Fully Proportional Representation as Resource Allocation: Approximability Results

Piotr Skowron
University of Warsaw
Warsaw, Poland
p.skowron@mimuw.edu.pl

Piotr Faliszewski
AGH University
Krakow, Poland
faliszew@agh.edu.pl

Arkadii Slinko
University of Auckland
Auckland, New Zealand
a.slinko@auckland.ac.nz

Abstract

We study the complexity of (approximate) winner determination under Monroe's and Chamberlin-Courant's multiwinner voting rules, where we focus on the total (dis)satisfaction of the voters (the utilitarian case) or the (dis)satisfaction of the worst-off voter (the egalitarian case). We show good approximation algorithms for the satisfaction-based utilitarian cases, and inapproximability results for the remaining settings.

1 Introduction

We study the complexity of (approximate) winner determination under Monroe's [1995] and Chamberlin-Courant's [1983] multiwinner voting rules, where the goal is to select a group of candidates that best represent the voters. Multiwinner elections are important both for human societies (e.g., in indirect democracies for electing committees of representatives) and for software multiagent systems (e.g., for recommendation systems [Lu and Boutilier, 2011]), and thus it is important to have good multiwinner rules and good algorithms for them. Monroe's and Chamberlin-Courant's rules are particularly appealing because they create an explicit (and, in some sense, optimal) connection between the elected committee members and the voters; each voter knows his or her representative and each committee member knows to whom he or she is accountable. Further, the optimality of the connection ensures the proportionality of the representation (again, in some sense).

When choosing a K-member committee, Monroe's and Chamberlin-Courant's rules work as follows. We have a set of m candidates and a set of n voters, where each voter ranks the candidates according to the extent to which the voter feels represented by them. For each voter we assign a single candidate to be his or her representative, respecting the following rules: (a) altogether exactly K candidates are assigned to the voters (for the case of Monroe's rule, each candidate is assigned either to about $\frac{n}{K}$ voters or to none; for Chamberlin-Courant's rule there is no such restriction[1]), (b) the candidates are selected and assigned to the voters optimally in the following sense. We assume that there is a function $\alpha\colon \mathbb{N} \to \mathbb{N}$ such that $\alpha(i)$ measures how well a voter is represented by the candidate that this voter ranks as i'th best. (We use the same function α for each voter.) We can view α either as a *satisfaction function* (then it should be a decreasing one) or as a *dissatisfaction function* (then it should be an increasing one). For example, it is typical to use the Borda count scoring function whose m-candidate dissatisfaction variant is defined as $\alpha^m_{\mathrm{DB}} = i - 1$, and whose satisfaction variant is $\alpha^m_{\mathrm{SB}} = m - i$. In the utilitarian variants of the rules, the assignment should maximize (minimize) the sum of the voters' satisfactions (dissatisfactions) with their representatives. In the egalitarian variants, the assignment should maximize (minimize) the satisfaction (the dissatisfaction) of the worst-off voter. Thus, Monroe's and Chamberlin-Courant's rules do create a useful connection between the voters and their representatives, in a way that promotes candidates' accountability to the voters, and that represents voters' views proportionally.

Indeed, among common voting rules, Monroe's and Chamberlin-Courant's rules seem to be unique in having *both* the accountability and the proportionality property. For example, first-past-the-post system (where the voters are partitioned into districts with a separate single-winner Plurality election in each) can give very disproportionate results (forcing some of the voters to be represented by candidates they despise), party-list systems (where the voters vote for parties that then distribute the seats in the committee) make the candidates feel more accountable to their parties than to their voters, and under the single transferable vote rule (STV) it is difficult to tell which candidate represents which voters. However, Monroe's rule and Chamberlin-Courant's rule have one crucial drawback that makes them impractical. It is NP-hard to tell who the winners are!

Specifically, NP-hardness of winner determination under Monroe's and Chamberlin-Courant's rules was shown by Procaccia et al. [2008] and by Lu and Boutilier [2011]. Worse yet, these hardness results hold even if various natural parameters of the election are small [Betzler *et al.*, 2011]. Rare easy cases include those when, e.g., the committee to be elected is small, or voters have single-peaked preferences [Betzler *et al.*, 2011; Yu *et al.*, 2013]. (Similar hardness results are known for some appealing single-winner rules as well, e.g., for Kemeny's rule [Bartholdi *et al.*, 1989; Hemaspaandra *et*

[1] In effect, under Chamberlin-Courant's rule each committee member might be representing a different number of voters. The committee should take this into account in its operation.

al., 2005; Betzler *et al.*, 2009].) Thus Lu and Boutilier [2011] proposed to use approximation algorithms (as was also done for Kemeny [Ailon *et al.*, 2008; Coppersmith *et al.*, 2010; Kenyon-Mathieu and Schudy, 2007]) and have given the first such algorithm for the Chamberlin-Courant's system. Their procedure outputs an assignment that achieves no less than $1 - \frac{1}{e} \approx 0.63$ fraction of the optimal voter satisfaction. Yet while there is no difference between finding an assignment with the optimal satisfaction or with the optimal dissatisfaction, as soon as approximations come into play the distinction becomes tricky. For example, under Chamberlin-Courant's system with Borda count scoring function, a $\frac{1}{2}$-approximation algorithm may assign each voter to a candidate somewhere in the middle of this voter's preference order, even if there is a feasible solution that assigns each voter to his or her most preferred one. Yet, a 2-approximation algorithm focusing on voter dissatisfaction would give the optimal result.

In this paper, we follow the line of research focusing on approximate winner determination for Monroe's and Chamberlin-Courant's rules. Our first goal is to seek an algorithm that finds assignments where voter dissatisfaction is within a fixed bound of the optimal one. Unfortunately, we have shown that under standard complexity-theoretic assumptions such algorithms do not exist. Nonetheless, we did find good algorithms that focus on voter satisfaction. Specifically, we have obtained the following results:

1. Monroe's and Chamberlin-Courant's systems are hard to approximate up to any constant factor both for the dissatisfaction-based utilitarian and egalitarian cases (Theorem 1) and for the satisfaction-based egalitarian cases (Theorems 2 and 3).
2. For the satisfaction-based utilitarian cases (with the Borda count scoring function), for the Monroe's system we give a $(0.715 - \epsilon)$-approximation algorithm (often, the ratio is much better; see Section 4), and for Chamberlin-Courant's system we give a polynomial-time approximation scheme (that is, for each ϵ, $0 < \epsilon < 1$, we have a polynomial-time $(1 - \epsilon)$-approximation algorithm; see Theorem 8).

We believe that our algorithms are practically useful. Indeed, in Section 4 we describe a real-life scenario where our algorithm for the Monroe's rule achieves approximation ratio 0.96. Further, Skowron et al. [2013] have conducted a number of experiments on both real-life and synthetic data, showing that our algorithms do well on those instances.

We model Monroe's and Chamberlin-Courant's rules as special cases of a certain resource allocation problem. The alternatives are shareable resources, each with a certain cost and a maximum number of agents that may share it; each agent has preferences over the resources and is interested in getting exactly one. Our goal is to find an optimal allocation within a given global budget (see Section 2 for details). In addition to unifying the view of the two rules, the problem is interesting in its own right and brings in a connection to other resource allocation problems. In particular, it closely resembles multi-unit resource allocation with single-unit demand [Shoham and Leyton-Brown, 2009, Chapter 11] (see also the work of Chevaleyre et al. [2006] for a survey of the most fundamental issues in the multiagent resource allocation theory) and resource allocation with sharable indivisible goods [Chevaleyre *et al.*, 2006; Airiau and Endriss, 2010]. Our paper is also very close in spirit to the recent work of Darmann et al. [2012].

2 Preliminaries

Preferences. For each $n \in \mathbb{N}$, by $[n]$ we mean $\{1, \ldots, n\}$. We assume that there is a set $N = [n]$ of *agents* and a set $A = \{a_1, \ldots a_m\}$ of *alternatives*. Each alternative a has *capacity* $\mathrm{cap}_a \in \mathbb{N}$ and cost $c_a \in \mathbb{N}$. An alternative's capacity gives the total number of the agents that can be assigned to it, and its cost gives the price of selecting the alternative (irrespective of the number of agents assigned to the alternative). Further, each agent i has a *preference order* $\succ_i$ over A, i.e., a strict linear order of the form $a_{\pi(1)} \succ_i a_{\pi(2)} \succ_i \cdots \succ_i a_{\pi(m)}$ for some permutation π of $[m]$. For an alternative a, by $\mathrm{pos}_i(a)$ we mean the position of a in i'th agent's preference order. For example, if a is the most preferred alternative for i then $\mathrm{pos}_i(a) = 1$, and if a is the least preferred one then $\mathrm{pos}_i(a) = m$. A collection $V = (\succ_1, \ldots, \succ_n)$ of agents' preference orders is called a *preference profile*.

We will often include subsets of the alternatives in the descriptions of preference orders. For example, if A is the set of alternatives and B is some nonempty strict subset of A, then by $B \succ A - B$ we mean that for the preference order $\succ$ all alternatives in B are preferred to those outside of B.

A *positional scoring function* (PSF) is a function $\alpha^m \colon [m] \to \mathbb{N}$. A PSF α^m is an *increasing positional scoring function* (IPSF) if for each $i, j \in [m]$, if $i < j$ then $\alpha^m(i) < \alpha^m(j)$. Analogously, a PSF α^m is a *decreasing positional scoring function* (DPSF) if for each $i, j \in [m]$, if $i < j$ then $\alpha^m(i) > \alpha^m(j)$.

Intuitively, if β^m is an IPSF then $\beta^m(i)$ gives the *dissatisfaction* that an agent suffers when assigned to an alternative that is ranked i'th on his or her preference order. Thus, we assume that for each IPSF β^m it holds that $\beta^m(1) = 0$ (an agent is not dissatisfied by his or her top alternative). Similarly, a DPSF γ^m measures an agent's satisfaction and we assume that for each DPSF γ^m it holds that $\gamma^m(m) = 0$.

We will often speak of families α of IPSFs (DPSFs) of the form $\{\alpha^m \mid m \in \mathbb{N}, \alpha^m \text{ is a PSF}\}$, such that:

1. If we are dealing with IPSFs, then for each $m \in \mathbb{N}$ it holds that $(\forall i \in [m])[\alpha^{m+1}(i) = \alpha^m(i)]$.
2. If we are dealing with DPSFs, then for each $m \in \mathbb{N}$ it holds that $(\forall i \in [m])[\alpha^{m+1}(i+1) = \alpha^m(i)]$.

In other words, we build our families of IPSFs (DPSFs) by appending (prepending) values to functions with smaller domains. We assume that each function α^m from a family can be computed in polynomial time with respect to m. To simplify notation, we will refer to such families of IPSFs (DPSFs) as *normal* IPSFs (normal DPSFs). We are particularly interested in the Borda count family of IPSFs (DPSFs) defined as $\alpha^m_{\mathrm{DB}}(i) = i - 1$ ($\alpha^m_{\mathrm{SB}}(i) = m - i$).

The Problem. We consider a problem of finding a function $\Phi : N \to A$ that assigns each agent to some alternative (we will call Φ an *assignment function*). We say that Φ is feasible

if for each alternative a the number of agents assigned to it does not exceed its capacity cap_a. Further, we define the cost of assignment Φ to be $\mathrm{cost}(\Phi) = \sum_{a:\Phi^{-1}(a)\neq\emptyset} c_a$ (we pay c_a for each alternative a to which at least one agent is assigned.)

Given a normal IPSF (DPSF) α, we consider two *dissatisfaction functions*, $\ell_1^\alpha(\Phi)$ and $\ell_\infty^\alpha(\Phi)$, (two *satisfaction functions*, $\ell_1^\alpha(\Phi)$ and $\min^\alpha(\Phi)$), measuring the quality of the assignment Φ as follows:

1. $\ell_1^\alpha(\Phi) = \sum_{i=1}^n \alpha(\mathrm{pos}_i(\Phi(i)))$.
2. $\ell_\infty^\alpha(\Phi) = \max_{i=1}^n \alpha(\mathrm{pos}_i(\Phi(i)))$,
   $\min^\alpha(\Phi) = \min_{i=1}^n \alpha(\mathrm{pos}_i(\Phi(i)))$.

The former one follows the utilitarian approach of measuring agents' total dissatisfaction (satisfaction), whereas the latter one follows the egalitarian approach by considering the most dissatisfied (the least satisfied) agent only.

Definition 1 *Let α be a normal IPSF. An instance of α-*DU-ASSIGNMENT *problem (i.e., of the disatisfaction-based utilitarian assignment problem) consists of a set of agents $N = [n]$, a set of alternatives $A = \{a_1, \dots a_m\}$, a preference profile V of the agents, sequences $(\mathrm{cap}_{a_1}, \dots, \mathrm{cap}_{a_m})$ and $(c_{a_1}, \dots, c_{a_m})$ of alternatives' capacities and costs, respectively, and a budget $B \in \mathbb{N}$. We ask for an assignment function Φ such that: (1) $\mathrm{cost}(\Phi) \leq B$, (2) $\forall_{a\in A} \|\{i \mid \Phi(i) = a\}\| \leq \mathrm{cap}_a$, and (3) $\ell_1^\alpha(\Phi)$ is minimized.*

Problem α-SU-ASSIGNMENT (the satisfaction-based utilitarian assignment problem) is defined identically except that α is a normal DPSF and (3) we seek to maximize $\ell_1^\alpha(\Phi)$. If we replace ℓ_1^α with ℓ_∞^α in α-DU-ASSIGNMENT then we obtain problem α-DE-ASSIGNMENT, i.e., the dissatisfaction-based egalitarian variant. If we replace ℓ_1^α with $\min^\alpha$ in α-SU-ASSIGNMENT then we obtain problem α-SE-ASSIGNMENT, i.e., the satisfaction-based egalitarian variant.

Our four problems can be viewed as generalizations of Monroe's [1995] and Chamberlin-Courant's [1983] multi-winner voting systems (see the introduction for their definitions). To model Monroe's system, it suffices to set the budget $B = K$, the cost of each alternative to be 1, and the capacity of each alternative to be $\frac{\|N\|}{K}$ (for simplicity, throughout the paper we assume that K divides $\|N\|$). We will refer to thus restricted variants of our problems as MONROE variants. To represent Chamberlin-Courant's system, we replace alternatives' capacities with $\|N\|$. We will refer to thus restricted variants of our problems as CC variants.

Computational Issues. For many normal IPSFs α (e.g., for the Borda count), even the above-mentioned restricted versions of the original problem, namely, α-DU-MONROE, α-DE-MONROE, α-DU-CC, and α-DE-CC are NP-complete [Betzler *et al.*, 2011; Procaccia *et al.*, 2008] (the same holds for the satisfaction-based variants of the problems). Thus we seek approximate solutions.

Definition 2 *Let β be a real number such that $\beta \geq 1$ ($0 < \beta \leq 1$) and let α be a normal IPSF (a normal DPSF). An algorithm is a β-approximation algorithm for α-*DU-ASSIGNMENT *problem (for α-*SU-ASSIGNMENT *problem) if on each instance I it returns a feasible assignment Φ that meets the budget constraint and such that $\ell_1^\alpha(\Phi) \leq \beta \cdot \mathrm{OPT}$ (and such that $\ell_1^\alpha(\Phi) \geq \beta \cdot \mathrm{OPT}$), where OPT is the optimal total dissatisfaction (satisfaction) $\ell_1^\alpha(\Phi_{\mathrm{OPT}})$.*

We define β-approximation algorithms for the egalitarian variants analogously. Lu and Boutilier [2011] gave a $(1-\frac{1}{e})$-approximation algorithm for SU-CC family of problems.

Throughout this paper, we will consider each of the MONROE and CC variants of the problem and for each we will either prove inapproximability with respect to any constant β (under standard complexity-theoretic assumptions) or we will present an approximation algorithm. We use the following standard NP-complete problem [Garey and Johnson, 1979].

Definition 3 *An instance I of* X3C *consists of a set $U = [n]$ (called the ground set; we assume that n is divisible by 3) and a family $\mathcal{F} = \{F_1, \dots, F_m\}$ of 3-element subsets of U. We ask if there is a set $I \subseteq [m]$ such that $\|I\| = \frac{n}{3}$ and $\bigcup_{i\in I} F_i = U$.*

Note that X3C remains NP-complete even if we additionally assume that n is divisible by 2 and each member of U appears in at most 3 sets from $\mathcal{F}$ [Garey and Johnson, 1979].

We omit some of the proofs due to space constraints (however, all our proofs exist in a written form).

3 Hardness of Approximation

We now present our inapproximability results for Monroe's and Chamberlin-Courant's rules: There are no constant-factor approximation algorithms for the dissatisfaction-based variants of the rules (both utilitarian and egalitarian) and for the satisfaction-based egalitarian ones.

Naturally, these inapproximability results carry over to more general settings. For example, unless $\mathrm{P} = \mathrm{NP}$, there are no polynomial-time constant-factor approximation algorithms for the general dissatisfaction-based resource allocation problem. On the other hand, our results do not preclude good satisfaction-based approximation algorithms for the utilitarian case; see Section 4.

Theorem 1 *For each normal IPSF α and each constant factor β, $\beta > 1$, there are no polynomial-time β-approximation algorithms for either of α-*DU-MONROE*, α-*DU-CC*, α-*DE-MONROE*, and α-*DE-CC*, unless* $\mathrm{P} = \mathrm{NP}$.

Proof Due to space constraint, we give the proof for α-DU-MONROE only. Let us fix a normal IPSF α and let us assume, for the sake of contradiction, that there is some constant β, $\beta > 1$, and a polynomial-time β-approximation algorithm $\mathcal{A}$ for α-DU-MONROE.

Let I be an instance of X3C with ground set $U = [n]$ and family $\mathcal{F} = \{F_1, \dots, F_m\}$ of 3-element subsets of U. W.l.o.g., we assume that n is divisible by 6 and that each member of U appears in at most 3 sets from $\mathcal{F}$.

Given I, we build an instance I_M of α-DU-MONROE as follows. We set $N = U$ (that is, the elements of the ground set are the agents) and we set $A = A_1 \cup A_2$, where $A_1 = \{a_1, \dots, a_m\}$ is a set of alternatives corresponding to the sets from the family $\mathcal{F}$ and A_2, $\|A_2\| = \lceil \frac{n^2\cdot\alpha(3)\cdot\beta}{2} \rceil$, is a set of dummy alternatives needed for our construction. We

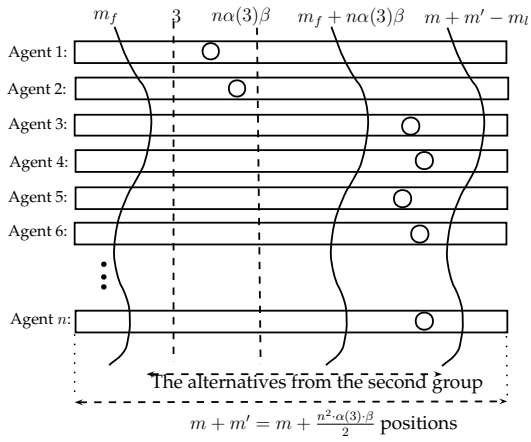

Figure 1: The alignment of the positions in the preference orders of the agents. The positions are numbered from left to right. The left wavy line shows the positions $m_f(\cdot)$, each no greater than 3. The right wavy line shows the positions $m_l(\cdot)$, each higher than $n \cdot \alpha(3) \cdot \beta$. The alternatives from A_2 (one such alternative is illustrated with a circle) are placed only between the peripheral wavy lines. Each alternative from A_2 is placed on the left from the middle wavy line exactly 2 times.

let $m' = \|A_2\|$ and we name the alternatives in A_2 so that $A_2 = \{b_1, \ldots, b_{m'}\}$. We set $K = \frac{n}{3}$.

We build agents' preference orders using the following algorithm. For each $j \in N$, set $M_f(j) = \{a_i \mid j \in F_i\}$ and $M_l(j) = \{a_i \mid j \notin F_i\}$. Set $m_f(j) = \|M_f(j)\|$ and $m_l(j) = \|M_l(j)\|$; as the frequency of the elements from U is bounded by 3, $m_f(j) \leq 3$. For each agent j we set his or her preference order to be of the form $M_f(j) \succ_j A_2 \succ_j M_l(j)$, where the alternatives in $M_f(j)$ and $M_l(j)$ are ranked in an arbitrary way and the alternatives from A_2 are placed at positions $m_f(j)+1, \ldots, m_f(j)+m'$ in the way described below (see Figure 1 for a high-level illustration of the construction).

We place the alternatives from A_2 in the preference orders of the agents in such a way that for each alternative $b_i \in A_2$ there are at most two agents that rank b_i among their $n \cdot \alpha(3) \cdot \beta$ top alternatives. The following construction achieves this effect. If $(i + j) \bmod n < 2$, then alternative b_i is placed at one of the positions $m_f(j) + 1, \ldots, m_f(j) + n \cdot \alpha(3) \cdot \beta$ in j's preference order. Otherwise, b_i is placed at a position with index higher than $m_f(j) + n \cdot \alpha(3) \cdot \beta$ (and, thus, at a position higher than $n \cdot \alpha(3) \cdot \beta$). This construction can be implemented because for each agent j there are exactly $m' \cdot \frac{2}{n} = n \cdot \alpha(3) \cdot \beta$ alternatives $b_{i_1}, b_{i_2}, \ldots, b_{i_{n\alpha(3)\beta}}$ such that $(i_k + j) \bmod n < 2$.

Let Φ be an assignment computed by $\mathcal{A}$ on I_M. We will show that $\ell_1^\alpha(\Phi) \leq n \cdot \alpha(3) \cdot \beta$ iff I is a *yes*-instance of X3C.

($\Leftarrow$) If there exists a solution for I (i.e., an exact cover of U with $\frac{n}{3}$ sets from $\mathcal{F}$), then we can easily show an assignment in which each agent j is assigned to an alternative from the top $m_f(j)$ positions of his or her preference order (namely, one that assigns each agent j to the alternative $a_i \in A_1$ that corresponds to the set F_i, from the exact cover of U, that contains j). Thus, for the optimal assignment Φ_{OPT} it holds that $\ell_1^\alpha(\Phi_{\mathrm{OPT}}) \leq \alpha(3) \cdot n$. Consequently, $\mathcal{A}$ must return an assignment with the total dissatisfaction at most $n \cdot \alpha(3) \cdot \beta$.

($\Rightarrow$) Let us now consider the opposite direction. We assume that $\mathcal{A}$ found an assignment Φ such that $\ell_1^\alpha(\Phi) \leq n \cdot \alpha(3) \cdot \beta$ and we will show that I is a *yes*-instance of X3C. Since we require each alternative to be assigned to either 0 or 3 agents, if some alternative b_i from A_2 were assigned to some 3 agents, at least one of them would rank it at a position worse than $n \cdot \alpha(3) \cdot \beta$. This would mean that $\ell_1^\alpha(\Phi) \geq n \cdot \alpha(3) \cdot \beta + 1$. Analogously, no agent j can be assigned to an alternative that is placed at one of the $m_l(j)$ bottom positions of j's preference order. Thus, only the alternatives in A_1 have agents assigned to them and, further, if agents x, y, z, are assigned to some $a_i \in A_1$, then it holds that $F_i = \{x, y, z\}$ (we will call each set F_i for which alternative a_i is assigned to some agents x, y, z *selected*). Since each agent is assigned to exactly one alternative, the selected sets are disjoint. Since the number of selected sets is $K = \frac{n}{3}$, it must be the case that the selected sets form an exact cover of U. Thus, I is a *yes*-instance of X3C. □

One may ask if the above result is not an artifact of our strict requirements regarding the budget. However, a similar proof shows that there is no (β, γ)-approximation algorithm that finds an assignment with the following properties: (1) the total dissatisfaction $\ell_1^\alpha(\Phi)$ is at most β times higher than the optimal one, (2) the number of alternatives to which agents are assigned is at most γK and (3) each selected alternative (i.e., each alternative that has agents assigned), is assigned to no more than $\gamma \lceil \frac{n}{K} \rceil$ and no less than $\frac{1}{\gamma} \lceil \frac{n}{K} \rceil$ agents.

While approximating the optimal dissatisfaction is difficult, there are good algorithms for the satisfaction-based utilitarian cases (see, e.g., the work of Lu and Boutilier [2011] and the next section). It is, thus, a bit surprising that satisfaction-based egalitarian cases are nonetheless hard to approximate.

Theorem 2 *For each normal DPSF α (with each entry polynomially bounded in the number of the alternatives) and each constant factor β, $0 < \beta \leq 1$, there is no polynomial-time β-approximation algorithm for α-*SE-MONROE *unless* $\mathrm{P} = \mathrm{NP}$.

For the case of SE-CC family of problems our inapproximability argument holds for the case of Borda DPSF only (though we believe that it can be adapted to other DPSFs as well). Further, in our previous theorems we were showing that existence of a respective constant-factor approximation algorithm implies that NP collapses to P. In the following theorem we will reduce the statement to a collapse of W[2] to FPT. (See the books of Niedermeier [2006] and Flum and Grohe [2006] for reference on parametrized complexity.)

Theorem 3 *For each constant factor β, $0 < \beta \leq 1$, no polynomial-time β-approximation algorithm for α_{SB}^m-*SE-CC *exists unless* $\mathrm{FPT} = \mathrm{W}[2]$.

4 Approximation Algorithms

We now turn to our positive results. Throughout this section we consider the Borda DPSF only; we show the first nontrivial approximation algorithms for the Monroe's system and the first polynomial-time approximation scheme (PTAS) for the Chamberlin-Courant's system.

Monroe's System. A natural iterative approach to solve α_{SB}-SU-MONROE is to pick, in each step, some not-yet-assigned alternative a_i (using some criterion) and assign it to those $\lceil \frac{N}{K} \rceil$ agents that (a) do not have any alternatives assigned to them, and (b) whose satisfaction of being represented by a_i is maximal. This idea—implemented formally in Algorithm 1—works very well in many cases. (For each positive integer k, we let $H_k = \sum_{i=1}^{k} \frac{1}{i}$ be the k'th harmonic number. Recall that $H_k = \Theta(\log k)$.)

Lemma 4 *Algorithm 1 is a polynomial-time* $(1 - \frac{K-1}{2(m-1)} - \frac{H_K}{K})$*-approximation algorithm for* α^m_{SB}-SU-MONROE.

Proof Our algorithm computes an optimal solution for $K \leq 2$. Thus we assume $K \geq 3$. Let us consider the situation in the algorithm after the i'th iteration of the outer loop (we have $i = 0$ if no iteration has been executed yet). So far, the algorithm has picked i alternatives and assigned them to $i\frac{n}{K}$ agents (recall that for simplicity we assume that K divides n evenly). Hence each agent has $\lceil \frac{m-i}{K-i} \rceil$ unassigned alternatives among his or her $i + \lceil \frac{m-i}{K-i} \rceil$ top-ranked alternatives. By pigeonhole principle, this means that there is an unassigned alternative a_ℓ that is ranked among top $i+\lceil \frac{m-i}{K-i} \rceil$ positions by at least $\frac{n}{K}$ agents. To see this, note that there are $(n-i\frac{n}{K})\lceil \frac{m-i}{K-i} \rceil$ slots for unassigned alternatives among the top $i+\lceil \frac{m-i}{K-i} \rceil$ positions in the preference orders of unassigned agents, and that there are $m - i$ unassigned alternatives. As a result, there must be an alternative a_ℓ for which the number of agents that rank a_ℓ among the top $i + \lceil \frac{m-i}{K-i} \rceil$ positions is at least: $\frac{1}{m-i}((n - i\frac{n}{K})\lceil \frac{m-i}{K-i} \rceil) \geq \frac{n}{m-i}(\frac{K-i}{K})(\frac{m-i}{K-i}) = \frac{n}{K}$. Hence, the $\lceil \frac{n}{K} \rceil$ agents assigned in the next step of the algorithm will have the total satisfaction at least $\lceil \frac{n}{K} \rceil \cdot (m - i - \lceil \frac{m-i}{K-i} \rceil)$. Thus, summing over the K iterations, the total satisfaction guaranteed by the assignment Φ computed by Algorithm 1 is at least the following value (to derive the fourth line from the third one we assume $K \geq 3$):

$$\begin{aligned}
\ell_1^{\alpha_b}(\Phi) &\geq \textstyle\sum_{i=0}^{K-1} \frac{n}{K} \cdot (m - i - \lceil \frac{m-i}{K-i} \rceil) \\
&\geq \textstyle\sum_{i=0}^{K-1} \frac{n}{K} \cdot (m - i - \frac{m-i}{K-i} - 1) \\
&= \textstyle\sum_{i=1}^{K} \frac{n}{K} \cdot (m - i - \frac{m-1}{K-i+1} + \frac{i-2}{K-i+1}) \\
&\geq \frac{n}{K}(\frac{K(2m-K-1)}{2} - (m-1)H_K) \\
&= (m-1)n(1 - \frac{K-1}{2(m-1)} - \frac{H_K}{K}).
\end{aligned}$$

If each agent were assigned to his or her top alternative, the total satisfaction would be equal to $(m-1)n$. Thus we get that $\frac{\ell_1^{\alpha_{\mathrm{SB}}}(\Phi)}{\mathrm{OPT}} \leq 1 - \frac{K-1}{2(m-1)} - \frac{H_K}{K}$. □

Note that in the above proof we measure the quality of our assignment against a perhaps-impossible, perfect solution, where each agent is assigned to his or her top alternative. This means that for relatively large m and K, and small $\frac{K}{m}$ ratio, the algorithm can achieve a close-to-ideal solution irrespective of the voters' preference orders (for example, for Polish parliamentary elections with $K = 460$ and $m \approx 6000$ the algorithm's approximation ratio would be about 0.96).

Algorithm 1: The algorithm for SU-MONROE.

Notation: $\Phi \leftarrow$ the iteratively build (partial) assignment
$\Phi^{\leftarrow} \leftarrow$ the set of agents assigned to alternatives under Φ
$\Phi^{\rightarrow} \leftarrow$ the set of alternatives used in the assignment Φ.

```
if K ≤ 2 then
    return the solution given by the algorithm of Betzler et
    al. [2011].
Φ = {}
for ℓ ← 1 to K do
    score ← {},   bests ← {}
    foreach a_i ∈ A \ Φ→ do
        agents ← sort N \ Φ← so that j ≺ k in agents
                 ⟹ pos_j(a_i) ≤ pos_k(a_i)
        bests[a_i] ← choose first ⌈N/K⌉ elements from agents
        score[a_i] ← Σ_{j∈bests}(m − pos_j(a_i))
    choose a_best such that a_best ∈ argmax_{a∈A\Φ→} score[a]
    foreach j ∈ bests[a_best] do
        Φ[j] ← a_best
```

On the flip side, to obtain a better approximation ratio we would have to use a more involved bound on the quality of the optimal solution. To see that this is the case, form an instance I of α_{SB}-SU-MONROE with n agents and m alternatives, where all the agents have the same preference order, and where the budget is K (and where K divides n). Each solution that assigns the K universally top-ranked alternatives to the agents is optimal. Thus the total satisfaction of the agents in the optimal solution is $\frac{n}{K}((m-1) + \cdots + (m-K)) = \frac{n}{K}\left(\frac{K(2m-K-1)}{2}\right) = n(m-1)\left(1 - \frac{K-1}{2(m-1)}\right)$. By taking large enough m and K (even for a fixed value of $\frac{m}{K}$), the fraction $1 - \frac{K-1}{2(m-1)}$ can be arbitrarily close to the approximation ratio of our algorithm (the argument is in spirit of identifying maximally robust elections, as studied by Shiryaev, Yu, and Elkind [2013]).

For small values of K we can solve α_{SB}-SU-MONROE optimally in polynomial time using the result of Betzler et al. [2011]. This means that we can essentially disregard the $\frac{H_K}{K}$ part of the approximation ratio of Algorithm 1 and that the quality of the solution produced by Algorithm 1 most strongly depends on the ratio $\frac{K-1}{m-1}$. We expect this ratio to be small in typical settings. For the remaining cases, e.g., when $K > \frac{m}{2}$, we can use a sampling-based randomized algorithm.

The idea is to randomly pick K alternatives and assign them optimally to the agents (in polynomial-time, as is implicitly shown in in [Betzler *et al.*, 2011]). If K is large relative to m then it is likely that such a random sample would include a large chunk of some optimal solution. In the lemma below we assess the expected satisfaction obtained with a single sampling step and the probability that a single sampling step gives satisfaction close to the expected one.

Lemma 5 *A single sampling step of the randomized algorithm for* α^m_{SB}-SU-MONROE *achieves expected approximation ratio of* $\frac{1}{2}(1 + \frac{K}{m} - \frac{K^2}{m^2-m} + \frac{K^3}{m^3-m^2})$. *For* $K \geq 8$, *the probability* p_ϵ *that the relative deviation between the obtained total satisfaction and the expected total satisfaction is higher than* ϵ, *is at most* $\exp(-\frac{K\epsilon^2}{128})$.

Algorithm 2: The algorithm for SU-CC.

Notation: As in Algorithm 1. $\mathrm{W}(\cdot)$ is Lambert's W-Function.
$\mathrm{num_pos}_x(a) \leftarrow \|\{i \in [n] \setminus \Phi^{\leftarrow} : pos_i(a) \leq x\}\|$
(the number of not-yet assigned agents that rank alternative a in one of their first x positions)

$\Phi = \{\}, x = \lceil \frac{m\mathrm{W}(K)}{K} \rceil$
for $i \leftarrow 1$ **to** K **do**
  choose a_i such that $a_i \in \mathrm{argmax}_{a \in A \setminus \Phi^{\rightarrow}} \mathrm{num_pos}_x(a)$
  foreach $j \in [n] \setminus \Phi^{\leftarrow}$ **do**
    if $pos_j(a_i) \leq x$ **then**
      $\Phi[j] \leftarrow a_i$
foreach $j \in A \setminus \Phi^{\leftarrow}$ **do**
  $a \leftarrow \mathrm{argmin}_{a' \in \Phi^{\rightarrow}} pos_j(a')$
  $\Phi[j] \leftarrow a$

The threshold for $\frac{K}{m}$, where the randomized algorithm is (in expectation) better than the greedy algorithm is about 0.57. By combining the two algorithms, we can guarantee an expected approximation ratio of $0.715 - \epsilon$, for each fixed constant ϵ. Formally, for a given fixed ϵ our algorithm, to which we will refer as the combined algorithm, works as follows:

1. Let I be the input instance and let λ be the required probability of success. If $m \leq 1 + \frac{2}{\epsilon}$ then solve I using brute-force; if $K \leq 8$ or $\frac{H_K}{K} \geq \frac{\epsilon}{2}$ then solve I using the algorithm of Betzler et al. [2011] for fixed K.
2. Compute solution Φ_1 using Algorithm 1. Run the sampling step for $\frac{-512 \log(1-\lambda)}{K\epsilon^2}$ times and let Φ_2 be the best solution found. Return the best solution in $\{\Phi_1, \Phi_2\}$.

Theorem 6 *For each fixed ϵ, the combined algorithm gives a $(0.715 - \epsilon)$-approximate solution for the problem α_{SB}-SU-MONROE with probability λ, in time polynomial with respect to the input size and $-\log(1-\lambda)$.*

Chamberlin-Courant's System. Our algorithms for Monroe's system apply to the Chamberlin-Courant's setting as well, improving upon the result of Lu and Boutilier [2011]. However, we provide an even better algorithm and derive a polynomial-time approximation scheme.

The idea of our method (Algorithm 2) is to compute a certain threshold value x and to greedily find an assignment that (approximately) maximizes the number of agents assigned to alternatives that they rank at position x or better. For nonnegative real numbers, Lambert's W function, $\mathrm{W}(x)$, is defined to be the solution of $x = \mathrm{W}(x)e^{\mathrm{W}(x)}$ ($\mathrm{W}(x) = O(\log x)$).

Lemma 7 *Algorithm 2 is a polynomial-time $(1 - \frac{2\mathrm{W}(K)}{K})$-approximation algorithm for α^m_{SB}-SU-CC.*

Proof Let $x = \frac{m\mathrm{W}(K)}{K}$. We will first give an inductive proof that, for each i, $0 \leq i \leq K$, after the i'th iteration of the outer loop at most $n(1 - \frac{\mathrm{W}(K)}{K})^i$ agents are unassigned. Then we will derive the approximation ratio of our algorithm.

For $i = 0$, the inductive hypothesis holds because $n(1 - \frac{\mathrm{W}(K)}{K})^0 = n$. For each i, let n_i denote the number of unassigned agents after the i'th iteration. Thus after the i'th iteration there are n_i unassigned agents, each with x unassigned alternatives among his or her x most-preferred alternatives. As a result, at least one unassigned alternative is present in at least $\frac{n_i x}{m-i}$ of the highest x positions of the unassigned agents. This means that after the $(i+1)$'th iteration the number of unassigned agents is $n_{i+1} \leq n_i - \frac{n_i x}{m-i} \leq n_i(1 - \frac{x}{m}) = n_i(1 - \frac{\mathrm{W}(K)}{K})$. If for a given i the inductive hypothesis holds, that is, if $n_i \leq n(1 - \frac{\mathrm{W}(K)}{K})^i$, then $n_{i+1} \leq n(1 - \frac{\mathrm{W}(K)}{K})^i(1 - \frac{\mathrm{W}(K)}{K}) = n(1 - \frac{\mathrm{W}(K)}{K})^{i+1}$. Thus the hypothesis holds and, as a result, we have that $n_k \leq n(1 - \frac{\mathrm{W}(K)}{K})^K \leq n(\frac{1}{e})^{\mathrm{W}(K)} = \frac{n\mathrm{W}(K)}{K}$.

Let Φ be the assignment computed by our algorithm. To compare it against the optimal solution, it suffices to observe that the optimal solution has total satisfaction at most $\mathrm{OPT} \leq (m-1)n$, that each agent selected during the first K steps has satisfaction at least $m - x = m - \frac{m\mathrm{W}(K)}{K}$, and that the agents not assigned within the first K steps have satisfaction no worse than 0. Thus: $\frac{\ell_1^{\alpha_{\mathrm{SB}}}(\Phi)}{\mathrm{OPT}} \geq \frac{(n - \frac{n\mathrm{W}(K)}{K})(m - \frac{m\mathrm{W}(K)}{K})}{(m-1)n} \geq (1 - \frac{\mathrm{W}(K)}{K})(1 - \frac{\mathrm{W}(K)}{K}) \geq 1 - \frac{2\mathrm{W}(K)}{K}$. □

Since for each $\epsilon > 0$ there is a value K_ϵ such that for each $K > K_\epsilon$ it holds that $\frac{2\mathrm{W}(K)}{K} < \epsilon$, and α_{SB}-SU-CC problem can be solved optimally in polynomial time for each fixed constant K [Betzler *et al.*, 2011], there is a polynomial-time approximation scheme (PTAS) for α_{SB}-SU-CC (i.e., a family of algorithms such that for each fixed β, $0 < \beta < 1$, there is a polynomial-time β-approximation algorithm for α_{SB}-SU-CC in the family; note that in PTASes one assumes β to be a fixed constant).

Theorem 8 *There is a PTAS for α_{SB}-SU-CC.*

Independently, Oren [2012] gave a randomized sampling-based $(1 - \frac{1}{K+1})(1 + \frac{1}{m})$-approximation algorithm.

5 Conclusions

We have defined a certain resource allocation problem that generalizes Monroe's and Chamberlin-Courant's multiwinner voting rules. Since these rules are hard to compute [Procaccia *et al.*, 2008; Lu and Boutilier, 2011; Betzler *et al.*, 2011], we have investigated the possibility of computing approximate solutions (see, e.g., [Caragiannis *et al.*, 2012; Ailon *et al.*, 2008; Coppersmith *et al.*, 2010; Kenyon-Mathieu and Schudy, 2007] for such results on single-winner rules). For the case of maximizing total voters' satisfaction, we have obtained the first nontrivial approximation algorithms for Monroe's rule (our randomized algorithm obtains approximation ratios arbitrarily close to 0.715) and the first PTAS for Chamberlin-Courant's rule (the remaining cases turned out to be hard to approximate); Skowron et al. [2013] verified empirically that our algorithms do well in practice. Our algorithms are particularly appealing when voters submit truncated ballots because they focus on top parts of voters' preference orders.

Acknowledgments. We would like to thank the referees for their helpful comments. Supported in part by Poland's National Science Center grants UMO-2012/06/M/ST1/00358, DEC-2011/03/B/ST6/01393 and by AGH University grant 11.11.230.015.

References

[Ailon *et al.*, 2008] N. Ailon, M. Charikar, and A. Newman. Aggregating inconsistent information: Ranking and clustering. *Journal of the ACM*, 55(5):Article 23, 2008.

[Airiau and Endriss, 2010] S. Airiau and U. Endriss. Multiagent resource allocation with sharable items: Simple protocols and nash equilibria. In *Proceedings of AAMAS-2010*, pages 167–174, May 2010.

[Bartholdi *et al.*, 1989] J. Bartholdi, III, C. Tovey, and M. 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.*, 2009] N. Betzler, M. Fellows, J. Guo, R. Niedermeier, and F. Rosamond. Fixed-parameter algorithms for Kemeny scores. *Theoretical Computer Science*, 410(45):45544570, 2009.

[Betzler *et al.*, 2011] N. Betzler, A. Slinko, and J. Uhlmann. On the computation of fully proportional representation. Technical Report SSRN: http://ssrn.com/abstract=1952497, U. of Auckland, November 2011.

[Caragiannis *et al.*, 2012] I. Caragiannis, J. Covey, M. Feldman, C. Homan, C. Kaklamanis, N. Karanikolas, A. Procaccia, and J. Rosenschein. On the approximability of Dodgson and Young elections. *Artificial Intelligence*, 187:31–51, 2012.

[Chamberlin and Courant, 1983] B. Chamberlin and P. Courant. Representative deliberations and representative decisions: Proportional representation and the borda rule. *American Political Science Review*, 77(3):718–733, 1983.

[Chevaleyre *et al.*, 2006] Y. Chevaleyre, P. Dunne, U. Endriss, J. Lang, M. Lematre, N. Maudet, J. Padget, S. Phelps, J. Rodríguez-Aguilar, and P. Sousa. Issues in multiagent resource allocation. *Informatica*, 30:3–31, 2006.

[Coppersmith *et al.*, 2010] D. Coppersmith, L. Fleisher, and A. Rurda. Ordering by weighted number of wins gives a good ranking for weighted tournaments. *ACM Transactions on Algorithms*, 6(3):Article 55, 2010.

[Darmann *et al.*, 2012] A. Darmann, E. Elkind, S. Kurz, J. Lang, J. Schauer, and G. Woeginger. Group activity selection problem. In *Proceedings of WINE-12*, pages 156–169, December 2012.

[Flum and Grohe, 2006] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer-Verlag, 2006.

[Garey and Johnson, 1979] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[Hemaspaandra *et al.*, 2005] E. Hemaspaandra, H. Spakowski, and J. Vogel. The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391, 2005.

[Kenyon-Mathieu and Schudy, 2007] C. Kenyon-Mathieu and W. Schudy. How to rank with few errors. In *Proceedings of STOC-2007*, pages 95–103, June 2007.

[Lu and Boutilier, 2011] T. Lu and C. Boutilier. Budgeted social choice: From consensus to personalized decision making. In *Proceedings of IJCAI-2011*, pages 280–286, 2011. See also the version that appeared in COMSOC-10.

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

[Niedermeier, 2006] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, 2006.

[Oren, 2012] J. Oren. Personal communication, 2012.

[Procaccia *et al.*, 2008] A. Procaccia, J. Rosenschein, and A. Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30(3):353–362, April 2008.

[Shiryaev *et al.*, 2013] D. Shiryaev, L. Yu, and E. Elkind. On elections with robust winners. In *Proceedings of AAMAS-2013*, May 2013.

[Shoham and Leyton-Brown, 2009] Y. Shoham and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

[Skowron *et al.*, 2013] P. Skowron, P. Faliszewski, and A. Slinko. Achieving fully proportional representation is easy in practice. In *Proceedings of AAMAS-2013*, May 2013.

[Yu *et al.*, 2013] L. Yu, H. Chan, and E. Elkind. Multiwinner elections under preferences that are single-peaked on a tree. In *Proceedings of IJCAI-2013*, August 2013.