

Approximation Algorithms and Mechanism Design for Minimax Approval Voting¹

Ioannis Caragiannis, Dimitris Kalaitzis, and Evangelos Markakis

Abstract

We consider approval voting elections in which each voter votes for a (possibly empty) set of candidates and the outcome consists of a set of k candidates for some parameter k , e.g., committee elections. We are interested in the minimax approval voting rule in which the outcome represents a compromise among the voters, in the sense that the maximum distance between the preference of any voter and the outcome is as small as possible. This voting rule has two main drawbacks. First, computing an outcome that minimizes the maximum distance is computationally hard. Furthermore, any algorithm that always returns such an outcome provides incentives to voters to misreport their true preferences.

In order to circumvent these drawbacks, we consider approximation algorithms, i.e., algorithms that produce an outcome that approximates the minimax distance for any given instance. Such algorithms can be considered as alternative voting rules. We present a polynomial-time 2-approximation algorithm that uses a natural linear programming relaxation for the underlying optimization problem and deterministically rounds the fractional solution in order to compute the outcome; this result improves upon the previously best known algorithm that has an approximation ratio of 3. We are furthermore interested in approximation algorithms that are resistant to manipulation by (coalitions of) voters, i.e., algorithms that do not motivate voters to misreport their true preferences in order to improve their distance from the outcome. We complement previous results in the literature with new upper and lower bounds on strategyproof and group-strategyproof algorithms.

1 Introduction

Approval voting is a very popular voting protocol mainly used for committee elections [2]. In such a protocol, the voters are allowed to vote for, or approve of, as many candidates as they like. In the last three decades, many scientific societies and organizations have adopted approval voting for their council elections. The solution concept that has been used in almost all such elections in practice is the minisum solution, *i.e.*, output the committee which, when seen as a 0/1-vector, minimizes the sum of the Hamming distances to the ballots. We assume throughout the paper that the committee should be of some predefined size k . Then the minisum solution consists of the k candidates with the highest number of approvals.

This solution however may ignore some voters' preferences in certain instances and does not take fairness issues into account. We demonstrate this with the following example with four voters, five candidates, and $k = 2$. Each row represents the preference of the corresponding voter. The minisum solution contains the candidates $\{a, b\}$. The distances of the voters from this outcome are 1, 0, 2, and 5 for voters 1, 2, 3, and 4, respectively (counting the number of alternatives in which the voter disagrees with the outcome). Instead, the solution $\{a, c\}$ has distances 3, 2, 2, and 3, respectively, and suggests a better compromise among the voters since everybody is relatively close to the outcome.

¹A slightly different version of the paper appeared in *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI 10)*. This work was partially supported by the European Union under the Euro-NF Network of Excellence and Cost Action IC0602 "Algorithmic Decision Theory", and by a "Caratheodory" basic research grant from the University of Patras.

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
1	1	1	0	0	1
2	1	1	0	0	0
3	1	1	1	1	0
4	0	0	1	1	1

Recently, a new voting rule, the *minimax solution*, was introduced as a means to achieve a compromise between the voters' preferences [3]. The minimax solution picks the k candidates for which the maximum (Hamming) distance of any voter from the outcome is minimized. Since this rule minimizes the disagreement with the least satisfied voter, it tends to result in outcomes that are more widely acceptable than the minisum solution. On the negative side the minimax solution has two main drawbacks that prevent its applicability: (i) the problem of computing the minimax solution is NP-hard, and (ii) voters may have incentives to misreport their preference in order to improve the distance of their true preference from the outcome. Our main goal in this paper is to tackle these issues by resorting to approximation algorithms.

Approximation algorithms tackle the computational hardness of an optimization problem by producing (in polynomial-time) solutions provably close to optimal ones for any problem instance; see [9] for a coverage of early work in the field. We refer to the optimization problem of computing the minimax solution as k -minimax approval. [4] present a 3-approximation algorithm for the problem; given an instance, the algorithm produces a solution (i.e., a set of k candidates) so that its distance from any voter's preference is at most 3 times the maximum distance of the voters from the minimax solution. The algorithm is very simple to describe and we will refer to it here as the k -completion algorithm: it arbitrarily picks a voter and computes a set of k candidates which has minimum distance from this voter. An immediate question is whether algorithms with better approximation ratios exist. Another interesting question is whether we can have good approximations by non-dictatorial algorithms. Note that the k -completion algorithm is dictatorial as it is based only on one voter's preferences.

The issue of resistance to manipulation is the very subject of Mechanism Design; see [6] for an introduction to the field. In our context, it translates to algorithms for k -minimax approval which, given a profile, compute an approximate solution in such a way that no single voter or a coalition of voters have any incentive to misreport their preferences in order to decrease their distance from the outcome. The corresponding properties of resistance to manipulation by single voters and coalitions of voters are known as strategyproofness and group-strategyproofness, respectively. [4] prove that the minimax solution is not resistant to manipulation while the k -completion algorithm is. They also pose the question of computing the best possible bound on the approximation ratio of algorithms that are resistant to manipulation. This question falls within the line of research on mechanisms without monetary transfers [8] and, in particular, approximate mechanism design without money [7].

We make progress in both directions. Concerning the approximability of k -minimax approval by polynomial-time algorithms, we first establish a connection between the property of Pareto-efficiency and approximability. As a corollary, we obtain that Minisum (i.e., the algorithm that returns a minisum solution) has approximation ratio at most $3 - \frac{2}{k+1}$ for k -minimax approval. Our strongest result in this direction is an algorithm based on linear programming that achieves an improved approximation ratio of 2; this is a significant improvement compared to the previously best known algorithms. The algorithm is based on rounding the fractional solution of a natural linear programming relaxation for k -minimax approval. This result is the best possible that can be obtained using the particular LP relaxation which has an integrality gap of 2.

In the direction of algorithms resistant to manipulation, we observe that a variation of Minisum is strategyproof and present a Pareto-efficient refinement of the k -completion algorithm. Due to Pareto-efficiency, the latter algorithm has approximation ratio $3 - \frac{2}{k+1}$ as well. We also present the first inapproximability results for algorithms that are resistant to manipulation, making progress on the question posed in [4]. In particular, we present a lower bound of $2 - \frac{2}{k+1}$ on the approximation ratio of any strategyproof algorithm and a negative result which states that a slightly stronger notion of group-strategyproofness cannot be achieved by algorithms with approximation ratio different than $3 - \frac{2}{k+1}$ and infinity. Our lower bounds are not based on any computational complexity assumption and, hence, hold for exponential-time algorithms as well.

2 Notation and Definitions

We fix some notation used in the following. We typically use n to denote the number of voters and m for the number of candidates. We denote the set of candidates by A . A preference is simply a subset of A . A profile P is a tuple $P = (P_1, \dots, P_n)$ where P_i denotes the preference of voter i (i.e., the set of candidates she approves). Throughout the paper we make the reasonable assumption that $n > k$. When this is not explicitly mentioned (e.g., in some lower bound proofs), we can complete the profile by adding indifferent voters (that approve no candidate).

We extend the notion of (Hamming) distance to subsets of A as follows. We say that the distance of two sets Q and T is the total number of candidates in which they differ, i.e.,

$$d(Q, T) = |Q \setminus T| + |T \setminus Q| = |Q| + |T| - 2|Q \cap T|.$$

Note that this is precisely the Hamming distance of the sets, when seen as binary vectors where the i th coordinate of each vector equals 1 if the i th candidate belongs to the set.

3 Approximation Algorithms

We begin by establishing a connection between Pareto-efficiency and low approximation ratio.

Definition 1. Given a profile P , a size- k set $K \subseteq A$ is called Pareto-efficient with respect to P if there is no other size- k set $K' \subseteq A$ such that $d(K', P_{i^*}) < d(K, P_{i^*})$ for some voter i^* and $d(K', P_i) \leq d(K, P_i)$ for any other voter i . An algorithm for k -minimax approval is Pareto-efficient if, on any input profile P , its outcome is Pareto-efficient with respect to P .

The next lemma significantly extends the class of 3-approximation algorithms for minimax-approval and will be proved very useful later. Interestingly, Minisum is Pareto-efficient; the proof follows by the definition of Pareto-efficiency and the fact that Minisum minimizes the sum of the distances of the outcome from the voters.

Lemma 2. *Any Pareto-efficient algorithm for k -minimax approval has approximation ratio at most $3 - \frac{2}{k+1}$.*

Proof. Let P be a profile and let O and K be the minimax solution and the outcome returned by a non-optimal Pareto-efficient algorithm on input P . Let $OPT = \max_i \{d(O, P_i)\}$. We will show that $d(K, P_i)/OPT \leq 3 - \frac{2}{k+1}$ for every voter i .

First assume that $OPT \geq k + 1$. Then, by applying the triangle inequality, we obtain

$$\begin{aligned} \frac{d(K, P_i)}{OPT} &\leq \frac{d(K, O) + d(O, P_i)}{OPT} \\ &\leq 1 + 2k/OPT \leq 3 - \frac{2}{k+1} \end{aligned}$$

for each voter i . The second inequality follows since the distance of any two size- k sets is at most $2k$ and $d(O, P_i) \leq OPT$.

Now, assume that $OPT < k + 1$. Since the solution returned by the algorithm is non-optimal for the particular profile P , there exists a voter i^* such that $d(K, P_{i^*}) < d(O, P_{i^*})$. Indeed, if this was not the case, then K would not be Pareto-efficient with respect to P . By the definition of the distance, we observe that $d(K, P_{i^*})$ has the same parity with $d(O, P_{i^*})$, and the above argument implies that

$$d(K, P_{i^*}) \leq d(O, P_{i^*}) - 2.$$

Now, using this observation and by applying the triangle inequality twice, we have

$$\begin{aligned} \frac{d(K, P_i)}{OPT} &\leq \frac{d(K, P_{i^*}) + d(P_{i^*}, P_i)}{OPT} \\ &\leq \frac{d(K, P_{i^*}) + d(O, P_{i^*}) + d(O, P_i)}{OPT} \\ &\leq \frac{2d(O, P_{i^*}) + d(O, P_i) - 2}{OPT} \\ &\leq 3 - \frac{2}{OPT} \leq 3 - \frac{2}{k+1} \end{aligned}$$

for any voter i . This completes the proof. \square

We now present an algorithm based on linear programming. On an input profile P , the algorithm uses the following equivalent integer linear program for k -minimax approval.

$$\begin{aligned} &\text{minimize} && q \\ &\text{subject to:} && \forall i \in N, q + 2 \sum_{a \in P_i} x_a \geq k + |P_i| \\ & && \sum_{a \in A} x_a = k \\ & && \forall a \in A, x_a \in \{0, 1\} \\ & && q \geq 0 \end{aligned}$$

The variable x_a denotes whether candidate a is included in the solution ($x_a = 1$) or not ($x_a = 0$). The first constraint essentially lower-bounds the value of variable q by the maximum distance of a voter from the size- k set that consists of the candidates included in the solution. The LP-based algorithm solves the LP relaxation in which the integrality constraint has been relaxed to $0 \leq x_a \leq 1$. In this way, a fractional solution is obtained with the x -variables having values in $[0, 1]$. Then, the algorithm includes the candidates with the k largest x -variables in the final solution (by breaking ties arbitrarily).

Theorem 3. *The LP-based algorithm has approximation ratio at most 2.*

Proof. Consider the application of the LP-based algorithm on a profile P . Denote by (q^*, x^*) the optimal fractional solution of the LP and let K be the outcome of the LP-based algorithm. We will show that, for each voter i , her preference P_i has distance at most $2q^*$ from the set K . Since q^* is a lower bound on the cost of the optimal integral solution for the particular instance of k -minimax approval, we will have obtained the desired 2-approximation bound.

Denote by Y_i the set of candidates in the preference of voter i that belong to the set K , i.e., $Y_i = P_i \cap K$. Let j be a voter whose preference P_j has maximum distance from K . The

first constraint of the LP implies that

$$q^* \geq k + |P_j| - 2 \sum_{a \in P_j} x_a$$

and, using the fact that the x -variables of the LP are upper-bounded by 1 (due to the third LP constraint), we obtain that $q^* \geq |k - |P_j||$. Observe that if $|Y_j| = \min\{k, |P_j|\}$, then $d(K, P_j) = |k - |P_j||$, i.e., the solution of the algorithm is optimal in this case. So, in the following, we assume that $|Y_j| < \min\{k, |P_j|\}$.

For the sake of contradiction, assume that $d(K, P_j) > 2q^*$. By the definition of distance and the first LP constraint, we obtain

$$k + |P_j| - 2|Y_j| > 2q^* \geq 2 \left(k + |P_j| - 2 \sum_{a \in P_j} x_a^* \right)$$

and, equivalently,

$$0 > k + |P_j| + 2|Y_j| - 4 \sum_{a \in P_j} x_a^*. \quad (1)$$

Since none of the candidates in $P_j \setminus Y_j$ was selected in the solution, this means that the x -variables corresponding to the $k - |Y_j|$ candidates in $K \setminus Y_j$ are not smaller than any x -variable corresponding to a candidate in $P_j \setminus Y_j$, i.e., for each candidate a in $K \setminus Y_j$, it holds that $x_a^* \geq \max_{a' \in P_j \setminus Y_j} \{x_{a'}^*\}$. By summing over all candidates in $K \setminus Y_j$, we have

$$\begin{aligned} \sum_{a \in K \setminus Y_j} x_a^* &\geq (k - |Y_j|) \max_{a' \in P_j \setminus Y_j} \{x_{a'}^*\} \\ &\geq \frac{k - |Y_j|}{|P_j| - |Y_j|} \sum_{a' \in P_j \setminus Y_j} x_{a'}^*. \end{aligned} \quad (2)$$

By the definition of set Y_j , we have that every candidate of $K \setminus Y_j$ also belongs to $A \setminus P_j$. Hence

$$\sum_{a \in A \setminus P_j} x_a^* \geq \sum_{a \in K \setminus Y_j} x_a^*. \quad (3)$$

Furthermore, using the third LP constraint, we have

$$\sum_{a \in P_j \setminus Y_j} x_a^* = \sum_{a \in P_j} x_a^* - \sum_{a \in Y_j} x_a^* \geq \sum_{a \in P_j} x_a^* - |Y_j|. \quad (4)$$

Putting (2), (3), and (4) together, we have

$$\sum_{a \in A \setminus P_j} x_a^* \geq \frac{k - |Y_j|}{|P_j| - |Y_j|} \sum_{a \in P_j} x_a^* - \frac{|Y_j|(k - |Y_j|)}{|P_j| - |Y_j|}.$$

Now, observe that the left hand side in the above inequality satisfies (due to the second LP constraint)

$$\sum_{a \in A \setminus P_j} x_a^* = \sum_{a \in A} x_a^* - \sum_{a \in P_j} x_a^* = k - \sum_{a \in P_j} x_a^*.$$

Hence, the above inequality yields

$$k - \sum_{a \in P_j} x_a^* \geq \frac{k - |Y_j|}{|P_j| - |Y_j|} \sum_{a \in P_j} x_a^* - \frac{|Y_j|(k - |Y_j|)}{|P_j| - |Y_j|},$$

and, equivalently,

$$\sum_{a \in P_j} x_a^* \leq \frac{k|P_j| - |Y_j|^2}{k + |P_j| - 2|Y_j|}. \quad (5)$$

Now, (1) and (5) yield to the following contradiction:

$$\begin{aligned} 0 &> k + |P_j| + 2|Y_j| - 4 \frac{k|P_j| - |Y_j|^2}{k + |P_j| - 2|Y_j|} \\ &= \frac{(k - |P_j|)^2}{k + |P_j| - 2|Y_j|} \geq 0, \end{aligned}$$

We conclude that $d(K, P_j) \leq 2q^*$ as desired. \square

Given that the rounding in the LP-based algorithm is performed in an extremely simple way, one might hope that a more clever rounding could yield an improved algorithm. Unfortunately, the particular LP relaxation has an integrality gap of 2 and well-known arguments from the theory of approximation algorithms [9] imply that this is the best possible bound that can be obtained using the particular LP relaxation. Consider a profile with at least $2k$ candidates and denote by A' a size- $2k$ set of candidates. There are sufficiently many voters so that each one approves a different set of k candidates from A' . Clearly, for any k -size subset Q of A' , there exists a voter whose preference does not include any of the candidates in Q . Hence, the minimax solution on the particular instance has cost at least $2k$. The claim follows by observing that the solution with the x -variables set to $1/2$ and $q = k$ satisfies the constraints of the LP relaxation.

4 Resistance to Manipulation

Let us first formally define strategyproofness in our setting. Given a profile P and an algorithm R , we denote by $R(P)$ the outcome of the algorithm on profile P . We also denote by P_{-i} the preferences of all voters besides i . Hence, we can also write P as (P_{-i}, P_i) . Strategyproofness means that no voter i has an incentive to unilaterally change her preference so as to reduce the distance of P_i from the outcome of the algorithm.

Definition 4. An algorithm R is strategyproof (SP) if for any voter i , for any profile P , and for any $P'_i \subseteq A$:

$$d(P_i, R(P_{-i}, P_i)) \leq d(P_i, R(P_{-i}, P'_i)).$$

We begin with an example demonstrating that the minimax solution is not SP. Consider the profile at the left table below with $k = 2$; a similar example is presented in [4]. In this

	a	b	c			a	b	c
1	1	1	0		1	1	1	0
2	0	1	1		2	0	0	1
3	0	1	0		3	0	1	0

profile, the sets $\{a, b\}$ and $\{b, c\}$ are those with distance at most 2 from all voters. Assume

that $\{a, b\}$ is the minimax solution returned for the particular profile (the other case is symmetric). Now, assume that voter 2 has $\{c\}$ as her preference (see the right table). Now, the only set that has distance at most 2 from each voter's preference is $\{b, c\}$, i.e., exactly the preference of voter 2 in the first profile. This implies that voter 2 has an incentive to misreport her preference as $\{c\}$ instead of $\{b, c\}$ and demonstrates that minimax is not SP. The same example can show that the LP-based algorithm is not SP either.

Note that both solutions mentioned above are minisum solutions as well. This implies that Minisum is not SP in general. However, we can introduce a simple tie-breaking rule which assigns distinct ids to the candidates and ties for the last positions of the outcome are resolved by selecting the candidates with the smallest id. Then, Minisum equipped with the smallest-id-first tie-breaking rule can be easily proved to be strategyproof. Note that the particular assumption on the way ties are broken does not affect the Pareto-efficiency of Minisum. We summarize the discussion on Minisum to the following statement. Compared to the k -completion algorithm, Minisum is certainly non-dictatorial.

Theorem 5. *Minisum with the smallest-id tie-breaking rule is SP and has approximation ratio at most $3 - \frac{2}{k+1}$ for k -minimax approval.*

Let us remark here that the fact that a variation of Minisum is SP indicates that k -minimax approval is sufficiently restricted as a setting since well-known impossibility results state that, in general, strategyproofness is only achievable by dictatorial algorithms; see [6].

In the following, we present a lower bound on the approximation ratio of SP algorithms. We outline the main argument with the following example with $k = 1$ (we essentially adapt to our model an argument used in [7] in a slightly different context).

Consider the application of an SP algorithm on the following profile with $k = 1$. Without

	a_1	a_2	a_3	a_4	a_5	a_6
1	1	1	1	0	0	0
2	0	0	0	1	1	1

loss of generality, let $\{a_1\}$ be the outcome of the algorithm for this profile (the other cases can be handled symmetrically). Now consider the profile below. Again, the outcome should

	a_1	a_2	a_3	a_4	a_5	a_6
1	1	0	0	0	0	0
2	0	0	0	1	1	1

be the same otherwise voter 1 would have an incentive to misreport her preference from $\{a_1\}$ to $\{a_1, a_2, a_3\}$ and improve her distance from the outcome returned by the algorithm; this would violate strategyproofness. The maximum distance in the second profile is 4. The minimax solution approves one of the three rightmost candidates and has maximum distance 2. Hence, the approximation ratio is 2 in this case.

The extension of this argument for higher values of k yields a slightly weaker lower bound.

Theorem 6. *Any SP algorithm for k -minimax approval has approximation ratio at least $2 - \frac{2}{k+1}$.*

Proof. Consider a profile with $m \geq 4k$ candidates and two voters 1 and 2 that approve the disjoint size- $2k$ sets P_1 and P_2 , respectively. Let K be the outcome of an SP algorithm on this particular profile. Assume that $P_1 \cap K \leq k/2$ (the other case is handled similarly). Now, consider the profile in which voter 1 approves the set P_1 and voter 2 approves the set K . We argue that the outcome of the algorithm is again K . Indeed, if this was not

the case and the outcome was a set $K' \neq K$, voter 2 would have an incentive to misreport her preference as P_2 instead of K in order to decrease the distance of her true preference from the outcome. The distance of voters 1 and 2 from the outcome in the second profile is $d(K, P_1) = 3k - 2|K \cap P_1|$ and 0, respectively.

Let t be an integer such that

$$\frac{3k - 2|K \cap P_1| - 2}{4} \leq t \leq \frac{3k - 2|K \cap P_1| + 2}{4}.$$

Since $|K \cap P_1| \leq k/2$ and $|P_1| = 2k$, it holds that $t \leq |P_1 \setminus K|$. Consider the size- k set O which consists of the alternatives in $K \cap P_1$, t alternatives from $P_1 \setminus K$, and $k - |K \cap P_1| - t$ alternatives from $K \setminus P_1$. We have

$$d(O, K) = 2t \leq \frac{3k - 2|K \cap P_1| + 2}{2}$$

and

$$d(O, P_1) = 3k - 2|K \cap P_1| - 2t \leq \frac{3k - 2|K \cap P_1| + 2}{2}.$$

Hence, the approximation ratio of the algorithm for the second profile is at least

$$\begin{aligned} \frac{3k - 2|K \cap P_1|}{\max\{d(O, K), d(O, P_1)\}} &\geq 2 - \frac{4}{3k - 2|K \cap P_1| + 2} \\ &\geq 2 - \frac{2}{k+1}. \end{aligned}$$

The last inequality follows since $|K \cap P_1| \leq k/2$. \square

We now move to stronger notions of resistance to manipulation. For a set (or coalition) of voters S , we denote by P_{-S} the preferences of the voters not in S .

Definition 7. An algorithm R is group-strategyproof (GSP) if for any coalition S of voters, and for any profile P , there is no profile P'_S of the voters in S such that:

$$d(P_i, R(P_{-S}, P_S)) > d(P_i, R(P_{-S}, P'_S)) \quad \forall i \in S.$$

It is not hard to see that Minisum is not GSP. In contrast, the k -completion algorithm can be easily implemented so that it is GSP. The reason for this is that a coalition that does not contain the dictator cannot affect the outcome and the dictator has no incentive to participate in any coalition since her distance from the outcome is anyway minimum. We present a refinement of the k -completion algorithm which can be proved to be simultaneously GSP and Pareto-efficient. Then, Lemma 2 implies that its approximation ratio is at most $3 - \frac{2}{k+1}$. The algorithm uses an ordering of the voters with the dictator being first and an ordering of the candidates. Now, we can think of a candidate a as a binary vector z_a such that the i -th coordinate of the vector is 1 if voter i approves candidate a and 0 otherwise. For each candidate a , it computes its score as

$$\text{sc}(a) = \sum_{i=1}^n z_a(i) \cdot 2^{n-i}$$

and picks the k candidates with highest scores by breaking ties according to the candidate ordering.

The Pareto-efficiency and strategyproofness of this algorithm become apparent by the following interpretation of its execution. Initially, it considers all possible size- k sets as

possible outcomes. Among them, it keeps the ones that have the same minimum distance from the preference of voter 1. Then, among them, it keeps the ones that have the same minimum distance from the preference of voter 2, and so on. After considering voter n , it returns as an outcome one among the sets kept at that point.

Our last result concerns a stronger definition of group-strategyproofness.

Definition 8. An algorithm R is strongly group-strategyproof (strongly GSP) if for any coalition S of voters, and for any profile P , there is no profile P'_S of the voters in S such that:

$$d(P_i, R(P_{-S}, P_S)) \geq d(P_i, R(P_{-S}, P'_S)) \quad \forall i \in S$$

with strict inequality for at least one voter of S .

The rationale behind this concept is that we demand the algorithm to be resistant to coalitions in which some voters may change their preference profile in order to help other members of the coalition (without necessarily gaining something for themselves). We make a connection between Pareto-efficiency and strong group-strategyproofness. We show that the former property is necessary in order to guarantee the existence of good approximation algorithms satisfying the latter. Of course, it is not sufficient. For example, minisum is Pareto-efficient but not even GSP. We also point out that this property is not necessary for group-strategyproofness since there are implementations of the k -completion algorithm that are not Pareto-efficient.

Lemma 9. *Any strongly GSP algorithm for k -minimax approval that has finite approximation ratio is Pareto-efficient.*

Proof. Consider a strongly GSP algorithm with finite approximation ratio. First observe that in each profile in which all voters approve the same set S of k candidates, the algorithm must return S as the outcome. If this is not the case for some profile of this kind, then the approximation ratio would be infinite.

Assume now that the algorithm returns a size- k set K on some profile which is not Pareto-efficient. Then, there exists another size- k set K' such that $d(K', P_{i^*}) < d(K, P_{i^*})$ for some voter i^* and $d(K', P_i) \leq d(K, P_i)$ for any other voter i . Now, the voters have an incentive to misreport the set K' and improve their distance from the outcome. \square

Lemmas 2 and 9 imply that if a strongly GSP algorithm with finite approximation ratio exists, then it must have approximation ratio at most $3 - \frac{2}{k+1}$. We complement this corollary with the following tight lower bound.

Theorem 10. *Any strongly GSP algorithm for k -minimax approval has approximation ratio at least $3 - \frac{2}{k+1}$.*

Proof. Consider an algorithm with approximation ratio strictly better than $3 - \frac{2}{k+1}$. We will actually prove that it is manipulable by two voters. Consider the profile with $3k + 1$ candidates and $3k + 1$ voters in which the preference of voter i contains only candidate i . Denote by K the outcome of the algorithm for the particular profile. Let i^* be a voter that has a candidate not in K in her preference and consider the profile in which voter i^* approves the $2k + 1$ candidates outside K . Now, since the algorithm has approximation ratio strictly better than $3 - 2/(k+1)$, the outcome for the new profile should include a candidate i' not in K . Hence, voters i^* and i' have an incentive to manipulate the algorithm; voter i^* misreports her preference and does not decrease her distance and voter i' strictly decrease her distance from the outcome. \square

Together with the above discussion, Theorem 10 leads to the following interesting statement.

Corollary 11. *Strongly GSP algorithms for k -minimax approval have at most two possible values for their approximation ratio: it can be either exactly $3 - \frac{2}{k+1}$ or infinity.*

5 Discussion

As a conclusion, let us discuss an interesting (but not obvious at first glance) relation of k -minimax approval to facility location problems; see [8, 9] and the references therein. In facility location, we are given agents located at the nodes of a network and the objective is to locate a facility at a node so that the maximum distance of any agent to the facility is minimized. k -minimax approval can be thought of as a facility location problem on a hypercubic network. Recall that a hypercube of dimension m has 2^m nodes each associated with a distinct 0/1 vector. An edge connects two nodes if their vectors differ in exactly one coordinate. So, k -minimax approval on a profile with n voters and m candidates can be thought of as a facility location instance with n agents (corresponding to the voters) located at some nodes of a hypercube of dimension m (the vector of such a node corresponds to the preference of a voter) with the objective being to put a facility to a node with exactly k 1s in its vector (corresponding to a size- k set of candidates) so that the maximum distance of any agent from the facility is minimized.

Besides this relation, the restriction on the type of nodes where the facility can be placed differentiates significantly k -minimax approval from standard facility location so that the best known approximation algorithm (implicit in [5]) for facility location on the hypercube does not carry over to our model. Furthermore, from the resistance to manipulation viewpoint, an important property of the standard facility location setting is single-peakedness in the agents preferences in the sense that the location of the agent is her mostly preferred location for the facility. This property does not hold in our model as there may be several among the possible locations an agent may prefer the most. A consequence of this peculiarity is that strategyproofness does not imply group-strategyproofness in k -minimax approval, in contrast to what is the case for single-peaked agent preferences [1] in facility location settings. We have demonstrated this when we observed that (a variation of) Minisum is SP but not GSP.

Our work leaves several challenging questions open. Concerning the approximability of k -minimax approval there is no known lower bound on the approximation ratio of polynomial-time algorithms besides the NP-hardness of the problem. It is interesting either to find such a lower bound or obtain a polynomial-time approximation scheme (PTAS), i.e., an algorithm that can achieve an approximation guarantee $1 + \epsilon$ for any constant $\epsilon > 0$ at the expense of a (possibly exponential) dependence of its running time on $1/\epsilon$. Progress in either direction will significantly improve our understanding of k -minimax approval. Experimental results in [4] provide evidence that local-search algorithms might have very low approximation ratios. Interestingly, we have a lower bound (very close to 3) for a natural and broad class of local-search algorithms that includes the ones considered in that paper; details will appear in the final version of the paper. As far as resistance to manipulation is concerned, our work leaves an intriguing gap between the upper bound of $3 - \frac{2}{k+1}$ and the lower bound of $2 - \frac{2}{k+1}$ on the approximation ratio of SP or GSP algorithms for k -minimax approval when $k \geq 2$. Furthermore, detecting whether strongly GSP algorithms with finite approximation ratio exist or not is of interest; here, we have made several unsuccessful attempts in both directions.

References

- [1] Barberà, S.; Berga, D; and Moreno, B. 2010. Individual versus group strategy-proofness: When do they coincide? *Journal of Economic Theory*, in press.
- [2] Brams, S. J.; Fishburn, P. C. 2007. *Approval Voting*, 2nd edition, Springer.
- [3] Brams, S. J.; Kilgour, D. M.; and Sanver, M. R. 2007. A minimax procedure for electing committees. *Public Choice*, 132(3-4): 401-420.
- [4] LeGrand, R.; Markakis, E.; and Mehta, A. 2007. Some results on approximating the minimax solution in approval voting. *Proceedings of 6th AAMAS*, 1193-1195.
- [5] Li, M.; Ma, B.; and Wang, S. 1999. Finding similar regions in many strings. *Proceedings of 31st ACM STOC*, 473-482.
- [6] Nisan, N. 2007. Introduction to mechanism design (for computer scientists). *Algorithmic Game Theory*, Cambridge University Press, 209-241.
- [7] Procaccia, A. D.; and Tennenholtz, M. 2009. Approximate mechanism design without money. *Proceedings of 10th ACM EC*, 177-186.
- [8] Schummer J.; and Vohra, R. V. 2007. Mechanism design without money. *Algorithmic Game Theory*, Cambridge University Press, 209-241.
- [9] Vazirani, V. V. 2001. *Approximation algorithms*. Springer.

Ioannis Caragiannis
Department of Computer Engineering
and Informatics & RACTI
University of Patras, Greece
Email: caragian@ceid.upatras.gr

Dimitris Kalaitzis
Department of Computer Engineering
and Informatics & RACTI
University of Patras, Greece
Email: kalaitzis@cti.gr

Evangelos Markakis
Department of Informatics
Athens University of Economics
& Business, Greece
Email: markakis@gmail.com