Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and welfare*, 6(2):157–165, 1989.

[Betzler *et al.*, 2013] Nadja Betzler, Arkadii Slinko, and Johannes Uhlmann. On the computation of fully proportional representation. *Journal of Artificial Intelligence Research*, 47:475–519, 2013.

[Black, 1958] Duncan Black. *The theory of committees and elections*. Cambridge University Press, 1958.

[Conitzer, 2010] Vincent Conitzer. Making decisions based on the preferences of multiple agents. *Communications of the ACM*, 53(3):84–94, 2010.

[Dummet, 1984] Michael Dummet. *Voting Procedures*. Oxford University Press, 1984.

[Elkind *et al.*, 2011] Edith Elkind, Jérôme Lang, and Abdallah Saffidine. Choosing collectively optimal sets of alternatives based on the Condorcet criterion. In *Twenty-Second International Joint Conference on Artificial Intelligence*, pages 186–191. AAAI Press, July 2011.

[Elkind *et al.*, 2014] Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. Properties of multiwinner voting rules. In *International Conference on Autonomous Agents and Multiagent Systems*, pages 53–60. International Foundation for Autonomous Agents and Multiagent Systems, May 2014.

[Garey and Johnson, 1979] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Series of books in the mathematical sciences. W. H. Freeman, 1979.

[Kilgour, 2010] D. Marc. Kilgour. Approval balloting for multi-winner elections. In Jean-Franois Laslier and M. Remzi Sanver, editors, *Handbook on Approval Voting*, pages 105–124. Springer, 2010.

[Monroe, 1995] Burt L. Monroe. Fully proportional representation. *The American Political Science Review*, 89(4):925–940, 1995.

[Procaccia *et al.*, 2008] Ariel D. Procaccia, Jeffrey S. Rosenschein, and Aviv Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30(3):353–362, 2008.

[Sandholm, 1999] T. W. Sandholm. Distributed rational decision making. In G. Weiss, editor, *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*, pages 201–258. The MIT Press, 1999.

[Skowron *et al.*, 2013] Piotr Skowron, Piotr Faliszewski, and Arkadii Slinko. Fully proportional representation as resource allocation: Approximability results. In *Twenty-Third International Joint Conference on Artificial Intelligence*, pages 353–359. AAAI Press, August 2013.