Socially Desirable Approximations for Dodgson's Voting Rule*

Ioannis Caragiannis
University of Patras & RACTI
Building B, 26500 Rio, Greece
caragian@ceid.upatras.gr

Christos Kaklamanis
University of Patras & RACTI
Building B, 26500 Rio, Greece
kakl@ceid.upatras.gr

Nikos Karanikolas
University of Patras & RACTI
Building B, 26500 Rio, Greece
nkaranik@ceid.upatras.gr

Ariel D. Procaccia
Harvard SEAS
Cambridge, MA 02138, USA
arielpro@seas.harvard.edu

ABSTRACT

In 1876 Charles Lutwidge Dodgson suggested the intriguing voting rule that today bears his name. Although Dodgson's rule is one of the most well-studied voting rules, it suffers from serious deficiencies, both from the computational point of view—it is $\mathcal{NP}$-hard even to approximate the Dodgson score within sublogarithmic factors—and from the social choice point of view—it fails basic social choice desiderata such as monotonicity and homogeneity.

In a previous paper [Caragiannis et al., SODA 2009] we have asked whether there are approximation algorithms for Dodgson's rule that are monotonic or homogeneous. In this paper we give definitive answers to these questions. We design a monotonic exponential-time algorithm that yields a 2-approximation to the Dodgson score, while matching this result with a tight lower bound. We also present a monotonic polynomial-time $\mathcal{O}(\log m)$-approximation algorithm (where m is the number of alternatives); this result is tight as well due to a complexity-theoretic lower bound. Furthermore, we show that a slight variation of a known voting rule yields a monotonic, homogeneous, polynomial-time $\mathcal{O}(m \log m)$-approximation algorithm, and establish that it is impossible to achieve a better approximation ratio even if one just asks for homogeneity. We complete the picture by studying several additional social choice properties; for these properties, we prove that algorithms with an approximation ratio that depends only on m do not exist.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

General Terms

Algorithms, Theory, Economics

Keywords

Social choice, Dodgson's voting rule, Approximation algorithms

*This work is partially supported by the European Union under IST FET Integrated Project FP6-015964 AEOLUS and Cost Action IC0602 "Algorithmic Decision Theory", and by a "Caratheodory" basic research grant from the University of Patras.

1. INTRODUCTION

Social choice theory is concerned with aggregating the preferences of a set of n *agents* over a set of m *alternatives*. It is often assumed that each agent holds a private ranking of the alternatives; the collection of agents' rankings is known as a *preference profile*. The preference profile is reported to a *voting rule*, which then singles out the winning alternative.

When there are two alternatives (and an odd number of agents), majority voting is unanimously considered a perfect method of selecting the winner. However, when there are at least three alternatives it is sometimes unclear which alternative is best. In the Eighteenth Century the marquis de Condorcet, perhaps the founding father of the mathematical theory of voting, suggested a solution by extending majority voting to multiple alternatives [10]. An alternative x is said to beat alternative y in a *pairwise election* if a majority of agents prefer x to y, i.e., rank x above y. An alternative that beats every other alternative in a pairwise election is easy to accept as the winner of the entire election; in the modern literature such an alternative is known as a *Condorcet winner*. Unfortunately, there are preference profiles for which no alternative is a Condorcet winner.

Almost a century after Condorcet, a refinement of Condorcet's ideas was proposed by Charles Lutwidge Dodgson (today better known by his pen name Lewis Carroll), despite apparently being unfamiliar with Condorcet's work [5]. Dodgson proposed selecting the alternative "closest" to being a Condorcet winner, in the following sense. The *Dodgson score* of an alternative x is the number of exchanges between adjacent alternatives in the agents' rankings that must be introduced in order for x to become a Condorcet winner (see Section 2 for an example). A *Dodgson winner* is an alternative with minimum Dodgson score.

Although Dodgson's rule is intuitively appealing, it has been heavily criticized over the years for failing to satisfy desirable properties that are considered by social choice theorists to be extremely basic. Most prominent among these properties are *monotonicity* and *homogeneity*; a voting rule is said to be *monotonic* if it is indifferent to pushing a winning alternative upwards in the preferences of the agents, and is said to be *homogeneous* if it is invariant under duplication of the electorate. In fact, several authors have commented that it is somewhat unfair to attribute the above-mentioned rule to Dodgson, since Dodgson himself seems to have questioned it due to its serious defects (see, e.g., the papers by Tideman [21, p. 194] and Fishburn [11, p. 474]).

To make matters worse, the rise of computational complexity theory, a century after the conception of Dodgson's rule, has made it clear that it suffers from yet another serious deficiency: it is intractable to single out the winner of the election. Indeed, it is the first voting rule where winner determination was known to be $\mathcal{NP}$-hard [4]; even the computation of the Dodgson score of a given alternative is $\mathcal{NP}$-hard. The question of the exact complexity of winner determination under Dodgson's rule was resolved by Hemaspaandra et al. [13]: it is complete for the class Θ_2^p. These results have sparked great interest in Dodgson's rule among computer scientists, making it "one of the most studied voting rules in computational social choice" [6].

In previous work with numerous colleagues [8], we have largely taken the computational complexity point of view by considering the computation of the Dodgson score as an optimization problem. Among other results, we have given two polynomial-time algorithms that guarantee an approximation ratio of $\mathcal{O}(\log m)$ to the Dodgson score (where m is the number of alternatives); this bound is asymptotically tight with respect to polynomial-time algorithms (unless $\mathcal{P} = \mathcal{NP}$). Approximating Dodgson's rule (using slightly different notions of approximation) has also been considered by Homan and Hemaspaandra [14], McCabe-Dansted et al. [16], and Tideman [22, pages 199-201].

Taking the social choice point of view, our main conceptual contribution in [8] was the suggestion that an algorithm that approximates the Dodgson score is a voting rule in its own right in the sense that it naturally induces a voting rule that selects an alternative with minimum score according to the algorithm. Hence, such algorithms should be evaluated not only by their computational properties (e.g., approximation ratio and complexity) but also by their social choice properties (e.g., monotonicity and homogeneity). In other words, they should be "socially desirable". This issue was very briefly explored in the foregoing paper: we have shown that one of our two approximation algorithms satisfies a weak flavor of monotonicity, whereas the other does not. Both algorithms, as well as Dodgson's rule itself, are neither monotonic (in the usual sense) nor homogeneous, but this does not preclude the existence of monotonic or homogeneous approximation algorithms for the Dodgson score. Indeed, we have asked whether there exist such algorithms that yield a good approximation ratio [8, p. 1064].

In the following, we refer to algorithms approximating the Dodgson score (as well as to the voting rules they induce) using the term *Dodgson approximations*. A nice property that Dodgson approximations enjoy is that a finite approximation ratio implies *Condorcet-consistency*, i.e., a Condorcet winner (if one exists) is elected as the unique winner. One might wish for approximations of the Dodgson ranking (i.e., the ranking of the alternatives with respect to their Dodgson scores) directly instead of approximating the Dodgson score. Unfortunately, it is known that distinguishing whether an alternative is the Dodgson winner or in the last $\mathcal{O}(\sqrt{m})$ positions in the Dodgson ranking is $\mathcal{NP}$-hard [8]. This extreme inapproximability result provides a complexity-theoretic explanation of the discrepancies that have been observed in the social choice literature when comparing Dodgson's rule to simpler polynomial-time voting rules (see the discussion in [8]) and implies that, as long as we care about efficient algorithms, reasonable approximations of the Dodgson ranking are impossible. However, the cases where the ranking is hard to approximate are cases where the alternatives have very similar Dodgson scores. We would argue that in those cases it is not crucial, from Dodgson's point of view, which alternative is elected, since they are all almost equally close to being Condorcet winners. Put another way, if the Dodgson score is a measure of an alternative's quality, the goal is simply to elect a good alternative according to this measure.

Our results and techniques. In this paper we give definitive (and mostly positive) answers to the questions raised above; our results are tight.

In Section 3 we study monotonic Dodgson approximations. We first design an algorithm that we denote by M. Roughly speaking, this algorithm "monotonizes" Dodgson's rule by explicitly defining a winner set for each given preference profile, and assigning an alternative to the winner set if it is a Condorcet winner in some preference profile such that the former profile is obtained from the latter by pushing the alternative upwards. We prove the following result.

THEOREM 3.2. *M is a monotonic Dodgson approximation with an approximation ratio of* 2.

We furthermore show that there is no monotonic Dodgson approximation with a ratio smaller than 2 (Theorem 3.4), hence M is optimal among monotonic Dodgson approximations. Note that the lower bound is independent of computational assumptions, and, crucially, computing an alternative's score under M requires exponential time. This is to be expected since the Dodgson score is computationally hard to approximate within a factor better than $\Omega(\log m)$ [8].

It is now natural to ask whether there is a monotonic *polynomial-time* Dodgson approximation with an approximation ratio of $\mathcal{O}(\log m)$. We give a positive answer to this question as well. Indeed, we design a Dodgson approximation denoted by Q, and establish the following result.

THEOREM 3.10. *Q is a monotonic polynomial-time Dodgson approximation with an approximation ratio of* $\mathcal{O}(\log m)$.

The result relies on monotonizing an existing Dodgson approximation that is based on linear programming. The main obstacle is to perform the monotonization in polynomial time rather than looking at an exponential number of profiles, as described above. Our main tool is the notion of *pessimistic estimator*, which allows the algorithm to restrict its attention to a single preference profile. Pessimistic estimators are obtained by solving a linear program that is a variation of the one that approximates the Dodgson score.

In Section 4 we turn to homogeneity. We consider Tideman's *simplified Dodgson rule* [22, pages 199-201], which was designed to overcome the deficiencies of Dodgson's rule. The former rule is computable in polynomial time, and is moreover known to be monotonic and homogeneous. By scaling the score given by the simplified Dodgson rule we obtain a Dodgson approximation, denoted Td', that is identical as a voting rule, and moreover has the following properties.

THEOREM 4.1. *Td' is a monotonic, homogeneous, polynomial-time Dodgson approximation with an approximation ratio of* $\mathcal{O}(m \log m)$.

Note that the Dodgson score can be between 0 and $\Theta(nm)$, so this bound is far from trivial. The analysis is tight when there is an alternative that is tied against many other alternatives in pairwise elections (and hence has relatively high Dodgson score), whereas another alternative strictly loses in pairwise elections to few alternatives (so it has relatively low Dodgson score). By homogeneity the former alternative must be elected, since its score does not scale when the electorate is replicated (we elaborate in Section 4). This intuition leads to the following result which applies to any (even exponential-time) homogeneous Dodgson approximation.

THEOREM 4.2. *Any homogeneous Dodgson approximation has approximation ratio at least* $\Omega(m \log m)$.

In particular the upper bound given in Theorem 4.1 (which is achieved by an algorithm that is moreover monotonic and efficient) is asymptotically tight. The heart of our construction is the design of a preference profile such that an alternative is tied against $\Omega(m)$ other alternatives; this is equivalent to a construction of a family of subsets of a set U, $|U| = m$, such that each element of U appears in roughly half the subsets but the minimum cover is of size $\Omega(\log m)$.

In order to complete the picture, in Section 5 we discuss some other, less prominent, social choice properties not satisfied by Dodgson's rule [22, Chapter 13]: combinativity, Smith consistency, mutual majority, invariant loss consistency, and independence of clones. We show that any Dodgson approximation that satisfies one of these properties has an approximation ratio of $\Omega(nm)$ (in the case of the former two properties) or $\Omega(n)$ (in the case of the latter three). An $\Omega(nm)$ ratio is a completely trivial one, but we also consider an approximation ratio of $\Omega(n)$ to be impractical, as the number of agents n is very large in almost all settings of interest.

Discussion. Our results with respect to monotonicity are positive across the board. In particular, we find Theorem 3.2 surprising as it indicates that Dodgson's lack of monotonicity can be circumvented by slightly modifying the definition of the Dodgson score; in a sense this suggests that Dodgson's rule is not fundamentally far from being monotonic. Theorem 3.10 provides a striking improvement over the main result of [8]. Indeed, if one is interested in computationally tractable algorithms then an approximation ratio of $\mathcal{O}(\log m)$ is optimal; the theorem implies that we can additionally satisfy monotonicity without (asymptotically) increasing the approximation ratio. Our monotonization techniques may be of independent interest.

Our results regarding homogeneity, Theorem 4.1 and Theorem 4.2, can be interpreted both positively and negatively. Consider first the case where the number of alternatives m is small (e.g., in political elections). A major advantage of Theorem 4.1 is that it concerns Tideman's simplified Dodgson rule, which is already recognized as a desirable voting rule, as it is homogeneous, monotonic, Condorcet-consistent, and resolvable in polynomial time. The theorem lends further justification to this rule by establishing that it always elects an alternative relatively close (according to Dodgson's notion of distance) to being a Condorcet winner, that is, the spirit of Dodgson's ideas is indeed preserved by the "simplification" and (due to Theorem 4.2) this is accomplished in the best possible way.

Viewed negatively, when the number of alternatives is large (an extreme case is a multiagent system where the agents are voting over joint plans), Theorem 4.2 strengthens the criticism against Dodgson's rule: not only is the rule itself nonhomogeneous, but any (even exponential-time computable) conceivable variation that tries to roughly preserve Dodgson's notion of proximity to Condorcet is also nonhomogeneous. We believe that both interpretations of the homogeneity results are of interest to social choice theorists as well as computer scientists.

As an aside, note that almost all work in *algorithmic mechanism design* [18] seeks *truthful* approximation algorithms, that is, algorithms such that the agents cannot benefit by lying. However, it is well known that in the standard social choice setting, truthfulness cannot be achieved. Indeed, the Gibbard-Satterthwaite Theorem [12, 19] (see also [17]) implies that any minimally reasonable voting rule is not truthful. Therefore, social choice theorists strive for other socially desirable properties, and in particular the ones discussed above. To avoid confusion, we remark that although notions of monotonicity are often studied in mechanism design as ways of obtaining truthfulness (e.g., see [3]), in social choice theory monotonicity is a very basic desirable property in its own right (and has been so long before mechanism design was conceived).

Future work. In the future, we envision the extension of our agenda of socially desirable approximation algorithms to other important voting rules. Positive results in this direction would provide us with tools to circumvent the deficiencies of known voting rules without sacrificing their core principles; negative results would further enhance our understanding of such deficiencies. Note that these questions are relevant even with respect to tractable voting rules that do not satisfy certain properties, but seem especially interesting in the context of voting and rank aggregation rules that are hard to compute, e.g., Kemeny's and Slater's rules [1, 9, 15]. The work in this direction might involve well-known tractable, Condorcet-consistent, monotonic, and homogeneous rules such as Copeland and Maximin (see, e.g., [22]) in the same way that we use Tideman's simplified Dodgson rule in the current paper. It might also use different notions of approximation (such as additive or differential approximations) besides the standard definition of the approximation ratio as a multiplicative factor used in the current paper.

2. PRELIMINARIES

We consider a set of *agents* $N = \{0, 1, \ldots, n-1\}$ and a set of *alternatives* A, $|A| = m$. Each agent has linear preferences over the alternatives, that is, a ranking over the alternatives. Formally, the preferences of agent i are a binary relation $\succ_i$ over A that satisfies irreflexivity, asymmetry, transitivity and totality; given $x, y \in A$, $x \succ_i y$ means that i prefers x to y. We let $\mathcal{L} = \mathcal{L}(A)$ be the set of linear preferences over A. A *preference profile* $\succ = \langle \succ_0, \ldots, \succ_{n-1} \rangle \in \mathcal{L}^n$ is a collection of preferences for all the agents. A *voting rule* (also known as a *social choice correspondence*) is a function $f : \mathcal{L}^n \to 2^A \setminus \{\emptyset\}$ from preference profiles to nonempty subsets of alternatives, which designates the winner(s) of the election.

Let $x, y \in A$, and $\succ \in \mathcal{L}^n$. We say that x beats y in a *pairwise election* if $|\{i \in N : x \succ_i y\}| > n/2$, that is, if a (strict) majority of agents prefer x to y. A *Condorcet winner* is an alternative that beats every other alternative in a pairwise election. The *Dodgson score* of an alternative $x \in A$ with respect to a preference profile $\succ \in \mathcal{L}^n$, denoted $\mathrm{sc}_D(x, \succ)$, is the number of swaps between adjacent alternatives in the individual rankings that are required in order to make it a Condorcet winner. A *Dodgson winner* is an alternative with minimum Dodgson score.

Consider, for example, the profile $\succ$ in Table 1; in this example $N = \{0, \dots, 4\}$, $A = \{a, b, c, d, e\}$, and the ith column is the ranking reported by agent i. Alternative a loses in pairwise elections to b and e (two agents prefer a to b, one agent prefers a to e). In order to become a Condorcet winner, four swaps suffice: swapping a and e, and then a and b, in the ranking of agent 1 (after the swaps the ranking becomes $a \succ_1 b \succ_1 e \succ_1 c \succ_1 d$), and swapping a and d, and then a and e, in the ranking of agent 4. Agent a cannot be made a Condorcet winner with fewer swaps, hence we have $\mathrm{sc}_D(a, \succ) = 4$ in this profile. However, in the profile of Table 1 there is a Condorcet winner, namely agent b, hence b is the Dodgson winner with $\mathrm{sc}_D(b, \succ) = 0$.

0	1	2	3	4
a	b	e	e	b
b	e	b	c	e
c	a	c	d	d
d	c	a	a	a
e	d	d	b	c

Table 1: An example of the Dodgson score. For this profile $\succ$, it holds that $\mathrm{sc}_D(b, \succ) = 0$, $\mathrm{sc}_D(a, \succ) = 4$.

Given a preference profile $\succ \in \mathcal{L}^n$ and $x, y \in A$, the *deficit* of x against y, denoted $\mathrm{defc}(x, y, \succ)$, is the number of additional agents that must rank x above y in order for x to beat y in a pairwise election. Formally,

$$\mathrm{defc}(x, y, \succ) \quad = \quad \max\left\{0, \left\lceil \frac{n+1}{2} \right\rceil - |\{i \in N : x \succ_i y\}|\right\}.$$

In particular, if x beats y in a pairwise election then it holds that $\mathrm{defc}(x, y, \succ) = 0$. Note that if n is even and x and y are tied, that is, $|\{i \in N : x \succ_i y\}| = n/2$, then $\mathrm{defc}(x, y, \succ) = 1$. For example, in the profile of Table 1 we have that $\mathrm{defc}(a, b, \succ) = 1$, $\mathrm{defc}(a, c, \succ) = 0$, $\mathrm{defc}(a, d, \succ) = 0$, $\mathrm{defc}(a, e, \succ) = 2$.

We consider algorithms that receive as input an alternative $x \in A$ and a preference profile $\succ \in \mathcal{L}^n$, and return a score for x. We denote the score returned by an algorithm V on the input which consists of an alternative $x \in A$ and a profile $\succ \in \mathcal{L}^n$ by $\mathrm{sc}_V(x, \succ)$. We call such an algorithm V a *Dodgson approximation* if $\mathrm{sc}_V(x, \succ) \geq \mathrm{sc}_D(x, \succ)$ for every alternative $x \in A$ and every profile $\succ \in \mathcal{L}^n$. We also say that V has an *approximation ratio* of ρ if $\mathrm{sc}_D(x, \succ) \leq \mathrm{sc}_V(x, \succ) \leq \rho \cdot \mathrm{sc}_D(x, \succ)$, for every $x \in A$ and every $\succ \in \mathcal{L}^n$. A Dodgson approximation naturally induces a voting rule by electing the alternative(s) with minimum score. Hence, when we say that a Dodgson approximation satisfies a social choice property we are referring to the voting rule induced by the algorithm. Observe that the voting rule induced by a Dodgson approximation with finite approximation ratio is *Condorcet-consistent*, i.e., it always elects a Condorcet winner as the sole winner if one exists.

Let us give an example. Consider the algorithm V that, given an alternative $x \in A$ and a preference profile $\succ \in \mathcal{L}^n$, returns a score of $\mathrm{sc}_V(x, \succ) = m \cdot \sum_{y \in A \setminus \{x\}} \mathrm{defc}(x, y, \succ)$. It is easy to show that this algorithm is a Dodgson approximation and, furthermore, has approximation ratio at most m. Indeed, it is possible to make x beat y in a pairwise election by pushing x to the top of the preferences of $\mathrm{defc}(x, y, \succ)$ agents, and this requires at most $m \cdot \mathrm{defc}(x, y, \succ)$ swaps. By summing over all $y \in A \setminus \{x\}$, we obtain an upper bound of $\mathrm{sc}_V(x, \succ)$ on the Dodgson score of x. On the other hand, given $x \in A$, for every $y \in A \setminus \{x\}$ we require $\mathrm{defc}(x, y, \succ)$ swaps that push x above y in the preferences of some agent in order for x to beat y in a pairwise election. Moreover, these swaps do not decrease the deficit against any other alternative. Therefore, $\sum_{y \in A \setminus \{x\}} \mathrm{defc}(x, y, \succ) \leq \mathrm{sc}_D(x, \succ)$, and by multiplying by m we get that $\mathrm{sc}_V(x, \succ) \leq m \cdot \mathrm{sc}_D(x, \succ)$.

3. MONOTONICITY

In this section we present our results on monotonic Dodgson approximations. A voting rule is *monotonic* if a winning alternative remains winning after it is pushed upwards in the preferences of some of the agents. Dodgson's rule is known to be non-monotonic (see, e.g., [6]). The intuition is that if an agent ranks x directly above y and y above z, swapping x and y may not help y if it already beats x, but may help z defeat x.

As a warm-up we observe that the Dodgson approximation mentioned at the end of the previous section is monotonic as a voting rule. Indeed, consider a preference profile $\succ$ and a winning alternative x. Pushing x upwards in the preference of some of the agents can neither increase its score (since its deficit against any other alternative does not increase) nor decrease the score of any other alternative $y \in A \setminus \{x\}$ (since the deficit of y against any alternative in $A \setminus \{x, y\}$ remains unchanged and its deficit against x does not decrease).

3.1 Monotonizing Dodgson's Voting Rule

In the following we present a much stronger result. Using a natural monotonization of Dodgson's voting rule, we obtain a monotonic Dodgson approximation with approximation ratio at most 2. The main idea is to define the winning set of alternatives for a given profile first and then assign the same score to the alternatives in the winning set and a higher score to the non-winning alternatives. Roughly speaking, the winning set is defined so that it contains the Dodgson winners for the given profile as well as the Dodgson winners of other profiles that are necessary so that monotonicity is satisfied.

More formally, we say that a preference profile $\succ' \in \mathcal{L}^n$ is a *y-improvement* of $\succ$ for some alternative $y \in A$ if $\succ'$ is obtained by starting from $\succ$ and pushing y upwards in the preferences of some of the agents. In particular a profile is a y-improvement of itself for any alternative $y \in A$. The next statement is obvious.

Observation 3.1. *Let $y \in A$ and let $\succ, \succ' \in \mathcal{L}^n$ be profiles such that $\succ'$ is a y-improvement of $\succ$. Then,*

$$sc_D(y, \succ') \leq sc_D(y, \succ).$$

We monotonize Dodgson's voting rule as follows. Let M denote the new voting rule we are constructing. We denote by $W(\succ)$ the set of winners of M for profile $\succ \in \mathcal{L}^n$. Let $\Delta = \max_{y \in W(\succ)} \mathrm{sc}_D(y, \succ)$. The voting rule M assigns a score of $\mathrm{sc}_M(y, \succ) = \Delta$ to each alternative $y \in W(\succ)$ and a score of

$$\mathrm{sc}_M(y, \succ) = \max\{\Delta + 1, \mathrm{sc}_D(y, \succ)\}$$

to each alternative $y \notin W(\succ)$. All that remains is to define the set of winners $W(\succ)$ for profile $\succ$. This is done as follows: for each preference profile $\succ^* \in \mathcal{L}^n$ and each Dodgson winner y^* at $\succ^*$, include y^* in the winner set $W(\succ')$ of each preference profile $\succ' \in \mathcal{L}^n$ that is a y^*-improvement of $\succ^*$.

THEOREM 3.2. *M is a monotonic Dodgson approximation with an approximation ratio of 2.*

PROOF. Clearly, the voting rule M is a Dodgson approximation (i.e., $\text{sc}_M(y, \succ) \geq \text{sc}_D(y, \succ)$ for each alternative $y \in A$ and profile $\succ \in \mathcal{L}^n$). Furthermore, it is monotonic by definition; if y is a winner for a profile $\succ$, y stays a winner for each y-improvement of $\succ$. In the following, we show that it has an approximation ratio of 2 as well. We will need the following lemma; informally, it states that by pushing an alternative upwards we cannot significantly decrease (that is, improve) the Dodgson score of another alternative.

LEMMA 3.3. *Let $y \in A$ and let $\succ, \succ' \in \mathcal{L}^n$ be profiles such that $\succ'$ is a y-improvement of $\succ$. For any alternative $z \in A \setminus \{y\}$ it holds that $sc_D(z, \succ) \leq 2 \cdot sc_D(z, \succ')$.*

PROOF. The fact that the Dodgson score of alternative z at profile $\succ'$ is $\text{sc}_D(z, \succ')$ means that z can become a Condorcet winner by pushing it $\text{sc}_D(z, \succ')$ positions upwards in the preference of some agents; let $\succ''$ be the resulting profile (where z is a Condorcet winner). For $i \in N$ denote

$$S_i = \{x \in A : \ x \succ'_i z \wedge z \succ''_i x\}.$$

Clearly, $\sum_{i \in N} |S_i| = \text{sc}_D(z, \succ')$.

Next consider the profile $\succ$ and observe that for all alternatives $x \in A \setminus \{z\}$, $\text{defc}(z, x, \succ) \leq \text{defc}(z, x, \succ')$. Hence, z can become a Condorcet winner by pushing it upwards in each agent i so that it bypasses all the alternatives in S_i. This involves no swaps at agent i if $S_i = \emptyset$ while pushing z $|S_i| + 1$ positions upwards at agent i is sufficient to bypass the alternatives in S_i and, possibly, alternative y that may lie in between them in $\succ$ (and not in $\succ'$). Hence, the Dodgson score of z at profile $\succ$ is

$$\text{sc}_D(z, \succ) \leq \sum_{i \in N : S_i \neq \emptyset} (|S_i| + 1) \leq 2 \sum_{i \in N} |S_i| = 2 \cdot \text{sc}_D(z, \succ'),$$

and the lemma follows. $\square$

Now, consider a profile $\succ \in \mathcal{L}^n$. Let y^* be the alternative in $W(\succ)$ with highest Dodgson score (equal to Δ). If y^* is a Dodgson winner at $\succ$ then $\text{sc}_M(z, \succ) = \text{sc}_D(z, \succ)$ for each alternative $z \in A$, and we are done. Hence we can assume that y^* is not a Dodgson winner but it belongs to $W(\succ)$.

By definition there must be a profile $\succ^* \in \mathcal{L}^n$ such that y^* is a Dodgson winner of $\succ^*$ and $\succ$ is a y^*-improvement of $\succ^*$. By Observation 3.1, since $\succ$ is a y^*-improvement of $\succ^*$, we have

$$\text{sc}_D(y^*, \succ) \leq \text{sc}_D(y^*, \succ^*). \tag{1}$$

Since y^* is a Dodgson winner of $\succ^*$, we have

$$\text{sc}_D(y^*, \succ^*) \leq \text{sc}_D(z, \succ^*) \tag{2}$$

for each alternative $z \in W(\succ)$. Also, by applying Lemma 3.3, we have

$$\text{sc}_D(z, \succ^*) \leq 2 \cdot \text{sc}_D(z, \succ) \tag{3}$$

for each alternative $z \in W(\succ)$. Now, using the definition of M and inequalities (1), (2), and (3), for any alternative $z \in W(\succ)$ we have

$$\text{sc}_M(z, \succ) = \Delta = \text{sc}_D(y^*, \succ) \leq 2 \cdot \text{sc}_D(z, \succ).$$

It remains to establish the approximation ratio with respect to the alternatives in $A \setminus W(\succ)$. Let $y' \in A \setminus \{y^*\}$ be a Dodgson winner of $\succ$. Since $y' \in W(\succ)$ the above inequality implies that

$$\Delta \leq 2 \cdot \text{sc}_D(y', \succ). \tag{4}$$

Let $z' \in A \setminus W(\succ)$. By definition, $\text{sc}_M(z', \succ) = \text{sc}_D(z', \succ)$ when $\text{sc}_D(z', \succ) > \Delta + 1$, and we are done. Otherwise it holds that $\text{sc}_M(z', \succ) = \Delta + 1$. Since z' is not a Dodgson winner for $\succ$, we have $\text{sc}_D(z', \succ) \geq \text{sc}_D(y', \succ) + 1$ in this case, and using (4) we obtain

$$\text{sc}_M(z', \succ) = \Delta + 1 \leq 2 \cdot \text{sc}_D(y', \succ) + 1 \leq 2 \cdot \text{sc}_D(z', \succ).$$

To conclude, the score of each alternative under M is at most twice its Dodgson score. $\square$

In general, the Dodgson approximation M is computable in exponential time. However, it can be implemented to run in polynomial time when m is a constant; in this special case the number of different profiles with n agents is polynomial and the Dodgson score can be computed exactly in polynomial time [4].

The next statement shows that the voting rule M is the best possible monotonic Dodgson approximation. Note that it is not based on any complexity assumptions and, hence, it holds for exponential-time Dodgson approximations as well.

THEOREM 3.4. *A monotonic Dodgson approximation cannot have an approximation ratio smaller than 2.*

The proof of the theorem appears in Appendix A.

3.2 A monotonic polynomial-time $\mathcal{O}(\log m)$-approximation algorithm

In the following we present a monotonic polynomial-time Dodgson approximation that achieves an approximation ratio of $\mathcal{O}(\log m)$. Given the $\Omega(\log m)$ inapproximability bound for the Dodgson score [8], this rule is asymptotically optimal with respect to polynomial-time algorithms. To be precise, it is optimal within a factor of 4, assuming that problems in $\mathcal{NP}$ do not have quasi-polynomial-time algorithms.

In general, there are two main obstacles that we have to overcome in order to implement the monotonization in polynomial time. First, the computation of the Dodgson score and the decision problem of detecting whether a given alternative is a Dodgson winner on a particular profile are $\mathcal{NP}$-hard problems [4]. We overcome this obstacle by using a polynomial-time Dodgson approximation R instead of the Dodgson score itself. Even in this case, given a profile, we still need to be able to detect whether an alternative $y \in A$ is the winner according to R at some profile of which the current profile is a y-improvement; if this is the case, y should be included in the winning set. Note that, in general, this requires checking an exponential number of profiles in order to determine the winning set of the current one. We tackle this second obstacle using the notion of *pessimistic estimators*; these are quantities defined in terms of the current profile only and are used to identify its winning alternatives.

In order to define the algorithm R that we will monotonize we consider an alternative definition of the Dodgson score for an alternative $z^* \in A$ and a profile $\succ \in \mathcal{L}^n$. Define the set $S_k^{\succ_i}(z^*)$ to be the set of alternatives z^* bypasses as it is pushed k positions upwards in the preference of agent i. Denote by $\mathcal{S}^{\succ_i}(z^*)$ the collection of all possible such sets for agent i, i.e.,

$$\mathcal{S}^{\succ_i}(z^*) = \{S_k^{\succ_i}(z^*) : k = 1, ..., r_i(z^*, \succ) - 1\},$$

where $r_i(z^*, \succ)$ denotes the rank of alternative z^* in the preferences of agent $i \in N$ (e.g., the most and least preferred alternatives have rank 1 and m, respectively). Let $\mathcal{S}(z^*) = \bigcup_{i \in N} \mathcal{S}^{\succ_i}(z^*)$. Then, the problem of computing the Dodgson score of alternative z^* on the profile $\succ$ is equivalent to selecting sets from $\mathcal{S}$ of minimum total size so that at most one set is selected among the ones in $\mathcal{S}^{\succ_i}(z^*)$ for each agent $i \in N$ and each alternative $z \in A \setminus \{z^*\}$ appears in at least $\mathrm{defc}(z^*, z, \succ)$ selected sets. This can be expressed by the following integer linear program:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} \sum_{k=1}^{r_i(z^*,\succ)-1} k \cdot \mathbf{x}\left(S_k^{\succ_i}(z^*)\right) \qquad (5) \\ \text{subject to} \quad & \forall z \in A \setminus \{z^*\}, \\ & \sum_{i \in N} \sum_{S \in \mathcal{S}^{\succ_i}(z^*) : z \in S} \mathbf{x}(S) \geq \mathrm{defc}(z^*, z, \succ) \\ & \forall i \in N, \sum_{S \in \mathcal{S}^{\succ_i}(z^*)} \mathbf{x}(S) \leq 1 \\ & \forall S \in \mathcal{S}(z^*), \ \mathbf{x}(S) \in \{0, 1\} \end{aligned}$$

The binary variable $\mathbf{x}(S)$ indicates whether the set $S \in \mathcal{S}(z^*)$ is selected ($\mathbf{x}(S) = 1$) or not ($\mathbf{x}(S) = 0$). Now, consider the LP relaxation of the above ILP in which the last constraint is relaxed to $\mathbf{x}(S) \geq 0$. We define the voting rule R that sets $\mathrm{sc}_R(z^*, \succ)$ equal to the optimal value of the LP relaxation multiplied by H_{m-1}, where H_k is the kth harmonic number. In [8] it is shown that

$$\mathrm{sc}_D(y, \succ) \leq \mathrm{sc}_R(y, \succ) \leq H_{m-1} \cdot \mathrm{sc}_D(y, \succ)$$

for every alternative $y \in A$, i.e., R is a Dodgson approximation with an approximation ratio of H_{m-1}. The following observation is analogous to Observation 3.1.

Observation 3.5. *Let $y \in A$ and let $\succ, \succ' \in \mathcal{L}^n$ be profiles such that $\succ'$ is a y-improvement of $\succ$. Then,*

$$sc_R(y, \succ') \leq sc_R(y, \succ).$$

We now present a new voting rule Q by monotonizing R. The voting rule Q defines a set of alternatives $W(\succ)$ that is the set of winners on a particular profile $\succ$. Then, it sets $\mathrm{sc}_Q(y, \succ) = 2 \cdot \mathrm{sc}_R(y^*, \succ)$ for each alternative $y \in W(\succ)$, where y^* is the winner according to the voting rule R. In addition, it sets $\mathrm{sc}_Q(y, \succ) = 2 \cdot \mathrm{sc}_R(y, \succ)$ for each alternative $y \notin W(\succ)$.

In order to define the set $W(\succ)$ we will use another (slightly different) linear program defined for two alternatives $y, z^* \in A$ and a profile $\succ \in \mathcal{L}^n$. The new LP has the same set of constraints as the relaxation of (5) used in the definition of $\mathrm{sc}_R(z^*, \succ)$ and the following objective function:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} \sum_{k=1}^{r_i(z^*,\succ)-1} k \cdot \mathbf{x}\left(S_k^{\succ_i}(z^*)\right) \qquad (6) \\ & + \sum_{i \in N : y \succ_i z^*} \sum_{k=1}^{r_i(z^*,\succ) - r_i(y,\succ)-1} \mathbf{x}\left(S_k^{\succ_i}(z^*)\right) \end{aligned}$$

We define the *pessimistic estimator* $\mathrm{pe}(z^*, y, \succ)$ for alternative $z^* \in A$ with respect to another alternative $y \in A \setminus \{z^*\}$ and a profile $\succ \in \mathcal{L}^n$ to be equal to the objective value of LP (6) multiplied by H_{m-1}. As will become apparent shortly, the pessimistic estimator $\mathrm{pe}(z^*, y, \succ')$ upper-bounds the score of alternative z^* under R on every profile $\succ$ such that $\succ'$ is a y-improvement of $\succ$, hence the pessimism with respect to estimating the score of z^*. These pessimistic estimators will be our main tool in order to monotonize R.

We are now ready to complete the definition of the voting rule Q. The set $W(\succ)$ is defined as follows. First, it contains all the winners according to voting rule R. An alternative y that is not a winning alternative according to R is included in the set $W(\succ)$ if $\mathrm{pe}(z, y, \succ) \geq \mathrm{sc}_R(y, \succ)$ for every alternative $z \in A \setminus \{y\}$.

Clearly, Q is polynomial-time; computing the scores of all alternatives involves solving m^2 linear programs of polynomial size. We next show that it is monotonic as well. This is done in Lemma 3.8 after establishing some properties of pessimistic estimators. The first property (stated in Lemma 3.6) has a long and technically involved proof that is relegated to Appendix B. The second one (in Lemma 3.7) follows easily by the definition of pessimistic estimators.

Lemma 3.6. *Let $y, z^* \in A$ be different alternatives and let $\succ, \succ' \in \mathcal{L}^n$ be two profiles such that $\succ'$ is a y-improvement of $\succ$. Then, $pe(z^*, y, \succ') \geq pe(z^*, y, \succ)$.*

Lemma 3.7. *Let $y, z^* \in A$ be different alternatives and let $\succ \in \mathcal{L}^n$ be a profile. Then, $sc_R(z^*, \succ) \leq pe(z^*, y, \succ) \leq 2 \cdot sc_R(z^*, \succ)$.*

Proof. The lemma follows directly from the observations that the objective of the linear program (6) is lower-bounded by the objective of the linear programming relaxation of (5) and also upper-bounded by the latter multiplied by 2. □

Lemma 3.8. *The voting rule Q is monotonic.*

Proof. Let $y \in A$ and consider a profile $\succ \in \mathcal{L}^n$ such that $y \in W(\succ)$. We will show that $y \in W(\succ')$ for each profile $\succ'$ which is a y-improvement of $\succ$. This is clearly true if y is a winning alternative according to R at $\succ'$. If this is not the case, we distinguish between two cases:

Case 1. y is a winning alternative according to R at $\succ$. Then for any alternative $z \in A \setminus \{y\}$ we have

$$\begin{aligned} \mathrm{pe}(z, y, \succ') & \geq \mathrm{pe}(z, y, \succ) \geq \mathrm{sc}_R(z, \succ) \geq \mathrm{sc}_R(y, \succ) \\ & \geq \mathrm{sc}_R(y, \succ'), \end{aligned}$$

therefore $y \in W(\succ')$. The first inequality follows by Lemma 3.6, the second follows by Lemma 3.7, the third is true since y is the winner under R in profile $\succ$, and the fourth follows from Observation 3.5.

Case 2. y is not a winning alternative according to R at $\succ$. Since $y \in W(\succ)$ it must hold that $\mathrm{pe}(z, y, \succ) \geq \mathrm{sc}_R(y, \succ)$

for any alternative $z \in A \setminus \{y\}$. We therefore have that

$$\mathrm{pe}(z, y, \succ') \geq \mathrm{pe}(z, y, \succ) \geq \mathrm{sc}_R(y, \succ) \geq \mathrm{sc}_R(y, \succ')$$

for any alternative $z \in A \setminus \{y\}$ and, hence, $y \in W(\succ')$. The first inequality follows by Lemma 3.6 and the third is implied by Observation 3.5. □

The following lemma provides the desired bound on the approximation ratio.

LEMMA 3.9. *Q is a Dodgson approximation with an approximation ratio of $2H_{m-1}$.*

PROOF. We have to show that

$$\mathrm{sc}_D(y, \succ) \leq \mathrm{sc}_Q(y, \succ) \leq 2H_{m-1} \cdot \mathrm{sc}_D(y, \succ)$$

for any alternative $y \in A$ and profile $\succ \in \mathcal{L}^n$. This is clearly the case if y is a winning alternative according to R or $y \notin W(\succ)$ since $\mathrm{sc}_Q(y, \succ) = 2 \cdot \mathrm{sc}_R(y, \succ)$ (by the definition of Q) and

$$\mathrm{sc}_D(y, \succ) \leq \mathrm{sc}_R(y, \succ) \leq H_{m-1} \cdot \mathrm{sc}_D(y, \succ)$$

since R is a Dodgson approximation with approximation ratio H_{m-1}.

Now assume that y is not a winning alternative according to R but it belongs to $W(\succ)$. Let z be a winning alternative according to R. Since $y \in W(\succ)$, it must be the case that $\mathrm{pe}(z, y, \succ) \geq \mathrm{sc}_R(y, \succ)$. So, using in addition the definition of Q, Lemma 3.7, and the fact that R is a Dodgson approximation, we have

$$\begin{aligned} \mathrm{sc}_Q(y, \succ) &= 2 \cdot \mathrm{sc}_R(z, \succ) \geq \mathrm{pe}(z, y, \succ) \geq \mathrm{sc}_R(y, \succ) \\ &\geq \mathrm{sc}_D(y, \succ). \end{aligned}$$

Furthermore, using the definition of Q, the fact that z is a winning alternative under R, and the approximation bound of R, we have

$$\begin{aligned} \mathrm{sc}_Q(y, \succ) &= 2 \cdot \mathrm{sc}_R(z, \succ) \leq 2 \cdot \mathrm{sc}_R(y, \succ) \\ &\leq 2H_{m-1} \cdot \mathrm{sc}_D(y, \succ). \end{aligned}$$

□

We summarize the discussion above with the following statement.

THEOREM 3.10. *Q is a monotonic polynomial-time Dodgson approximation with an approximation ratio of $2H_{m-1}$.*

4. HOMOGENEITY

In this section we present our results on homogeneous Dodgson approximations. A voting rule is *homogeneous* if duplicating the electorate, that is, duplicating the preference profile, does not change the outcome of the election. An example (due to Brandt [6]) that demonstrates that Dodgson's rule fails homogeneity can be found in Table 2. The intuition is that if alternatives x and y are tied in a pairwise election, the deficit of x against y does not increase by duplicating the profile, whereas if x strictly loses to y in a pairwise election then the deficit scales with the number of copies.

$\times 2$	$\times 2$	$\times 2$	$\times 2$	$\times 2$	$\times 1$	$\times 1$
d	b	c	d	a	a	d
c	c	a	b	b	d	a
a	a	b	c	c	b	b
b	d	d	a	d	c	c

Table 2: An example that demonstrates that Dodgson's rule does not satisfy homogeneity. A column headed by $\times k$ represents k identical agents. In this profile, a is the Dodgson winner with a score of 3. By duplicating the electorate three times we obtain a profile in which the winner is d with a score of 6.

4.1 The Simplified Dodgson Rule

Tideman [22, pages 199-201] defines the following simplified Dodgson rule and proves that it is monotonic and homogeneous. Consider a profile $\succ \in \mathcal{L}^n$. If an alternative is a Condorcet winner, then this alternative is the sole winner. Otherwise, the simplified Dodgson rule assigns a score to each alternative and the alternative with the minimum score wins. According to the simplified Dodgson rule, the score of an alternative x is

$$\mathrm{sc}_{\mathrm{Td}}(x, \succ) = \sum_{y \in A \setminus \{x\}} \max\{0, n - 2 \cdot |\{i \in N : x \succ_i y\}|\}.$$

Observe that $\mathrm{sc}_{\mathrm{Td}}(x, \succ)$ can be smaller than the Dodgson score of x and, hence, this definition does not correspond to a Dodgson approximation. For example, in profiles with an even number of agents, $\mathrm{sc}_{\mathrm{Td}}(x, \succ)$ is 0 when x is tied against some alternatives and beats the rest. Hence, we present an alternative definition of the simplified Dodgson rule as a Dodgson approximation by scaling the original definition. If an alternative x is a Condorcet winner, then it has score $\mathrm{sc}_{\mathrm{Td}'}(x, \succ) = 0$. Otherwise:

$$\mathrm{sc}_{\mathrm{Td}'}(x, \succ) = m \cdot \mathrm{sc}_{\mathrm{Td}}(x, \succ) + m(\log m + 1).$$

It is clear that this alternative definition is equivalent to the original one of the simplified Dodgson rule, in the sense that it elects the same set of alternatives. It is also clear that $\mathrm{sc}_{\mathrm{Td}}(x, \succ)$ can be computed in polynomial time, and, as mentioned above, Td is known to be monotonic and homogeneous. Hence, in order to prove the following theorem it is sufficient to prove that Td$'$ is a Dodgson approximation and to bound its approximation ratio.

THEOREM 4.1. *Td$'$ is a monotonic, homogeneous, polynomial-time Dodgson approximation with an approximation ratio of $\mathcal{O}(m \log m)$.*

PROOF. We will show that, given any profile $\succ \in \mathcal{L}^N$ and alternative $x \in A$, it holds that $\mathrm{sc}_D(x, \succ) \leq \mathrm{sc}_{\mathrm{Td}'}(x, \succ) \leq m(\log m + 3) \cdot \mathrm{sc}_D(x, \succ)$. We will consider the case in which x is not a Condorcet winner since otherwise the inequalities clearly hold.

In order to show that Td$'$ is a Dodgson approximation we distinguish between two cases.

If the number of agents is odd, then $\mathrm{sc}_{\mathrm{Td}}(x, \succ)$ can be expressed in terms of the deficits of alternative x against the other alternatives as follows:

$$\mathrm{sc}_{\mathrm{Td}}(x, \succ) = \sum_{y \in A \setminus \{x\}} \max\{0, 2 \cdot \mathrm{defc}(x, y, \succ) - 1\}.$$

Observe that each term of the above sum is non-zero only when $\mathrm{defc}(x, y, \succ) > 0$. Since $2 \cdot \mathrm{defc}(x, y, \succ) - 1 \geq \mathrm{defc}(x, y, \succ)$ in this case, we have that

$$\begin{aligned} \mathrm{sc}_{\mathrm{Td}'}(x, \succ) &= m \cdot \mathrm{sc}_{\mathrm{Td}}(x, \succ) + m(\log m + 1) \\ &> m \cdot \mathrm{sc}_{\mathrm{Td}}(x, \succ) \\ &\geq m \sum_{y \in A \setminus \{x\}} \mathrm{defc}(x, y, \succ) \\ &\geq \mathrm{sc}_D(x, \succ). \end{aligned}$$

If the number of agents is even, then $\mathrm{sc}_{\mathrm{Td}}(x, \succ)$ can be expressed in terms of the deficits of alternative x against the other alternatives as follows:

$$\mathrm{sc}_{\mathrm{Td}}(x, \succ) = \sum_{y \in A \setminus \{x\}} \max\{0, 2 \cdot \mathrm{defc}(x, y, \succ) - 2\}.$$

Let

$$S_x = \{y \in A \setminus \{x\} : \mathrm{defc}(x, y, \succ) \geq 2\}$$

and

$$T_x = \{y \in A \setminus \{x\} : \mathrm{defc}(x, y, \succ) = 1\}.$$

We will now prove that it is sufficient to push x to the top of the preferences of at most $\log m + 1$ agents in order to cover the deficits against the alternatives in T_x. Since the number of agents n is even, the fact that $\mathrm{defc}(x, y, \succ) = 1$ means that exactly $n/2$ agents rank x above y. For every $i \in N$, let $A_i = \{y \in T_x : y \succ_i x\}$. By the pigeonhole principle there exists an agent i_1 such that $|A_{i_1}| \geq |T_x|/2$; we add i_1 to our cover, and denote $X_1 = T_x \setminus A_{i_1}$. Next, there must exist an agent i_2 such that $|A_{i_2} \cap X_1| \geq \frac{|X_1|}{2}$. We add i_2 to our cover, and define $X_2 = X_1 \setminus A_{i_2}$. Continuing inductively in this way, we cover all the alternatives in T_x after $\log |T_x| + 1 \leq \log m + 1$ steps.

Moreover, in order to make x beat the alternatives in S_x it suffices to push it to the top of the preferences of at most $\sum_{y \in S_x} \mathrm{defc}(x, y, \succ)$ agents. Hence, the Dodgson score of x is

$$\mathrm{sc}_D(x, \succ) \leq m \sum_{y \in S_x} \mathrm{defc}(x, y, \succ) + m(\log m + 1).$$

Now observe that each term of the sum in the equivalent definition of $\mathrm{sc}_{\mathrm{Td}}(x, \succ)$ is non-zero only when $\mathrm{defc}(x, y, \succ) \geq 2$ (i.e., when $y \in S_x$). Since $2 \cdot \mathrm{defc}(x, y, \succ) - 2 \geq \mathrm{defc}(x, y, \succ)$ in this case, we have that

$$\begin{aligned} \mathrm{sc}_{\mathrm{Td}'}(x, \succ) &= m \cdot \mathrm{sc}_{\mathrm{Td}}(x, \succ) + m(\log m + 1) \\ &\geq m \sum_{y \in S_x} \mathrm{defc}(x, y, \succ) + m(\log m + 1) \\ &\geq \mathrm{sc}_D(x, \succ). \end{aligned}$$

We have completed the proof that Td' is a Dodgson approximation. In order to prove the bound on the approximation ratio, in both cases we have

$$\begin{aligned} \mathrm{sc}_{\mathrm{Td}'}(x, \succ) &= m \cdot \mathrm{sc}_{\mathrm{Td}}(x, \succ) + m(\log m + 1) \\ &\leq 2m \sum_{y \in A \setminus \{x\}} \mathrm{defc}(x, y, \succ) + m(\log m + 1) \\ &\leq m(\log m + 3) \cdot \mathrm{sc}_D(x, \succ). \end{aligned}$$

The last inequality holds since $\mathrm{sc}_D(x, \succ)$ is lower bounded by both $\sum_{y \in A \setminus \{x\}} \mathrm{defc}(x, y, \succ)$ and 1. $\square$

4.2 Lower Bound

We next show that Td' is the asymptotically optimal homogeneous Dodgson approximation by proving a matching lower bound on the approximation ratio of homogeneous Dodgson approximations. The lower bound is not based on any complexity assumptions and holds for exponential-time Dodgson approximations as well. This is quite striking since, as stated in Theorem 4.1, Td' is also monotonic and polynomial-time.

THEOREM 4.2. *Any homogeneous Dodgson approximation has approximation ratio at least $\Omega(m \log m)$.*

The proof is based on the construction of a preference profile with an alternative $b \in A$ that defeats some of the alternatives in pairwise elections, and is tied against many others. Hence, it has a high Dodgson score. On the other hand, there is a second alternative that has a Dodgson score of two, simply because it has a deficit of two against another alternative. In order to obtain a good approximation ratio, the algorithm must not select b in this profile. However, when the profile is replicated, the Dodgson score of b does not increase: it is still tied against the same alternatives. In contrast, the Dodgson score of the other alternatives scales with the number of copies. By homogeneity, we cannot select b in the replicated profile, which yields the lower bound.

We can think of an agent as the subset of alternatives that are ranked above b. If b is tied against an alternative, then that alternative is a member of exactly half the subsets. The argument used in the proof of Theorem 4.1 implies that there is always a cover of logarithmic size; the proof of Theorem 4.2 establishes that this bound is tight. Indeed, the combinatorial core of the theorem's proof is the construction of a set cover instance with the following properties: each element of the ground set appears in roughly half the subsets, but every cover requires a logarithmic number of subsets (see Claim 4.3). This (apparently novel) construction is due to Noga Alon [2].

PROOF OF THEOREM 4.2. Given an integer $r \geq 3$, we construct a preference profile $\succ$ with $n = 2^r$ agents and $m = 2^{r+1} + 1$ alternatives. There is a set $X = \{x_1, x_2, \ldots, x_{2^r - 1}\}$ with $2^r - 1$ alternatives, two sets Y and Z with 2^{r-1} alternatives each, and two additional alternatives a and b.

For $i = 1, ..., 2^r - 1$, denote by X_i the set of alternatives x_j such that the inner product of the binary vectors corresponding to the binary representations of i and j equals 1 modulo 2. Denote by $\mathcal{X}$ the collection of all sets X_i for $i = 1, ..., 2^r - 1$.

CLAIM 4.3. *The sets of $\mathcal{X}$ have the following properties:*

1. *Each alternative $x \in X$ belongs to 2^{r-1} different sets of $\mathcal{X}$.*
2. *Each set of $\mathcal{X}$ contains exactly 2^{r-1} alternatives.*
3. *There are r different sets in $\mathcal{X}$ whose union contains all alternatives in X.*
4. *For each subcollection of at most $r - 1$ sets in $\mathcal{X}$, there exists an alternative of X that does not belong to their union.*

PROOF. Properties 1 and 2 follow easily by the definition of the sets in $\mathcal{X}$.

0	1	2	...	2^{r-1}	$2^{r-1}+1$	...	2^r-1
b	b	a	...	a	$X_{2^{r-1}+1}$	...	X_{2^r-1}
a	a	X_2	...	$X_{2^{r-1}}$	Z	...	Z
Y	X_1	Y	...	Y	b	...	b
Z	Z	b	...	b	Y	...	Y
X	Y	Z	...	Z	$X \setminus X_{2^{r-1}+1}$	...	$X \setminus X_{2^r-1}$
	$X \setminus X_1$	$X \setminus X_2$	...	$X \setminus X_{2^{r-1}}$	a	...	a

Table 3: The preference profile $\succ$ used in the proof of Theorem 4.2.

In order to establish property 3, it suffices to consider the r sets X_{2^i} for $i = 0, ..., r-1$, i.e., the ones whose binary representation has just one 1 in the $(i+1)$-th bit position.

Turning to property 4, we consider a binary r-vector $\mathbf{z} = \langle z_1, z_2, ..., z_r \rangle \in \{0,1\}^r$. Now, consider the set X_k and let $\langle b_1(k), b_2(k), ..., b_r(k) \rangle$ be the r-vector corresponding to the binary representation of k. We have that the equation $\sum_{j=1}^{r} b_j(k) \cdot z_j = 0 \bmod 2$ is true if and only if the alternative such that $\mathbf{z}$ is the binary representation of its index is not contained in set X_k. Since any homogeneous system of less than r linear equations modulo 2 with r unknowns has a nontrivial (i.e., nonzero) solution, it follows that for any subcollection of less than r sets in $\mathcal{X}$, there exists an alternative in X that is not contained in their union. $\square$

We construct the preference profile $\succ$ as follows (see Table 3):

- Agent 0 ranks b first, then a, then the alternatives in Y (in arbitrary order), then the alternatives in Z (also in arbitrary order), and then the alternatives of X (in arbitrary order).
- Agent 1 ranks b first, then a, then the alternatives in X_1 (in arbitrary order), then the alternatives in Z, then the alternatives in Y, and then the alternatives in $X \setminus X_1$ (in arbitrary order).
- For $i = 2, ..., 2^{r-1}$, agent i ranks a first, then the alternatives of X_i (in arbitrary order), then the alternatives in Y, then b, then the alternatives in Z, and then the alternatives in $X \setminus X_i$ (in arbitrary order).
- For $i = 2^{r-1}+1, ..., 2^r-1$, agent i ranks the alternatives in X_i (in arbitrary order) first, then the alternatives of Z, then b, then the alternatives in Y, then the alternatives in $X \setminus X_i$ (in arbitrary order), and then a.

The next four claims state important properties of the profile $\succ$.

Claim 4.4. *The Dodgson score of a is at most 2.*

Proof. After swapping a and b in the rankings of agents 0 and 1 alternative a is ranked first by a majority of agents, hence it clearly becomes the Condorcet winner. $\square$

Claim 4.5. *Alternative b has deficit at most 1 against any other alternative.*

Proof. By property 1 of Claim 4.3 and the construction of the profile, we have that b is ranked below any alternative x_i of $X \setminus X_1$ by 2^{r-1} agents, that is, b is tied with these alternatives in pairwise elections. It follows that $\text{defc}(b, x_i, \succ) = 1$. In addition, b is ranked above any alternative in $X_1 \cup Y \cup Z \cup \{a\}$ by $2^{r-1}+1$ agents, i.e., $\text{defc}(b, x, \succ) = 0$ for any alternative $x \in X_1 \cup Y \cup Z \cup \{a\}$. $\square$

Claim 4.6. $r2^{r-2} \leq sc_D(b, \succ) \leq (r-1)2^r$.

Proof. By property 4 of Claim 4.3, alternative b has to be pushed upwards in the rankings of at least $r-1$ among the agents $2, ..., 2^r-1$ in order to eliminate its deficit against the $2^{r-1}-1$ alternatives of $X \setminus X_1$. This requires at least $(r-1)2^{r-1}$ swaps in order to push above the alternatives of Y (in the case of agents $2, \ldots, 2^{r-1}$) or Z (in the case of agents $2^{r-1}+1, \ldots, 2^r-1$) in the rankings of $r-1$ agents, plus at least $2^{r-1}-1$ additional swaps in order to defeat each of the alternatives in $X \setminus X_1$; the total is $r2^{r-1}-1 \geq r2^{r-2}$ swaps.

The upper bound follows by properties 2 and 3 of Claim 4.3, since b becomes a Condorcet winner by pushing it above the alternatives of X in the rankings of at most $r-1$ additional agents, and using at most $|X_i| + |Y| = 2^r$ or $|X_i| + |Z| = 2^r$ swaps per agent. $\square$

Claim 4.7. *Any alternative besides b has deficit at least 2 against some other alternative.*

Proof. Alternative a is ranked higher than alternative b by $2^{r-1}-1$ agents; so, it holds that $\text{defc}(a, b, \succ) = 2$. Moreover, alternative a is ranked higher than the alternatives in X, Y and Z by $2^{r-1}+1$ agents. So, $\text{defc}(x, a, \succ) = 2$ for any alternative $x \in X \cup Y \cup Z$. $\square$

Now, consider a homogeneous Dodgson approximation H. If it selects b as the winner of profile $\succ$ then, using Claims 4.4 and 4.6, and since H is a Dodgson approximation, we have

$$\text{sc}_H(a, \succ) \geq \text{sc}_H(b, \succ) \geq r2^{r-2} \geq r2^{r-3}\text{sc}_D(a, \succ).$$

Hence, H has an approximation ratio of at least

$$r2^{r-3} = \frac{m-1}{16} \cdot \log \frac{m-1}{2} = \Omega(m \log m) \quad .$$

Assume otherwise that the winner under H is some alternative $x \in A \setminus \{b\}$. Consider the preference profile $\succ'$ obtained by making $r(r-1)2^{2r-3}$ copies of the profile $\succ$. By Claim 4.5, we have that b has deficit at most 1 against any other alternative in the new profile as well; its Dodgson score in the new profile is in the range defined in Claim 4.6, i.e., $\text{sc}_H(b, \succ') \leq (r-1)2^r$. By the definition of the deficit and Claim 4.7, we have that alternative x has deficit at least $r(r-1)2^{2r-3}$ against some other alternative and, hence, its Dodgson score in the new profile is $\text{sc}_H(x, \succ') \geq r(r-1)2^{2r-3}$.

By the homogeneity property, x should be a winner under H in the profile $\succ'$. Then,

$$\begin{aligned} \text{sc}_H(b, \succ') &\geq \text{sc}_H(x, \succ') \geq \text{sc}_D(x, \succ') \\ &\geq r(r-1)2^{2r-3} \geq r2^{r-3}\text{sc}_D(b, \succ'). \end{aligned}$$

Therefore, H has approximation ratio $r2^{r-3} = \Omega(m \log m)$ in this case as well. $\square$

5. ADDITIONAL PROPERTIES

In this section we briefly summarize our results with respect to several additional social choice properties that are not satisfied by Dodgson's rule. In general, our lower bounds with respect to these properties are at least linear in n, the number of agents. Since n is almost always large, these results should strictly be interpreted as impossibility results, that is, normally an upper bound of $\mathcal{O}(n)$ is not useful. We now (informally) formulate the five properties in question; for more formal definitions the reader is referred to [22].

We say that a voting rule satisfies *combinativity* if, given two preference profiles where the rule elects the same winning set, the rule would also elect this winning set under the profile obtained from appending one of the original preference profiles to the other. Note that combinativity implies homogeneity.

A *dominating set* is a nonempty set of alternatives such that each alternative in the set beats every alternative outside the set in pairwise elections. The *Smith set* is the unique inclusion-minimal dominating set. A voting rule satisfies *Smith consistency* if winners under the rule are always contained in the Smith set.

We say that a voting rule satisfies *mutual majority consistency* if, given a preference profile where more than half the agents rank a subset of alternatives $X \subseteq A$ above $A \setminus X$, only alternatives from X can be elected. A voting rule satisfies *invariant loss consistency* if an alternative that loses to every other alternative in pairwise elections cannot be elected. Clearly, mutual majority consistency implies invariant loss consistency.

Independence of clones was introduced by Tideman [21]; see also the paper by Schulze [20]. For ease of exposition we use a slightly weaker definition previously employed by Brandt [6]; since we are proving a lower bound, a weaker definition only strengthens the bound. Given a preference profile, two alternatives $x, y \in A$ are considered *clones* if they are adjacent in the rankings of all the agents, that is, their order with respect to every alternative in $A \setminus \{x, y\}$ is identical everywhere. A voting rule is *independent of clones* if a losing alternative cannot be made a winning alternative by introducing clones.

We have the following theorem. The proof is relegated to Appendix C.

THEOREM 5.1. *Let V be a Dodgson approximation. If V satisfies combinativity or Smith consistency, then its approximation ratio is at least $\Omega(nm)$. If V satisfies mutual majority consistency, invariant loss consistency, or independence of clones, then its approximation ratio is at least $\Omega(n)$.*

6. REFERENCES

[1] N. Ailon, M. Charikar, and A. Newman. Aggregating inconsistent information: Ranking and clustering. In *Proc. of 37th STOC*, pages 684–693, 2005.

[2] N. Alon. Private communication, 2009.

[3] A. Archer and É. Tardos. Truthful mechanisms for one-parameter agents. In *Proc. of 42nd FOCS*, pages 482–491, 2001.

[4] J. Bartholdi, C. A. Tovey, and M. A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6:157–165, 1989.

[5] D. Black. *Theory of Committees and Elections.* Cambridge University Press, 1958.

[6] F. Brandt. Some remarks on Dodgson's voting rule. *Mathematical Logic Quarterly*, 55(4):460–463, 2009.

[7] F. Brandt, F. Fischer, and P. Harrenstein. The computational complexity of choice sets. *Mathematical Logic Quarterly*, 55(4):444–459, 2009.

[8] I. Caragiannis, J. A. Covey, M. Feldman, C. M. Homan, C. Kaklamanis, N. Karanikolas, A. D. Procaccia, and J. S. Rosenschein. On the approximability of Dodgson and Young elections. In *Proc. of 20th SODA*, pages 1058–1067, 2009.

[9] D. Coppersmith, L. Fleischer, and A. Rudra. Ordering by weighted number of wins gives a good ranking for weighted tournaments. In *Proc. of 17th SODA*, pages 776–782, 2006.

[10] Marquis de Condorcet. Essai sur l'application de l'analyse à la probabilité de décisions rendues à la pluralité de voix. Imprimerie Royal, 1785. Facsimile published in 1972 by Chelsea Publishing Company, New York.

[11] P. C. Fishburn. Condorcet social choice functions. *SIAM Journal on Applied Mathematics*, 33(3):469–487, 1977.

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

[13] E. Hemaspaandra, L. A. Hemaspaandra, and J. Rothe. Exact analysis of Dodgson elections: Lewis Carroll's 1876 voting system is complete for parallel access to NP. *Journal of the ACM*, 44(6):806–825, 1997.

[14] C. Homan and L. A. Hemaspaandra. Guarantees for the success frequency of an algorithm for finding Dodgson election winners. *Journal of Heuristics*, 15(4):403–423, 2009.

[15] C. Kenyon-Mathieu and W. Schudy. How to rank with few errors. In *Proc. of 39th STOC*, pages 95–103, 2007.

[16] J. C. McCabe-Dansted, G. Pritchard, and A. M. Slinko. Approximability of Dodgson's rule. *Social Choice and Welfare*, 31(2):311–330, 2008.

[17] N. Nisan. Introduction to mechanism design (for computer scientists). In N. Nisan, T. Roughgarden, É. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, chapter 9. Cambridge University Press, 2007.

[18] N. Nisan and A. Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35(1–2):166–196, 2001.

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

[20] M. Schulze. A new monotonic and clone-independent single-winner election method. *Voting Matters*, 17:9–19, 2003.

[21] N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

[22] N. Tideman. *Collective Decisions and Voting.*

$\times k$	$\times k$	$\times k$	$\times k$	$\times k$	$\times k$	$\times 2k$	$\times 2k$	$\times 2k$	$\times 2k$	$\times 2k$	$\times 2k$
c	c	a	a	b	b	a	a	b	b	c	c
b	b	c	c	a	a	d	e	d	e	d	e
a	a	b	b	c	c	b	b	c	c	a	a
d	e	d	e	d	e	e	d	e	d	e	d
e	d	e	d	e	d	c	c	a	a	b	b

Table 4: The preference profile $\succ$ used in the proof of Theorem 3.4.

Ashgate, 2006.

APPENDIX

A. PROOF OF THEOREM 3.4

Let k be a positive integer. We use the preference profile $\succ$ (see Table 4) with $18k$ agents and five alternatives a, b, c, d, and e.

The deficit of d and e against any of the alternatives a, b, and c is $3k+1$ and d is tied against e, i.e.,

$$\text{defc}(d, e, \succ) = \text{defc}(e, d, \succ) = 1.$$

Moreover,

$$\text{defc}(a, c, \succ) = \text{defc}(b, a, \succ) = \text{defc}(c, b, \succ) = k + 1$$

while the remaining deficits are zero.

Hence, the Dodgson score of d and e is at least $9k + 4$. Also, observe that c is ranked either below a or two positions above a by each agent. It follows that a must be pushed two positions upwards in $k + 1$ agents in order to become a Condorcet winner. This is also sufficient since a becomes a Condorcet winner by pushing it two positions upwards in $k + 1$ agents among those in the first and second column of profile $\succ$. Hence, $\text{sc}_D(a, \succ) = 2(k + 1)$. With a similar argument we obtain that

$$\text{sc}_D(b, \succ) = \text{sc}_D(c, \succ) = 2(k + 1)$$

as well.

Now, consider any monotonic approximation algorithm M' for Dodgson. Given $\succ$, if it returns alternative d or e as a winner then

$$\begin{aligned} \text{sc}_{M'}(a, \succ) &\geq \min\{\text{sc}_{M'}(d, \succ), \text{sc}_{M'}(e, \succ)\} \\ &\geq \min\{\text{sc}_D(d, \succ), \text{sc}_D(e, \succ)\} \\ &\geq 9k + 4 > 4(k + 1) = 2 \cdot \text{sc}_D(a, \succ), \end{aligned}$$

i.e., the approximation ratio is more than 2.

If the algorithm returns alternative a as the winner, then consider the profile $\succ^a$ in which a is pushed one position upwards in the $2k$ agents in the fifth and sixth column of profile $\succ$. Due to monotonicity, a should be the winner in the profile $\succ^a$ as well. Also, note that now c can become a Condorcet winner by pushing it just one position upwards in $k+1$ among the $2k$ agents in the fifth and sixth columns of profile $\succ^a$ and, hence, $\text{sc}_D(c, \succ^a) = k+1$ while the Dodgson score of a in $\succ^a$ is the same as in $\succ$. We have

$$\begin{aligned} \text{sc}_{M'}(c, \succ^a) &\geq \text{sc}_{M'}(a, \succ^a) \geq \text{sc}_D(a, \succ^a) \\ &= 2(k + 1) = 2 \cdot \text{sc}_D(c, \succ^a), \end{aligned}$$

i.e., the approximation ratio is at least 2.

If the algorithm returns alternative b as the winner, then consider the profile $\succ^b$ in which b is pushed one position upwards in the $2k$ agents in the first and second column of profile $\succ$. Due to monotonicity, we have that b is a winner in the profile $\succ^b$ as well. Also, note that now a can become a Condorcet winner by pushing it just one position upwards in $k+1$ among the $2k$ agents in the first and second columns of profile $\succ^b$ and, hence, $\text{sc}_D(a, \succ^b) = k+1$. Using a similar calculation we get that the approximation ratio is at least 2.

Finally, if the algorithm returns alternative c as the winner, then consider the profile $\succ^c$ in which c is pushed one position upwards in the $2k$ agents in the third and fourth columns of profile $\succ$. This time we have that $\text{sc}_D(b, \succ^c) = k + 1$, and we obtain the lower bound of 2 as before. We have thus covered all the possible cases, and the theorem follows. $\square$

We remark that in the proof above it is possible to replace each column of k or $2k$ agents with a single agent or two agents, respectively. However, the current proof has the advantage of demonstrating that the approximation ratio cannot improve if the number of agents is assumed to be large. Clearly, it is also possible to increase the number of alternatives by adding more alternatives at the bottom of the agents' preferences.

B. PROOF OF LEMMA 3.6

Consider the LP (6) used in the definition of $\text{pe}(z^*, y, \succ')$ and let $\mathbf{x}$ be an optimal solution. We will show how to transform this solution to a feasible solution $\bar{\mathbf{x}}$ for the LP (6) used in the definition of $\text{pe}(z^*, y, \succ)$ in such a way that the objective of the latter is not larger than that of the former LP.

We partition the set of agents N into three disjoint sets of agents N_1, N_2, and N_3 and define the solution $\bar{\mathbf{x}}$ as follows:

Denote by N_1 the set of agents such that either y is ranked at the same position in $\succ$ and $\succ'$ or y is ranked below z^* in both profiles $\succ$ and $\succ'$. In this case, we set $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_k^{\succ'_i})$, for each agent $i \in N_1$ and $k = 1, 2, ..., r_i(z^*, \succ) - 1$. Observe that $S_k^{\succ_i} = S_k^{\succ'_i}$ for each agent $i \in N_1$ and $k = 1, ..., r_i(z^*, \succ) - 1$.

Denote by N_2 the set of agents such that y is ranked above z^* in both profiles $\succ$ and $\succ'$ and, furthermore $r_i(y, \succ) > r_i(y, \succ')$. In this case, for each agent $i \in N_2$ we set

- $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_k^{\succ'_i})$ for $k = 1, ..., r_i(z^*, \succ) - r_i(y, \succ) - 1$;
- $\bar{\mathbf{x}}(S^{\succ_i}_{r_i(z^*,\succ)-r_i(y,\succ)}) = 0$;
- $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_{k-1}^{\succ'_i})$ for $k = r_i(z^*, \succ) - r_i(y, \succ) + 1, ..., r_i(z^*, \succ) - r_i(y \succ') - 1$;
- $\bar{\mathbf{x}}(S^{\succ_i}_{r_i(z^*,\succ)-r_i(y,\succ')}) = \mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ)-r_i(y,\succ')-1}) + \mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ)-r_i(y,\succ')})$;

- $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_k^{\succ'_i})$ for $k = r_i(z^*, \succ) - r_i(y, \succ') + 1, ..., r_i(z^*, \succ) - 1$.

Denote by $N_3 = N \setminus (N_1 \cup N_2)$ the set of agents such that y is ranked above z^* in $\succ'$ but below z^* in $\succ$. Note that $r_i(z^*, \succ') = r_i(z^*, \succ) + 1$ in this case. For each agent $i \in N_3$ we set

- $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_k^{\succ'_i})$ for $k = 1, ..., r_i(z^*, \succ) - r_i(y, \succ') - 1$;
- $\bar{\mathbf{x}}(S^{\succ_i}_{r_i(z^*,\succ)-r_i(y,\succ')}) = \mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ)-r_i(y,\succ')}) + \mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ)-r_i(y,\succ')+1})$;
- $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_{k+1}^{\succ'_i})$ for $k = r_i(z^*, \succ) - r_i(y, \succ') + 1, ..., r_i(z^*, \succ) - 1$.

Now consider the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ)$. Clearly, since the solution $\mathbf{x}$ is non-negative, the solution $\bar{\mathbf{x}}$ is non-negative as well. The definitions above guarantee that

$$\sum_{S \in \mathcal{S}^{\succ_i}} \bar{\mathbf{x}}(S) = \sum_{S \in \mathcal{S}^{\succ'_i}} \mathbf{x}(S)$$

for each agent $i \in N$ and, hence, the second set of constraints is satisfied since $\mathbf{x}$ satisfies the second set of constraints in the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ')$.

Furthermore, observe that $\mathrm{defc}(z^*, z, \succ) = \mathrm{defc}(z^*, z, \succ')$ for each alternative $z \in A \setminus \{y, z^*\}$ since the relative ranking of z and z^* in the preference of each agent is the same in both profiles $\succ$ and $\succ'$. Also, the definition of solution $\bar{\mathbf{x}}$ guarantees that

$$\sum_{S \in \mathcal{S}^{\succ_i}: z \in S} \bar{\mathbf{x}}(S) = \sum_{S \in \mathcal{S}^{\succ'_i}: z \in S} \mathbf{x}(S)$$

for each agent $i \in N$ and each alternatives $z \in A \setminus \{y, z^*\}$. Hence, the first set of constraints is satisfied for each alternative $z \in A \setminus \{y, z^*\}$ since the solution $\mathbf{x}$ satisfies the first set of constraints of the LP (6) in the definition of $\mathrm{pe}(z^*, y, \succ')$. Concerning alternative y, observe that $\mathrm{defc}(z^*, y, \succ) = \max\{0, \mathrm{defc}(z^*, y \succ') - |N_3|\}$. If $|N_3| \geq \mathrm{defc}(z^*, y, \succ')$ then, clearly,

$$\sum_{i \in N} \sum_{S \in \mathcal{S}^{\succ_i}: y \in S} \bar{\mathbf{x}}(S) \geq 0 = \mathrm{defc}(z^*, y, \succ).$$

Otherwise (if $|N_3| < \mathrm{defc}(z^*, y, \succ')$), observe that the definition of the solution $\bar{\mathbf{x}}$ guarantees that

$$\sum_{S \in \mathcal{S}^{\succ_i}: y \in S} \bar{\mathbf{x}}(S) \geq \sum_{S \in \mathcal{S}^{\succ'_i}: y \in S} \mathbf{x}(S)$$

for each agent $i \in N_1 \cup N_2$ (actually, the two sums are equal when $i \in N_1$), while no set in $\mathcal{S}^{\succ_i}$ contains alternative y when $i \in N_3$. We have

$$\begin{aligned}
& \sum_{i \in N} \sum_{S \in \mathcal{S}^{\succ_i}: y \in S} \bar{\mathbf{x}}(S) \\
= & \sum_{i \in N_1 \cup N_2} \sum_{S \in \mathcal{S}^{\succ_i}: y \in S} \bar{\mathbf{x}}(S) \\
\geq & \sum_{i \in N_1 \cup N_2} \sum_{S \in \mathcal{S}^{\succ'_i}: y \in S} \mathbf{x}(S) \\
\geq & \sum_{i \in N} \sum_{S \in \mathcal{S}^{\succ'_i}: y \in S} \mathbf{x}(S) - \sum_{i \in N_3} \sum_{S \in \mathcal{S}^{\succ'_i}: y \in S} \mathbf{x}(S) \\
\geq & \sum_{i \in N} \sum_{S \in \mathcal{S}^{\succ'_i}: y \in S} \mathbf{x}(S) - |N_3| \\
\geq & \ \mathrm{defc}(z^*, y, \succ') - |N_3| \\
= & \ \mathrm{defc}(z^*, y \succ),
\end{aligned}$$

which implies that the solution $\bar{\mathbf{x}}$ satisfies the first constraint of the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ)$ for alternative y. Note that the two last inequalities follows since the solution $\mathbf{x}$ satisfies the constraints in the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ')$.

So far, we have shown that $\bar{\mathbf{x}}$ is a feasible solution for the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ)$. It remains to upper-bound its objective value by the optimal objective value of the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ')$. In order to simplify the calculations, we denote by obj_1 the objective value of the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ)$ for the solutions $\bar{\mathbf{x}}$ and by obj_2 the objective value of the LP (6) used in the definition of $\mathrm{pe}(z^*, y, \succ')$ for the solutions $\mathbf{x}$. We also denote by $\mathrm{obj}_1(N')$ and $\mathrm{obj}_2(N')$ the two objective values when the sums in the objective function of the corresponding LP are restricted to agents of the set $N' \subseteq N$. In this way, it is $\mathrm{obj}_1 = \mathrm{obj}_1(N_1) + \mathrm{obj}_1(N_2) + \mathrm{obj}_1(N_3)$ and $\mathrm{obj}_2 = \mathrm{obj}_2(N_1) + \mathrm{obj}_2(N_2) + \mathrm{obj}(N_3)$ and the theorem will follow by proving that $\mathrm{obj}_1(N_1) \leq \mathrm{obj}_2(N_1)$, $\mathrm{obj}_1(N_2) \leq \mathrm{obj}_2(N_2)$, and $\mathrm{obj}_1(N_3) \leq \mathrm{obj}_2(N_3)$ (in the (in)equalities (7), (8), and (9) below).

Since $r_i(z^*, \succ) = r_i(z^*, \succ')$, $r_i(y, \succ) = r_i(y, \succ')$ when $y \succ_i z^*$, and $\bar{\mathbf{x}}(S_k^{\succ_i}) = \mathbf{x}(S_k^{\succ'_i})$ for $k = 1, ..., r_i(z^*, \succ) - 1$ for each agent $i \in N_1$, we have

$$\begin{aligned}
\mathrm{obj}_1(N_1) = & \sum_{i \in N_1} \sum_{k=1}^{r_i(z^*,\succ)-1} k \cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \\
& + \sum_{i \in N_1: y \succ_i z^*} \sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ)-1} \bar{\mathbf{x}}(S_k^{\succ_i}) \\
= & \sum_{i \in N_1} \sum_{k=1}^{r_i(z^*,\succ')-1} k \cdot \mathbf{x}(S_k^{\succ'_i}) \\
& + \sum_{i \in N_1: y \succ'_i z^*} \sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ')-1} \mathbf{x}(S_k^{\succ'_i}) \\
= & \ \mathrm{obj}_2(N_1). \qquad (7)
\end{aligned}$$

By the definition of $\mathrm{obj}_1(N_2)$ and by just rearranging the sums, we have

$$\begin{aligned}
\mathrm{obj}_1(N_2) &= \sum_{i\in N_2} \sum_{k=1}^{r_i(z^*,\succ)-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&\quad + \sum_{i\in N_2: y\succ_i z^*} \sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ)-1} \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&= \sum_{i\in N_2} \Bigg(\sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ)-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&\quad + (r_i(z^*,\succ)-r_i(y,\succ))\cdot \bar{\mathbf{x}}(S^{\succ_i}_{r_i(z^*,\succ)-r_i(y,\succ)}) \\
&\quad + \sum_{k=r_i(z^*,\succ)-r_i(y,\succ)+1}^{r_i(z^*,\succ)-r_i(y,\succ')-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&\quad + \big(r_i(z^*,\succ)-r_i(y,\succ')\big)\cdot \bar{\mathbf{x}}(S^{\succ_i}_{r_i(z^*,\succ)-r_i(y,\succ')}) \\
&\quad + \sum_{k=r_i(z^*,\succ)-r_i(y,\succ')+1}^{r_i(z^*,\succ)-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \Bigg) \\
&\quad + \sum_{i\in N_2: y\succ_i z^*} \sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ)-1} \bar{\mathbf{x}}(S_k^{\succ_i}).
\end{aligned}$$

By the definition of the variables $\bar{\mathbf{x}}$ on the sets of each agent $i\in N_2$ and since $r_i(z^*,\succ)=r_i(z^*,\succ')$ and since $y\succ_i z^*$ implies $y\succ'_i z^*$ for each agent $i\in N_2$, we have

$$\begin{aligned}
\mathrm{obj}_1(N_2) &= \sum_{i\in N_2} \Bigg(\sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ)-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) \\
&\quad + \big(r_i(z^*,\succ')-r_i(y,\succ)\big)\cdot 0 \\
&\quad + \sum_{k=r_i(z^*,\succ')-r_i(y,\succ)+1}^{r_i(z^*,\succ')-r_i(y,\succ')-1} k\cdot \mathbf{x}(S_{k-1}^{\succ'_i}) \\
&\quad + \big(r_i(z^*,\succ')-r_i(y,\succ')\big)\Big(\mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ')-r_i(y,\succ')-1}) \\
&\quad +\mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ')-r_i(y,\succ')})\Big) \\
&\quad + \sum_{k=r_i(z^*,\succ')-r_i(y,\succ')+1}^{r_i(z^*,\succ')-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) \Bigg) \\
&\quad + \sum_{i\in N_2: y\succ'_i z^*} \sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ)-1} \mathbf{x}(S_k^{\succ'_i}).
\end{aligned}$$

Again, by rearranging the sums and by the definition of $\mathrm{obj}_2(N_2)$, we obtain

$$\begin{aligned}
\mathrm{obj}_1(N_2) &= \sum_{i\in N_2} \Bigg(\sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ)-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) \\
&\quad + \sum_{k=r_i(z^*,\succ')-r_i(y,\succ)}^{r_i(z^*,\succ')-r_i(y,\succ')-1} (k+1)\cdot \mathbf{x}(S_k^{\succ'_i}) \\
&\quad + \sum_{k=r_i(z^*,\succ')-r_i(y,\succ')}^{r_i(z^*,\succ')-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) \Bigg) \\
&\quad + \sum_{i\in N_2: y\succ'_i z^*} \sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ)-1} \mathbf{x}(S_k^{\succ'_i}) \\
&= \sum_{i\in N_2} \sum_{k=1}^{r_i(z^*,\succ')-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) \\
&\quad + \sum_{i\in N_2: y\succ'_i z^*} \sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ')-1} \mathbf{x}(S_k^{\succ'_i}) \\
&= \mathrm{obj}_2(N_2). \qquad (8)
\end{aligned}$$

By the definition of $\mathrm{obj}_1(N_3)$ and since y is ranked below z^* in $\succ$ for $i\in N_3$ (i.e., $\{i\in N_3 : y\succ_i z^*\}=\emptyset$), we have

$$\begin{aligned}
\mathrm{obj}_1(N_3) &= \sum_{i\in N_3} \sum_{k=1}^{r_i(z^*,\succ)-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&\quad + \sum_{i\in N_3: y\succ_i z^*} \sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ)-1} \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&= \sum_{i\in N_3} \Bigg(\sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ')-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \\
&\quad + \big(r_i(z^*,\succ)-r_i(y,\succ')\big)\cdot \bar{\mathbf{x}}(S^{\succ_i}_{r_i(z^*,\succ)-r_i(y,\succ')}) \\
&\quad + \sum_{k=r_i(z^*,\succ)-r_i(y,\succ')+1}^{r_i(z^*,\succ)-1} k\cdot \bar{\mathbf{x}}(S_k^{\succ_i}) \Bigg) + 0.
\end{aligned}$$

By the definition of the variables $\bar{\mathbf{x}}$ on the sets of each agent $i\in N_3$, we have

$$\begin{aligned}
\mathrm{obj}_1(N_3) &= \sum_{i\in N_3} \Bigg(\sum_{k=1}^{r_i(z^*,\succ)-r_i(y,\succ')-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) \\
&\quad + \big(r_i(z^*,\succ)-r_i(y,\succ')\big)\Big(\mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ)-r_i(y,\succ')}) \\
&\quad +\mathbf{x}(S^{\succ'_i}_{r_i(z^*,\succ)-r_i(y,\succ')+1})\Big) \\
&\quad + \sum_{k=r_i(z^*,\succ)-r_i(y,\succ')+1}^{r_i(z^*,\succ)-1} (k+1)\cdot \mathbf{x}(S_k^{\succ'_i}) \Bigg).
\end{aligned}$$

Finally, by rearranging the sums, using the facts that the variables $\mathbf{x}$ are non-negative and $r_i(z^*,\succ)=r_i(z^*,\succ')-1$

0	...	$\lambda-3$	$\lambda-2$	...	$2\lambda-3$	$2\lambda-2$	...	$3\lambda-3$	$3\lambda-2$	...	$4\lambda-3$
a	...	a	b	...	b	$\Psi_{2\lambda-2}$	...	$\Psi_{3\lambda-3}$	a	...	a
Ψ_0	...	$\Psi_{\lambda-3}$	a	...	a	b	...	b	$\Psi_{3\lambda-2}$	...	$\Psi_{4\lambda-3}$
b	...	b	$\Psi_{\lambda-2}$	...	$\Psi_{2\lambda-3}$	a	...	a	b	...	b

0	1	2	...	$\lambda-1$	λ	...	$2\lambda-2$
a	a	a	...	a	b	...	b
b	b	b	...	b	Ψ_λ	...	$\Psi_{2\lambda-2}$
Ψ_0	Ψ_1	Ψ_2	...	$\Psi_{\lambda-1}$	a	...	a

Table 5: The preference profiles $\succ$ and $\succ'$ used in the proof of Theorem C.1. The subscripts of the sets Ψ_i are modulo k.

for $i \in N_3$, and by the definition of $\mathrm{obj}_2(N_3)$, we obtain

$$\begin{aligned}
\mathrm{obj}_1(N_3) &= \sum_{i\in N_3}\left(\sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ')-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) + \sum_{k=r_i(z^*,\succ')-r_i(y,\succ')}^{r_i(z^*,\succ')-1}(k-1)\cdot \mathbf{x}(S_k^{\succ'_i})\right)\\
&\le \sum_{i\in N_3}\sum_{k=1}^{r_i(z^*,\succ')-1} k\cdot \mathbf{x}(S_k^{\succ'_i}) + \sum_{i\in N_3 : y\succ'_i z^*}\sum_{k=1}^{r_i(z^*,\succ')-r_i(y,\succ')-1}\mathbf{x}(S_k^{\succ'_i})\\
&= \mathrm{obj}_2(N_3). \qquad (9)
\end{aligned}$$

This completes the proof of the lemma. □

C. PROOFS FOR SECTION 5

Each of the following subsections deals with a property (or two). Our main purpose is to prove Theorem 5.1, but in some cases we elaborate a bit regarding (simple) upper bounds.

C.1 Combinativity

An upper bound of $\mathcal{O}(nm)$ that satisfies combinativity can be obtained by selecting a Condorcet winner if one exists, and otherwise selecting some fixed alternative (by setting its score equal to $nm-1$ and setting the scores of all other alternatives equal to nm).

THEOREM C.1. *Let V be a Dodgson approximation. If V satisfies the combinativity property, then its approximation ratio is at least $\Omega(nm)$.*

PROOF. Let k, λ be positive integers such that k is even and divides $4\lambda-2$. Consider the following preference profile $\succ$ with $n = 4\lambda-2$ agents and $m = k+2$ alternatives. There is a set Ψ of k alternatives $\psi_0, ..., \psi_{k-1}$, and two additional alternatives a and b. For $i = 0, ..., k-1$, denote by Ψ_i the ordered set that contains the alternatives in Ψ ordered as $\psi_i, \psi_{i+1 \bmod k}, ..., \psi_{i+k-1 \bmod k}$.

The preference profile $\succ$ is the following (see Table 5):

- For $i = 0, ..., \lambda-3$, agent i ranks a first, then the alternatives of $\Psi_{i \bmod k}$, and then b.
- For $i = \lambda-2, ..., 2\lambda-3$, agent i ranks b first, then a, and then the alternatives of $\Psi_{i \bmod k}$.
- For $i = 2\lambda-2, ..., 3\lambda-3$, agent i ranks the alternatives of $\Psi_{i \bmod k}$ first, then b, and then a.
- For $i = 3\lambda-2, ..., 4\lambda-3$, agent i ranks a first, then the alternatives of $\Psi_{i \bmod k}$, and then b.

We have the following deficits with respect to $\succ$: $\mathrm{defc}(a, b, \succ) = 2$, $\mathrm{defc}(b, \psi_i, \succ) = \lambda$ for any $\psi_i \in \Psi$, and $\mathrm{defc}(\psi_i, a, \succ) = \lambda$ for any $\psi_i \in \Psi$. Alternative a has a Dodgson score of 2 since it suffices to push it one position upwards in the rankings of two agents in order to defeat b twice more. Alternative b has to defeat all the k alternatives in Ψ λ additional times, and hence has a Dodgson score at least $k\lambda$. Each alternative ψ_i of Ψ has to beat a λ times. By the definition of the sets Ψ_i, alternative ψ_i is ranked higher than alternative $\psi_{i+j \bmod k}$ by $(k-j)\frac{4\lambda-2}{k}$ agents, for $j = 1, 2, ..., k-1$ while it needs to defeat $\psi_{i+j \bmod k}$ 2λ times in total in order to beat it in their pairwise elections. Hence, for $j = k/2, ..., k-1$, we have

$$\begin{aligned}
\mathrm{defc}(\psi_i, \psi_{i+j \bmod k}, \succ) &= 2\lambda - (k-j)\frac{4\lambda-2}{k}\\
&= j\frac{4\lambda-2}{k} - 2\lambda + 2
\end{aligned}$$

and the Dodgson score of ψ_i is

$$\begin{aligned}
\mathrm{sc}_D(\psi_i, \succ) &\ge \mathrm{defc}(\psi_i, a, \succ) + \sum_{j=k/2}^{k-1}\mathrm{defc}(\psi_i, \psi_{i+j \bmod k}, \succ)\\
&= \lambda + \sum_{j=k/2}^{k-1}\left(j\frac{4\lambda-2}{k} - 2\lambda + 2\right)\\
&= \lambda + \frac{4\lambda-2}{k}\sum_{j=k/2}^{k-1} j - k(\lambda-1)\\
&= \frac{k\lambda}{2} + \frac{k}{4} + \frac{1}{2} > \frac{k\lambda}{2}.
\end{aligned}$$

Now consider the following preference profile $\succ'$ with $n' = 2\lambda-1$ agents and the same m alternatives (see Table 5).

- For $i = 0, ..., \lambda-1$, agent i ranks a first, then b, and then the alternatives of Ψ in arbitrary order.
- For $i = \lambda, ..., 2\lambda-2$, agent i ranks b first, then the alternatives of Ψ in arbitrary order, and then a.

In this preference profile, we have that alternative a is the Condorcet winner. If we combine the two preference profiles we get a new one $\succ''$ in which alternative b is the Condorcet winner. We can assume that V is Condorcet-consistent, since otherwise it would have an infinite approximation ratio. Hence, a should be the winner under V in $\succ'$

and b the winner under V in $\succ''$. Since V has the combinativity property, a is not the winner under V in $\succ$, i.e., its score $\mathrm{sc}_V(a,\succ)$ is either not smaller than $\mathrm{sc}_V(b,\succ)$ or not smaller than $\mathrm{sc}_V(\psi_i,\succ)$ for some alternative $\psi_i \in \Psi$. The minimum Dodgson score among these alternatives is at least $k\lambda/2$ while a has Dodgson score 2. Using the above, the fact that V is a Dodgson approximation, and the definition of n and m, we obtain that

$$\begin{aligned}\mathrm{sc}_V(a,\succ) &\geq \min_{y\in\Psi\cup\{b\}} \mathrm{sc}_V(y,\succ) \geq \min_{y\in\Psi\cup\{b\}} \mathrm{sc}_D(y,\succ) \geq \frac{k\lambda}{2} \\ &= \frac{k\lambda}{4}\mathrm{sc}_D(a,\succ) = \frac{(n+2)(m-2)}{16}\cdot \mathrm{sc}_D(a,\succ),\end{aligned}$$

i.e., V has an approximation ratio of $\Omega(nm)$. □

C.2 Smith Consistency

It is not difficult to define a trivial $\mathcal{O}(nm)$-approximation algorithm for the Dodgson score that satisfies Smith consistency. The algorithm selects a Condorcet winner if one exists. Otherwise, for each alternative in the Smith set, we set its score equal to $nm-1$, and we set the score of any other alternative equal to nm. Note that the Smith set can be computed in polynomial time [7]. The following statement shows that no asymptotic improvement is possible.

THEOREM C.2. *Let V be a Dodgson approximation. If V satisfies the Smith consistency property, then its approximation ratio is at least $\Omega(nm)$.*

PROOF. Let $k,t \geq 1$ be integers. We construct a preference profile $\succ$ with three alternatives a, b, and c that belong to the Smith Set, an alternative d, and a set of alternatives $X=\{x_1,...,x_{3t}\}$ that do not belong to the Smith Set, in a way that the Dodgson score of d is at most 3 and the Dodgson score of a, b, and c is at least $\Omega(nm)$.

The set X is partitioned into three sets $X_1=\{x_1,...,x_t\}$, $X_2=\{x_{t+1},...,x_{2t}\}$, and $X_3=\{x_{2t+1},...,x_{3t}\}$. We have $n=6k+1$ agents, $m=3t+4$ alternatives, and the preference profile $\succ$ of Table 6; note that the order of alternatives in the sets X_1, X_2, and X_3 is arbitrary.

$\times k$	$\times k$	$\times k$	$\times k$	$\times k$	$\times k$	$\times 1$
d	a	d	b	d	c	a
a	X_1	b	X_3	c	X_3	b
X_1	b	X_1	c	X_3	a	c
b	X_2	c	X_2	a	X_2	d
X_2	c	X_2	a	X_2	b	X_1
c	X_3	a	X_1	b	X_1	X_2
X_3	d	X_3	d	X_1	d	X_3

Table 6: The preference profile $\succ$ used in the proof of Theorem C.2.

Alternative a beats all alternatives besides c in a pairwise election and, furthermore, $\mathrm{defc}(a,c,\succ)=k$. Alternative b beats all alternatives besides a in a pairwise election and, furthermore, $\mathrm{defc}(b,a,\succ)=k+1$. Alternative c beats all alternatives besides b in a pairwise election and, furthermore, $\mathrm{defc}(c,b,\succ)=k+1$. Clearly the Smith set is $\{a,b,c\}$.

Alternative d beats all the alternatives in X and has $\mathrm{defc}(d,a,\succ)=\mathrm{defc}(d,b,\succ)=\mathrm{defc}(d,c,\succ)=1$. Clearly, the Dodgson score of d is (at most) 3 since it can be made a Condorcet winner by pushing it three positions in the ranking of the last agent.

Now, observe that, in order to defeat c in the preference of any agent where it is ranked lower, a has to be pushed at least $t+1$ positions upwards. This means that its Dodgson score is at least $k(t+1)$. Similarly, b (resp., c) can defeat a (resp., b) in the ranking of the last agent by rising one position upwards but needs at least $t+1$ pushes in the preference of any other agent where it is ranked below a (resp., b). So, we have that the Dodgson score of a is at least $k(t+1)$ and the Dodgson score of b and c is at least $1+k(t+1)$.

Since V has the Smith consistency property, some alternative among a, b, and c must be a winner. Using the bounds on the Dodgson scores and the definition of k and t, we have that

$$\begin{aligned}\mathrm{sc}_V(d,\succ) &\geq \min\{\mathrm{sc}_V(a,\succ),\mathrm{sc}_V(b,\succ),\mathrm{sc}_V(c,\succ)\} \\ &\geq \min\{\mathrm{sc}_D(a,\succ),\mathrm{sc}_D(b,\succ),\mathrm{sc}_D(c,\succ)\} \\ &\geq \frac{k(t+1)}{3}\mathrm{sc}_D(d,\succ) \\ &= \frac{(n-1)(m-1)}{54}\mathrm{sc}_D(d,\succ),\end{aligned}$$

i.e., V has an approximation ratio of $\Omega(nm)$. □

C.3 Mutual Majority Consistency and Invariant Loss Consistency

It is not difficult to see that the following trivial (superpolynomial-time) algorithm satisfies invariant loss consistency and has approximation ratio $\mathcal{O}(n)$ for the Dodgson score. For each alternative, set its score equal to nm if it loses to every other alternative and equal to its Dodgson score otherwise. The following statement shows that no asymptotic improvement is possible.

THEOREM C.3. *Let V be a Dodgson approximation. If V is mutual majority consistent or invariant loss consistent, then its approximation ratio is at least $\Omega(n)$.*

PROOF. Consider the following preference profile $\succ$ with $m\geq 4$ alternatives and $n=\lambda(m-1)$ agents where λ is odd and m is even (hence n is odd). There is a set Ψ of $m-1$ alternatives $\psi_0,...,\psi_{m-2}$ and one additional alternative a. For $i=0,...,m-2$, denote by Ψ_i the ordered set that contains the alternatives in Ψ ordered as $\psi_i,\psi_{i+1 \bmod (m-1)},...,\psi_{i+m-2 \bmod (m-1)}$.

The preference profile is the following:

- For $i=0,...,\lfloor\frac{\lambda(m-1)}{2}\rfloor-1$, agent i ranks a first, and then the alternatives of $\Psi_{i \bmod (m-1)}$.
- For $i=\lfloor\frac{\lambda(m-1)}{2}\rfloor,...,\lambda(m-1)-1$, agent i ranks the alternatives of $\Psi_{i \bmod (m-1)}$ first, and then a.

The Dodgson score of a is $sc_D(a,\succ)\leq m-1$ since it suffices to push a $m-1$ positions upwards in the ranking of one of the last $\lfloor n/2\rfloor$ agents in order to make it a Condorcet winner.

CLAIM C.4. *For every $\psi_i \in \Psi$, it holds that $sc_D(\psi_i,\succ) > \frac{\lambda(m-1)^2}{18}$.*

PROOF. Observe that ψ_i is ranked higher than $\psi_{i+j \bmod (m-1)}$ by $\lambda(m-1-j)$ agents. Hence, in order to beat $\psi_{i+j \bmod (m-1)}$ in the preferences of $\lceil n/2\rceil = \frac{\lambda(m-1)+1}{2}$

agents, the number of additional agents in whose preferences ψ_i must defeat $\psi_{i+j \bmod (m-1)}$ is

$$\begin{aligned}\mathrm{defc}(\psi_i, \psi_{i+j \bmod (m-1)}, \succ) &= \frac{\lambda(m-1)+1}{2} - \lambda(m-1-j) \\ &> \lambda\left(j - \frac{m}{2} + \frac{1}{2}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{sc}_D(\psi_i, \succ) &\geq \sum_{j=m/2}^{m-2} \mathrm{defc}(\psi_i, \psi_{i+j \bmod (m-1)}, \succ) \\ &> \lambda \sum_{j=m/2}^{m-2} \left(j - \frac{m}{2} + \frac{1}{2}\right) \\ &= \lambda \sum_{j=1}^{m/2-1} \left(j - \frac{1}{2}\right) \\ &= \frac{\lambda}{2}\left(\frac{m}{2} - 1\right)^2 \geq \frac{\lambda(m-1)^2}{18} \quad .\end{aligned}$$

The last inequality holds since $m \geq 4$. □

Clearly, every alternative in Ψ beats a in their pairwise election and, in particular, the alternatives in Ψ are ranked higher than a by more than half of the agents. Since V is mutual majority consistent or invariant loss consistent, a is not the winner under V in $\succ$, i.e., $\mathrm{sc}_V(a, \succ) \geq \min_{\psi_i \in \Psi} \mathrm{sc}_V(\psi_i, \succ)$. Using the above claim, we have that

$$\begin{aligned}\mathrm{sc}_V(a, \succ) &\geq \min_{\psi_i \in \Psi} \mathrm{sc}_V(\psi_i, \succ) \geq \min_{\psi_i \in \Psi} \mathrm{sc}_D(\psi_i, \succ) \\ &> \frac{\lambda(m-1)^2}{18} \geq \frac{\lambda(m-1)}{18}\mathrm{sc}_D(a, \succ) \\ &= \frac{n}{18}\mathrm{sc}_D(a, \succ)\end{aligned}$$

which means that V has an approximation ratio of $\Omega(n)$.

C.4 Independence of Clones

THEOREM C.5. *Let V be a Dodgson approximation. If V is independent of clones, then its approximation ratio is at least $\Omega(n)$.*

PROOF. Consider the preference profile $\succ$ with two alternatives a and b and n agents (n is a multiple of 4). The preferences are such that $\mathrm{defc}(a, b, \succ) = 0$ (i.e., a is a Condorcet winner) and $\mathrm{defc}(b, a, \succ) = 2$. Now, consider a Dodgson approximation V. If it selects b as a winner then, clearly, its approximation ratio is infinite. So, b should be a losing alternative.

Next, consider the profile $\succ'$ obtained by cloning alternative a four times so that the Dodgson score of the clones a_0, a_1, a_2, and a_3 of a is at least $n/4$. In order to do this, it suffices to replace a with the ranking of its clones $a_i \succ'_i a_{i+1 \bmod 4} \succ'_i a_{i+2 \bmod 4} \succ'_i a_{i+3 \bmod 4}$ in the ranking of agent i. The Dodgson score of b is at most 8 since it can become a Condorcet winner by pushing it to the top of the rankings of two agents. By the independence of clones property, b should be a losing alternative in this new profile, i.e., it should have $\mathrm{sc}_V(b, \succ') \geq \mathrm{sc}_V(a', \succ')$ for some clone a' of a. We obtain that

$$\begin{aligned}\mathrm{sc}_V(b, \succ') &\geq \mathrm{sc}_V(a', \succ') \geq \mathrm{sc}_D(a', \succ') \\ &\geq n/4 \geq n \cdot \mathrm{sc}_D(b, \succ')/32,\end{aligned}$$

which means that V has an approximation ratio of $\Omega(n)$. □