# Monotone Randomized Apportionment

JOSÉ CORREA, Universidad de Chile
PAUL GÖLZ, UC Berkeley, Cornell University
ULRIKE SCHMIDT-KRAEPELIN, TU Eindhoven
JAMIE TUCKER-FOLTZ, Harvard University
VICTOR VERDUGO, Pontificia Universidad Católica de Chile

Apportionment is the act of distributing the seats of a legislature among political parties (or states) in proportion to their vote shares (or populations). A famous impossibility by Balinski and Young (2001) shows that no apportionment method can be proportional up to one seat (*quota*) while also responding monotonically to changes in the votes (*population monotonicity*). Grimmett (2004) proposed to overcome this impossibility by randomizing the apportionment, which can achieve quota as well as perfect proportionality and monotonicity — at least in terms of the *expected number* of seats awarded to each party. Still, the correlations between the seats awarded to different parties may exhibit bizarre non-monotonicities. When parties or voters care about joint events, such as whether a coalition of parties reaches a majority, these non-monotonicities can cause paradoxes, including incentives for strategic voting.

In this paper, we propose monotonicity axioms ruling out these paradoxes, and study which of them can be satisfied jointly with Grimmett's axioms. Essentially, we require that, if a set of parties all receive more votes, the probability of those parties jointly receiving more seats should increase. Our work draws on a rich literature on *unequal probability sampling* in statistics (studied as *dependent randomized rounding* in computer science). Our main result shows that a sampling scheme due to Sampford (1967) satisfies Grimmett's axioms and a notion of higher-order correlation monotonicity.

## 1  INTRODUCTION

Across modern democracies, a long-lasting ideal has been that the legislature should "be an exact portrait, in miniature, of the people" it represents, in the words of John Adams [2]. Many democratic systems aim to achieve this maxim by partitioning the seats of their legislature over political blocs *in proportion* to the sizes of these blocs. Which blocs the countries consider differs: federal states divide up seats between the member states (in proportion to state populations), whereas 85 countries around the world divide up seats between political parties (in proportion to their vote shares) [1]. Mathematically speaking, however, both settings pose the same task of *apportionment*: dividing a fixed number $h \in \mathbb{N}$ of seats across $n$ blocs in proportion to the blocs' sizes. For the sake of exposition, we use the language of apportionment over parties throughout this paper, but our work equally applies to other apportionment settings.

Though it sounds almost trivial, the question of how to proportionally apportion seats possesses surprising mathematical and political complexity. The cause of this complexity is the indivisibility of seats — if a party receives 8.4% of the votes and hence deserves 8.4 out of $h = 100$ seats, should it receive 8 or 9, or even some other number? Since the 18th century, political luminaries and mathematicians have devised various *apportionment methods*, i.e., functions that take in the number of votes received by all parties and the legislature size $h$, and return how many seats are to be filled by each party. Which of these methods should be used has led to intense debate, which often revolved around the methods' mathematical properties and actively involved mathematicians [31].

Take, for example, *Hamilton's method*, the apportionment method put forward by the first US secretary of the treasury, Alexander Hamilton. His method first calculates, for each party, the number of seats it proportionally deserves (8.4, in the example above) and immediately assigns the floor of this number to the party (e.g., $\lfloor 8.4 \rfloor = 8$), which is known as the party's *lower quota*. Then, the method goes through the parties in decreasing order of their *residue* (e.g., $8.4 - \lfloor 8.4 \rfloor = 0.4$) and assigns one more seat to each party until the desired house size $h$ is reached. One strength of Hamilton's method is that it satisfies *quota*: a party's number of seats will never be below the floor of its proportional entitlement or above its ceiling. On the flip side, this method exhibits paradoxical non-monotonicities when the votes change. For instance, it violates *population monotonicity*, which means that, when party 1 gains voters and party 2 loses voters from one election to the next, party 1 might lose a seat while party 2 gains one. Such monotonicity violations are not mere mathematical inconveniences but led to Hamilton's method being abandoned for the apportionment over US states in 1901 [4, p. 42] and for the apportionment over parties in Germany in 2008 [16]. In fact, no apportionment method is immune to such criticism: as Balinski and Young showed, *any* method violates quota or population monotonicity [4].

In 2004, the mathematician Geoffrey Grimmett proposed a simple solution to this impasse: allowing apportionment methods to be random [19]. Following his proposal, a party deserving 8.4 seats would be "rounded down" to 8 seats with 60% probability and "rounded up" to 9 seats with 40% probability. In general, Grimmett's random apportionment of seats would always satisfy quota ex post and *ex-ante proportionality*, which means that a party's *expected number* of seats is perfectly equal to its proportional share. This latter property implies that the expected number of seats satisfies all reasonable forms of monotonicity, circumventing the Balinski–Young impossibility.

In designing such a randomized method, one must carefully correlate the rounding decisions across parties to ensure that the total number of seats is indeed $h$. Grimmett's method, like Hamilton's, first awards each party its lower quota, and then (randomly) awards some parties one more seat, as follows: Let the parties' residues be $p_1, \ldots, p_n \in [0, 1)$, and let $k$ denote their sum, which is the number of seats not yet awarded. We line up $n$ intervals on the number line without gaps, where the $i^{\text{th}}$ interval has length $p_i$. We then shift all intervals to the right by a random amount $u$, drawn

uniformly from $[0, 1)$, and round up exactly those parties whose interval contains an integer. Note that there are exactly $k$ such integers in the interval $[u, k + u)$, so this step will add $k$ seats, as required. Moreover, the process satisfies ex-ante proportionality since, no matter where some party $i$'s interval is placed, the random shift $u$ will leave it containing an integer with probability exactly $p_i$.

Beyond Grimmett's method, the recipe "award lower quotas then round" opens up a vast space of randomized apportionment methods, all of which satisfy quota and ex-ante proportionality. To construct such a method, all we need is a *rounding rule*, i.e., a procedure that takes in residues $p_1, \ldots, p_n \in [0, 1)$ summing to some integer $k$, and randomly selects a subset $S \subseteq \{1, \ldots, n\}$ of size $k$ such that each $1 \leq i \leq k$ is contained in $S$ with probability $p_i$. Fortuitously, this very setting has been studied in mathematical statistics as $\pi ps$ *("probability proportional to size") sampling without replacement* [7, 33]. In fact, the seminal book by Brewer and Hanif [7] lists 50 such methods,[1] out of which one, *systematic sampling* [23], yields exactly Grimmett's method when applied as above. Deville and Tille [13] and Hedayat et al. [21] even developed general frameworks for $\pi ps$ sampling, which capture infinitely many rounding rules and hence randomized apportionment methods. In the past two decades, theoretical computer scientists have searched for the same kind of rounding rules under the name of *dependent randomized rounding* [15], which have proved powerful tools in algorithm design [e.g., 9, 25, 30] — frequently reinventing rules previously known in statistics [34].

But which randomized apportionment method out of this multitude should one choose, given that all these methods satisfy quota and ex-ante proportionality?[2] Naturally, we define additional desirable axioms and focus on those methods with good axiomatic properties. In this paper, we ask

> *Which monotonicity axioms are desirable in randomized apportionment? And can sampling schemes from the statistics literature help us satisfy them?*

## 1.1 A Motivating Example

We motivate our monotonicity axioms with the story of two elections on the island of Apportia. Despite its modest population of 1100 people, Apportia boasts as many as six political parties, three leaning left and three leaning right. Apportia's parliament consists of just 11 legislators, elected at-large using Grimmett's randomized apportionment method. In the previous election, votes were cast according to the middle columns in Table 1.

Table 1. Vote totals for six parties in two fictional elections.

| Party | Ideology | Previous election | | | New election | | |
|---|---|---|---|---|---|---|---|
| | | Votes | Lower quota | Residue | Votes | Lower quota | Residue |
| 1 | left | 110 | 1 | 0.1 | 110 | 1 | 0.1 |
| 2 | right | 270 | 2 | 0.7 | (+20) 290 | 2 | 0.9 |
| 3 | left | 210 | 2 | 0.1 | 210 | 2 | 0.1 |
| 4 | right | 160 | 1 | 0.6 | (+30) 190 | 1 | 0.9 |
| 5 | left | 70 | 0 | 0.7 | (-60) 10 | 0 | 0.1 |
| 6 | right | 280 | 2 | 0.8 | (+10) 290 | 2 | 0.9 |

Did the left-wing parties have a chance of forming a majority coalition? The proportional number of votes needed to win a seat in Apportia is $1100/11 = 100$. Thus, after Grimmett's method had awarded each party its lower quota, the left-wing coalition collectively controlled three seats (one

---

[1]Though several only apply if $k = 2$ or give inclusion probabilities only approximately equal to the $p_i$.

[2]Grimmett himself offers "no justification for [his] scheme apart from fairness and ease of implementation" [19].
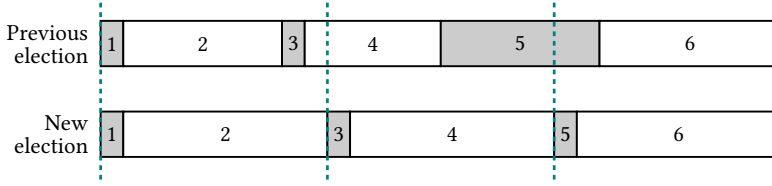
Fig. 1. Illustration of Grimmett's method [19] for the two elections from Table 1. Intervals corresponding to left-wing parties are gray. Dashed lines indicate the position of integers before the random shift.

from party 1, two from party 3, and none from party 5), the right-wing parties controlled five, with three seats remaining unapportioned. Hence, a left-wing majority would have only been possible if all remaining seats had gone to the left-wing parties. Upon inspection of the residues, however, one can see that this was impossible: Assuming we order the parties numerically for Grimmett's algorithm (see Fig. 1, top), the intervals for parties 1 and 3 are contained within the same unit interval, which means that it is impossible for both intervals to contain an integer, no matter the random shift $u$. Since a left-wing majority would have required each of the three parties to receive another seat, the right-wing parties obtained a majority with probability 1. (Arguably, this majority is deserved since the right-wing parties received 65% of the votes.)

Nonetheless, this poor showing by the left-wing parties was discouraging for the voters of party 5, which is why, for the next election, 60 of these voters switch to vote for right-wing parties: 20 voters switch to party 2, 30 to party 4, and 10 to party 6. (The right-most columns in Table 1 show the new tallies.) This shift leaves the lower quotas unchanged, but changes the residues. As can be seen on the bottom of Fig. 1, as long as the random shift $u$ falls in the range $(9/10, 1)$, all three left-wing parties will get an extra seat. This is a paradox: a shift of voters from left-wing to right-wing parties *increased* the chances of a left-wing government from zero to $1/10$. In other words, even though the vote shares of all left-wing parties weakly dropped and the vote shares of all right-wing parties grew, the left-wing coalition might well profit from (and hence encourage) this change. Our monotonicity axioms will rule out such paradoxes.[3]

## 1.2 Our Approach and Results

Our first contribution is the introduction of a family of monotonicity axioms that can guide the choice and future development of randomized apportionment methods. Unlike the work of Gölz et al. [17], who lift classical monotonicity axioms into the randomized realm through a rather technical notion of decomposition (see Section 1.3), our axioms are *native to randomized apportionment* and their violation directly implies paradoxes and perverse incentives. Conceptually, our axioms express that a set of parties that gains in votes while other parties lose votes should be entitled to higher chances of *simultaneous representation*, not just to higher chances of representation for each party separately, which is implied by ex-ante proportionality. Our axioms differ in which gains and losses of votes they apply to, how they define simultaneous representation, and whether they consider just a single, growing coalition or the comparison of a growing coalition with a shrinking one.

In Section 3, we begin our investigation in the domain of rounding rules. We find that not only the systematic rounding underlying Grimmett's method, but also pipage rounding and the

---

[3]It may appear as if this example relied on the ordering of the parties, so a natural attempt to fix Grimmett's algorithm would be to randomize the order. In fact, Grimmett [19] explicitly proposes randomizing the order as "it seems desirable to reduce to a minimum any correlations which depend on this extraneous element." However, the reader may verify that there was no order in which the left-wing parties could have reached a majority in the previous election. Thus, overcoming this paradox will require new apportionment methods.

rounding rule with maximum entropy (i.e., *conditional Poisson sampling*) violate simple forms of our monotonicity, and hence admit the paradox described in the previous subsection.

One rounding rule from the statistics literature, however, bucks this trend: we show that *Sampford sampling* [29] satisfies the axiom of *selection monotonicity*.

THEOREM 1.1 (SAMPFORD ROUNDING SATISFIES SELECTION MONOTONICITY). *Let* $\vec{p}, \vec{p}' \in [0, 1)^n$ *be two residue vectors, summing up to the same integer* $k$. *Let* $T$ *be a set of* $k$ *parties such that* $p'_i \geq p_i$ *for all parties* $i \in T$ *and* $p'_i \leq p_i$ *for all* $i \notin T$. *Then, the probability that Sampford sampling rounds up the set* $T$ *is at least as high for* $\vec{p}'$ *as for* $\vec{p}$.

Proving this theorem reduces to establishing a clean inequality between the partial derivatives of Sampford's probability mass function with respect to changes in the residues. But establishing this inequality is a considerable technical challenge because the Sampford's probability-mass expression contains a denominator that sums over all sets of $k$ parties. The key to our main result is establishing a surprising equality between this denominator and an expected-value term related to a Poisson trial, which allows us to bound and take derivatives of the denominator. Sampford sampling is also the only rounding rule among those we consider that satisfies a two-coalition version of selection monotonicity, which is logically incomparable to the axiom above. Because Sampford sampling possesses the same properties that computer scientists have leveraged when applying dependent randomized rounding (in particular, strong concentration properties like *conditional negative association* [6]), replacing pipage rounding with Sampford sampling adds our monotonicity axiom "for free", which we believe has potential applications, for example in mechanism design. Sampford sampling is the first, and only, rounding rule known to satisfy selection monotonicity, and there are no other obvious contenders. As for general results about this axiom, we are able to show that it implies that the rounding rule be Lipschitz continuous.

In Section 4, we return to the general apportionment setting and consider even stronger axioms of monotonicity, which refer to the probability of coalitions exceeding arbitrary seat thresholds. We show that, if the axioms are defined with respect to shifts in the raw vote counts (rather than the parties' share of votes), such axioms cannot be satisfied by any apportionment method. Even when shifts are measured in terms of vote share, the two-coalition axiom is incompatible with a natural regularity condition on randomized apportionment methods, namely, having *full support*. Whether the most direct generalization of our axiom, *threshold monotonicity*, can be satisfied, is left as an open question, though we conjecture that it is indeed satisfied by Sampford sampling. For the restricted case where coalitions consist of only two parties, we show that threshold monotonicity is satisfied by Grimmett's method.

## 1.3 Related Work

Deterministic apportionment has been the subject of a deep axiomatic treatment, much of it developed by Balinski and Young [see 4, 28]. Two of the most prominent axioms in this literature, like the ones we develop here, are forms of monotonicity that rule out paradoxes [4, Ch. 5]: population monotonicity, which we explained already, and *house monotonicity*, which states that increasing the house size may not cause any party to lose seats. Under mild assumptions, population monotonicity is incompatible with quota [4, Thm. 6.1] and is satisfied exactly by the widely used family of divisor methods [4, Thm. 4.2]. Coalitions have mainly been considered in this literature in the context of whether apportionment methods encourage multiple parties to merge into a single party [3]; in randomized apportionment, merging leaves the coalition's expected number of seats unaffected. In addition, population monotonicity can be restated in terms of coalitions, in a way that resembles our pairwise axioms: if all parties in coalition $T_1$ weakly gain in votes and all parties

in $T_2$ weakly lose votes, then each party in $T_1$ must weakly gain seats or each one in $T_2$ must weakly lose seats.

We have already touched on the literature on $\pi$ps sampling without replacement above. A central motivation for this literature is the *Horvitz–Thompson (HT) estimator* [22], which we illustrate in a toy example (loosely following [29]): Suppose we have farms numbered $1, \ldots, n$, and each farm $i$ has a known size of land $x_i$, and an unknown yield $y_i$. Following Horvitz and Thompson [22], we can get an unbiased estimate of the total yield, $\sum_{1 \leq i \leq n} y_i$, by sampling $k$ farms with probability proportional to their size $x_i$ and without replacement; summing up the yield per acre $y_i/x_i$ for each sampled farm $i$, and multiplying the result by $\sum_{1 \leq i \leq n} x_i/k$.[4] The variance of this estimate, which depends on how homogeneous the yield per acre is across farms, can be estimated from the same sample [36] as long as the joint sampling probabilities of pairs of farms are known. As a result, the $\pi$ps literature scrupulously studied pairwise inclusion probabilities [e.g., 7], but paid little attention to higher-order correlations, which gives our work a distinct lens on the topic.

By contrast, higher-order correlations have been very central to the study of dependent randomized rounding in computer science. As we explain in Section 3.4, typical applications of dependent rounding lean heavily on concentration bounds. In a pursuit of stronger concentration results, researchers on dependent rounding have leveraged increasingly strong notions of negative correlation, ranging from the relatively weak *negative cylinder dependence* [26] over *negative association* [35] to the *strong Rayleigh property* [5]. The axioms we pursue fundamentally differ from those above in that they are *relational* rather than *punctual* [32], i.e., in that they constrain how the rounding rule's distributions relates across different residue vectors, rather than just imposing separate constraints for each input. Whereas we only consider dependent rounding subject to a cardinality constraint [15, 30], dependent rounding schemes have since been developed for more general combinatorial constraints, including bipartite matchings [15] and matroids [10, 27].

A source of inspiration for our paper is the work by Gölz et al. [17], which, like us, aims to refine the space of randomized apportionment methods by imposing monotonicity axioms in addition to quota and ex-ante proportionality. The monotonicity axioms they arrive at, however, have a different flavor from ours: They say that a randomized apportionment method satisfies population monotonicity (respectively, house monotonicity) if it can be represented as a probability distribution over deterministic apportionment methods satisfying population monotonicity (respectively, house monotonicity). These axioms can be interpreted as preventing paradoxes, but only in artificial scenarios in which the votes or house size change *after the randomization of the apportionment has been performed*. The authors show that population monotonicity is incompatible with quota, but that house monotonicity can be guaranteed alongside quota and ex-ante proportionality. In Appendix A, we show that house monotonicity is unlikely to be compatible with our axioms: no apportionment method satisfying full support, and none of the methods we consider in this paper, satisfy house monotonicity.

## 2 PRELIMINARIES

Throughout this paper, let $[n] := \{1, \ldots, n\}$ denote a set of *parties*. For a set $S$ and natural number $k$, we set $\binom{S}{k} := \{T \subseteq S \mid |T| = k\}$. An *apportionment method* $a$ is a function that takes in the number of *votes* $v_1, \ldots, v_n \in \mathbb{R}_{\geq 0}$ for each party and house size $h$, and returns a vector-valued random variable $a(\vec{v}, h) \in \mathbb{N}^n$, summing to $h$, which lists the number of seats awarded to each party. For a fixed vector of votes $\vec{v}$ and a house size $h$, we denote party $i$'s *standard quota* (i.e., its proportionally

---

[4] A simpler estimator would sample $k$ farms uniformly at random, add up their yields $y_i$, and multiply the sum by $n/k$. This estimator is also unbiased but can have much higher variance because it often misses large farms with high yield.

deserved share of seats) by $q_i := h\,v_i/\sum_{j\in[n]} v_j$ and its *residue* by $p_i := q_i - \lfloor q_i \rfloor$. We will only consider apportionment methods that satisfy the following two axioms of Grimmett [19]:

**Quota.** For any $\vec{v}, h$, the number of seats awarded to party $i$ is always $\lfloor q_i \rfloor$ or $\lceil q_i \rceil$ ex post.

**Ex-ante proportionality.** For any $\vec{v}, h$, the expected number of seats awarded to party $i$ equals $q_i$.

Similarly, a *rounding rule* $r$ is a function that takes in a vector of *target probabilities* $\vec{p} \in [0, 1)^n$ for each party, which must sum up to an integer $k$, and returns a set-valued random variable $r(\vec{p}) \in \binom{[n]}{k}$, i.e., the values taken on by the random variable are sets of $k$ parties. In addition, $r$ must satisfy for each $\vec{p}$ and $i \in [n]$ that $i$'s probability of being included in $r(\vec{p})$ equals its target probability $p_i$. We are overloading $p_i$ by using it both for general target probabilities and residues induced by a vote profile $\vec{v}$. We do so as they are mathematically equivalent objects and will play the same role in Section 3 on rounding rules and Section 4 on apportionment rules, respectively. We say that a rounding rule is *continuous* if it is a continuous as a function from its domain, thought of as a subset of $\mathbb{R}^n$, to $\mathbb{R}^{\binom{n}{k}}$.

We can easily translate apportionment methods into rounding rules and vice versa. On the one hand, we can interpret any (quota and ex-ante proportional) apportionment method $a$ as a rounding rule: for given target probabilities $\vec{p}$ summing to $k \in \mathbb{N}$, $a(\vec{p}, k)$ is a random variable ranging over integer vectors adding to $k$. Quota ensures that this vector is binary and can thus be interpreted as a random set of $k$ parties. Ex-ante proportionality ensures that the target probabilities are met. Conversely, any rounding rule $r$ *induces an apportionment method*, as we have described in the introduction: Given the votes $\vec{v}$ and house size $h$, one computes the standard quotas $q_i$ and residues $p_i$, awards each party its lower quota $\lfloor q_i \rfloor$, at which point some number $k = \sum_{i\in[n]} p_i$ of seats remain unallocated. By randomly awarding an extra seat to the $k$ parties in $r(\vec{p})$, one obtains an apportionment method satisfying quota and ex-ante proportionality.

We define axioms for rounding rules and apportionment methods directly in Sections 3 and 4.

## 3 MONOTONICITY OF ROUNDING RULES

We begin our investigation of montonicity in the subsetting of rounding rules, for two reasons: First, the rounding setting allows us to capture the paradox in our motivating example (Section 1.1) in a more direct, circumscribed way: it is paradoxical that an exodus of voters from a bloc of parties could increase the probability that *exactly these parties receive an extra seat*. The rounding setting allows us to analyze this paradox — and bring it to an end — before considering in Section 4 more expansive definitions that extend to situations in which the lower quotas, and thus the meaning of getting an "extra seat," change.

The second reason for starting in the rounding setting is the importance of this setting beyond apportionment. In particular, our positive results in this section immediately contribute to the understanding of $\pi$ps sampling, and have potential implications for dependent randomized rounding in computer science, as we sketch in Section 3.4.

### 3.1 Axioms and Rounding Rules

We rule out the paradox from our introduction with the axiom of *selection monotonicity*, which says that, if the target probabilities of all parties in some set $T$ go up, and the target probabilities of all other parties go down, it should become (weakly) more likely that the set $T$ is rounded up:

*Definition 3.1 (Selection monotonicity).* Let $\vec{p}, \vec{p}' \in [0, 1)^n$ be two residue vectors, summing up to the same integer $k$. Let $T$ be a set of $k$ parties such that $p_i' \geq p_i$ for all parties $i \in T$ and $p_i' \leq p_i$ for all $i \notin T$. A rounding rule $r$ satisfies selection monotonicity if it always holds that

$$\mathbb{P}_{S\sim r(\vec{p}')}[S = T] \geq \mathbb{P}_{S\sim r(\vec{p})}[S = T].$$

As a warm-up to our proof, we will also consider a variant which, like the axiom of population monotonicity in classical apportionment theory, implies a disjunction: if a coalition $T_1$ is growing and a coalition $T_2$ is shrinking, $T_1$'s probability of being jointly rounded up should increase *or* $T_2$'s probability of being jointly rounded up should decrease:

*Definition 3.2 (Pairwise selection monotonicity).* Let $\vec{p}, \vec{p}' \in [0, 1)^n$ be two residue vectors, summing up to the same integer $k$. Let $T_1, T_2$ be two sets of $k$ parties (which might overlap) such that $p'_i \geq p_i$ for all parties $i \in T_1$ and $p'_i \leq p_i$ for all parties $i \in T_2$. A rounding rule $r$ satisfies pairwise selection monotonicity if it always holds that

$$\mathbb{P}_{S \sim r(\vec{p}')}[S = T_1] \geq \mathbb{P}_{S \sim r(\vec{p})}[S = T_1] \quad \text{or} \quad \mathbb{P}_{S \sim r(\vec{p}')}[S = T_2] \leq \mathbb{P}_{S \sim r(\vec{p})}[S = T_2].$$

Note that pairwise selection monotonicity is not implied by selection monotonicity, since the target probabilities of parties outside of $T_1$ and $T_2$ are allowed to vary arbitrarily.

We now introduce several rounding rules from the literature on $\pi$ps sampling:

**Systematic rounding [23].** We have already introduced this rule in the introduction, in the context of Grimmett's randomized apportionment method [19]. A random shift $u \sim Uniform([0, 1))$ is drawn and party $i$ is selected if the interval $[u + \sum_{1 \leq j < i} p_j, u + \sum_{1 \leq j \leq i} p_j)$ contains an integer.

**Pipage rounding [13, 30].**[5] While at least two parties $i, j$ with target probabilities $0 < p_i, p_j < 1$ exist, fix the first two such parties. If $p_i + p_j \leq 1$, update $p_i \leftarrow p_i + p_j, p_j \leftarrow 0$ with probability $p_i/(p_i + p_j)$, and update $p_i \leftarrow 0, p_j \leftarrow p_i + p_j$ with the remaining probability. Else, update $p_i \leftarrow 1$, $p_j \leftarrow p_i + p_j - 1$ with probability $(1 - p_j)/(2 - p_i - p_j)$, and update $p_i \leftarrow p_i + p_j - 1, p_j \leftarrow 1$ with the remaining probability. Once no such $p_i, p_j$ exist, select the $k$ parties with target probability 1.

**Conditional Poisson rounding [11].** This rounding rule implements the probability distribution over $\binom{[n]}{k}$ with maximal entropy, subject to attaining the target probabilities. To implement this rule algorithmically, one must determine the unique *working probabilities* $\pi_1, \ldots, \pi_n \in \mathbb{R}_{\geq 0}$ such that choosing each set $T \in \binom{[n]}{k}$ with probability $\prod_{i \in T} \pi_i$ satisfies the target probabilities, which can be achieved using Newton's method [33].

**Sampford rounding [29].** This rounding rule can be implemented with the following rejective procedure: Randomly sample a party such that each party $i$ is selected with probability proportional to $p_i$; call the sampled party $i_1$. Then, randomly sample $k - 1$ parties $i_2, \ldots, i_k$, with replacement from $[n]$, choosing each party $i$ with probability proportional to $\frac{p_i}{1 - p_i}$. If the $k$ drawn parties are distinct, select them; else, start over. Sampford's rounding rule can also be implemented in a non-rejective, polynomial-time manner [18].

It is worth dwelling on how *strange* (or, in the words of Tillé [33], "very ingenious") Sampford's rounding rule is. If, say, all $k$ units were drawn with probability proportional to $\frac{p_i}{1 - p_i}$, each set $A$ of $k$ parties would be drawn with probability proportional to $\prod_{i \in A} \frac{p_i}{1 - p_i}$, which means that this would just be Conditional Poisson rounding, but for incorrect working probabilities $\pi_i = \frac{p_i}{1 - p_i}$. But, whereas the computation of the correct working probabilities for conditional Poisson rounding has no closed form, adjusting the selection probabilities for only one of the $k$ draws perfectly attains the desired inclusion probabilities (which is nontrivial to show).

Though it is not a rounding rule in our model, our analysis will frequently refer to *Poisson sampling* (not to be confused with *conditional* Poisson sampling). Poisson sampling is not a rounding rule because it samples sets of nondeterministic size: it simply performs, for each party $i$, an independent Bernoulli trial with success probability $p_i$, and contains $i$ in its output if this trial is a success. Clearly, the probability of sampling a set $A$ is just $\prod_{i \in A} p_i \prod_{i \notin A} (1 - p_i)$.

---

[5]This method was first defined in statistics as the *pivotal method* [13], but we use the term common in computer science.

## 3.2 Rules that Violate Selection Monotonicity

The axiom of selection monotonicity might look so innocuous that one would take its satisfaction for granted. For example, in Poisson sampling, an increase in the target probabilities of $T$ and a decrease in the other parties' target probabilities would obviously increase $\prod_{i \in T} p_i \prod_{i \notin T} (1 - p_i)$, the probability of sampling $T$. Alas, Poisson sampling need not select the right number of parties. Selection monotonicity would also be straightforward to show if we used the naïve working probabilities $\pi_i = p_i / (1 - p_i)$ in conditional Poisson rounding. However, even though the selection probabilities asymptotically converge to the targets [20, Thm. 7.3], they are not exactly met.[6]

Even though these "almost rounding rules" easily satisfy selection monotonicity, most proper rounding rules fail the axiom. Indeed, our example from the introduction already proved that systematic rounding (even with a randomized ordering of parties) fails selection monotonicity, and we give a similar counterexample for pipage rounding in Appendix B.

PROPOSITION 3.3. *Systematic rounding and pipage rounding violate selection monotonicity, even if the order of parties is uniformly shuffled.*

But, perhaps, this violation is to be expected in light of several characteristics common to systematic and pipage rounding. Both rounding rules depend a lot on the ordering of parties, can have strange correlations between the inclusion of pairs of parties, and tend to have probability distributions supported on only few $k$-subsets of parties. Sharing none of these properties, and being the "most natural" rounding rule given its maximum entropy characterization, conditional Poisson rounding would seem like a promising candidate.

Still, to our own surprise, conditional Poisson rounding also violates selection monotonicity. How we found the counterexample attests to how close conditional Poisson rounding gets to satisfying the axiom: For the special case of $n = 6$, $k = 3$, we derived a polynomial over $p_1, \ldots, p_6$ which is negative exactly for the vectors of target probabilities that would certify a violation of selection monotonicity. Out of 705 monomials, only a single one had a negative sign, which by setting the target probabilities just right, at different orders of magnitude, could be made to overpower all other monomials and render the term as a whole negative. We refer to Appendix B for the proof.

PROPOSITION 3.4. *Conditional Poisson rounding violates selection monotonicity.*

A final reason to temper optimism about satisfying selection monotonicity all too easily is that a strengthened axiom, in which the target probabilities outside $T$ may increase or decrease, cannot be satisfied by any rounding rule. To see this, consider a scenario with four parties, which initially all have residues $1/2$ (thus, $k = 2$ parties will be selected). By symmetry, we may assume that the pair $T := \{1, 2\}$ is selected with probability at least $1 / \binom{4}{2} = 1/6$. Now let party 3's target increase to 99% and party 4's target decrease to 1%. Since party 3 must now almost always be one of the two selected parties, the pair $\{1, 2\}$ must be jointly selected much more rarely than $1/6$, even though their target probabilities weakly increased. This example shows that the distribution of target probabilities outside of $T$ can necessitate major changes to the joint selection probability of $T$.

## 3.3 Sampford Rounding is Monotone

We now more closely consider Sampford rounding, the rounding rule that finally satisfies our selection monotonicity axioms. Let $S := r_{\text{Sampford}}(\vec{p})$ be a random variable describing the set of parties drawn by Sampford rounding. From the rejective sampling procedure described in Section 3.1,

---

[6]These two processes are closely related: the Poisson sample, conditioned on selecting exactly $k$ parties, has the same distribution as this naïve conditional Poisson rounding — hence the name.

it follows that the probability of a set $A$ being selected is

$$\mathbb{P}[S = A] = \frac{\sum_{i \in A} p_i \prod_{j \in A \setminus i} \frac{p_j}{1-p_j}}{\sum_{A' \in \binom{[n]}{k}} \sum_{i \in A'} p_i \prod_{j \in A' \setminus i} \frac{p_j}{1-p_j}}. \tag{1}$$

As an easy warm-up, and an indication that Sampford rounding is indeed well suited for monotonicity, we first show that it satisfies the pairwise variant of selection monotonicity:

THEOREM 3.5. *Sampford rounding satisfies pairwise selection monotonicity.*

PROOF. Fix the target probabilities $\vec{p}, \vec{p}'$, the sample size $k$, and the sets $T_1, T_2$. Considering the numerator of the probability mass term in Eq. (1), it is easy to see that the numerator of $\mathbb{P}[S = T_1]$ weakly increases when going from $\vec{p}$ to $\vec{p}'$, whereas the numerator of $\mathbb{P}[S = T_2]$ weakly decreases.

Now, we distinguish two cases, depending on whether the denominator, which is common to $T_1$ and $T_2$, increases from $\vec{p}$ to $\vec{p}'$ or not. If it increases, then, since the numerator for $T_2$ decreases, the probability of sampling $T_2$ must decrease, and we are done. Else, since the numerator for $T_1$ increases, the probability of sampling $T_1$ increases, concluding the proof. □

Whereas Sampford rounding satisfies pairwise selection monotonicity quite handily, we show in Appendix B that the three other rounding rules all violate this axiom.

What made this proof so easy was that we barely had to engage with the denominator term, a sum over exponentially many terms that indicates the probability of non-rejection for the sampling procedure. Engaging with how the probability of non-rejection changes with the target probabilities will be the major obstacle in proving our main result, that Sampford satisfies selection monotonicity. For this, we slightly reformulate the probability-mass term, and continue Eq. (1):

$$\begin{aligned} \mathbb{P}[S = A] &= \frac{\sum_{i \in A}(1-p_i) \prod_{j \in A} \frac{p_j}{1-p_j}}{\sum_{A' \in \binom{[n]}{k}} \sum_{i \in A'}(1-p_i) \prod_{j \in A'} \frac{p_j}{1-p_j}} \\ &= \frac{\sum_{i \in A}(1-p_i) \prod_{j \in A} p_j \prod_{j \notin A}(1-p_j)}{\sum_{A' \in \binom{[n]}{k}} \sum_{i \in A'}(1-p_i) \prod_{j \in A'} p_j \prod_{j \notin A'}(1-p_j)} = \frac{f(A)}{\sum_{A' \in \binom{[n]}{k}} f(A')}, \end{aligned} \tag{2}$$

where the second equality expands the fraction with $\prod_{i \in [n]}(1-p_i)$, and where we introduce the shorthand $f(A)$ for the term $\sum_{i \in A}(1-p_i) \prod_{j \in A} p_j \prod_{j \notin A}(1-p_j)$, which will appear extensively throughout this proof.

In the remainder of this section, we will prove Theorem 1.1, i.e., that Sampford rounding is selection monotone. By symmetry, we will assume that $A = [k]$. A simple argument (Appendix C) shows that it suffices to consider the special case where the target probability of one party in $A$ (w.l.o.g., party 1) infinitesimally grows and the target probability of one party outside of $A$ (w.l.o.g., party $n$) shrinks by the same infinitesimal amount. Since Sampford's probability mass function is differentiable, this reduces to the following inequality of its partial derivatives:

PROPOSITION 3.6. *Let $\vec{p} \in [0, 1)^n$ such that $\sum_{i \in [n]} p_i = k$ for some integer $1 \le k \le n - 1$, $p_1 > 0$, and $S = r_{Sampford}(\vec{p})$. Then,*

$$\left( \frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{P}[S = [k]] \ge 0.$$

A crucial quantity throughout our proof will be $s := \sum_{i \in [k]}(1 - p_i)$, which measures how far the residues in $[k]$ fall short of being 1. Alternatively, $s = \sum_{i \in [k]}(1 - p_i) = k - \sum_{i \in [k]} p_i = \sum_{i \notin [k]} p_i$ can be seen as how far the residues outside of $[k]$ exceed 0.

Clearly, to bound the partial derivative from Proposition 3.6, we will eventually need to apply the quotient rule, which requires us to understand and bound the numerator and denominator

in Eq. (2), as well as their derivatives. While the numerator's expression is fairly simple, the core challenge lies in bounding the denominator (and its derivative), which sums up over exponentially many sets $A'$ of parties.

We will overcome this hurdle in Lemma 3.8, which establishes a surprising equality between this denominator term and the expected value of a natural variable defined over a Poisson trial. Then, we dedicate Lemmas 3.9 to 3.11 to finding a concise expression of the derivative of the denominator and clean bounds in terms of our parameter $s$ for both of these terms. This established the foundation for proving Proposition 3.6.

For the remainder of the section (including all lemma statements) we fix $\vec{p} \in [0, 1)^n$ as defined in Proposition 3.6. To prove our key Lemma 3.8, we require the following auxiliary lemma:

LEMMA 3.7. *For all* $0 \leq \ell \leq n - 1$ *it holds that*

$$\sum_{A \in \binom{[n]}{\ell+1}} f(A) - \sum_{A \in \binom{[n]}{\ell}} f(A) = (k - \ell) \sum_{A \in \binom{[n]}{\ell}} \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j).$$

PROOF. We start by reformulating the summation as follows:

$$\sum_{A \in \binom{[n]}{\ell+1}} f(A) = \sum_{A \in \binom{[n]}{\ell+1}} \sum_{i \in A} (1 - p_i) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j)$$

$$= \sum_{A \in \binom{[n]}{\ell+1}} \sum_{i \in A} p_i \prod_{j \in A \setminus \{i\}} p_j \prod_{j \notin A \setminus \{i\}} (1 - p_j).$$

In the next step, we will relabel these sums over $A$ and $i$. Specifically, we apply a bijection from the set $\{(A, i) \mid A \in \binom{[n]}{\ell+1}, i \in A\}$ to the set $\{(A', i') \mid A' \in \binom{[n]}{\ell}, i' \notin A'\}$, which maps $(A, i)$ to $(A \setminus \{i\}, i)$. To see that this is a bijection, observe that it has the inverse $(A', i') \mapsto (A' \cup i', i')$. Then, the previous summation is equal to

$$= \sum_{A \in \binom{[n]}{\ell}} \sum_{i \notin A} p_i \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j),$$

where we omit primes for the sake of readability. Finally, we make use that $\sum_{i \in [n]} p_i = k$ to get

$$= \sum_{A \in \binom{[n]}{\ell}} \left( k - \ell + \ell - \sum_{i \in A} p_i \right) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j)$$

$$= \sum_{A \in \binom{[n]}{\ell}} (k - \ell) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) + \sum_{A \in \binom{[n]}{\ell}} \left( \sum_{i \in A} 1 - p_i \right) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j)$$

$$= (k - \ell) \sum_{A \in \binom{[n]}{\ell}} \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) + \sum_{A \in \binom{[n]}{\ell}} f(A),$$

which concludes the proof. □

Lemma 3.7 provides us with the necessary tool to prove the reformulation of the denominator. In the new expression, the right-hand side is reminiscent of the variance of the random variable expressing the size of a set sampled via a Poisson trial (a *Poisson binomial variable*). However, the expectation is taken over the absolute distance, rather than the squared distance, towards the mean.

LEMMA 3.8. *Let $B \subseteq [n]$ be a random variable distributed according to a Poisson trial, i.e., $B$ contains each $1 \le i \le n$ independently with probability $p_i$. Then,*

$$\sum_{A \in \binom{[n]}{k}} f(A) = \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] = \tfrac{1}{2} \cdot \mathbb{E}\left[\big||B| - k\big|\right].$$

PROOF. For the first equality, we simply write its left-hand side as a telescoping sum:

$$\sum_{A \in \binom{[n]}{k}} f(A) = \sum_{A \in \binom{[n]}{k}} f(A) - \sum_{A \in \binom{[n]}{0}} f(A) \qquad \text{(second term is 0)}$$

$$= \sum_{0 \le \ell \le k-1} \left(\textstyle\sum_{A \in \binom{[n]}{\ell+1}} f(A) - \sum_{A \in \binom{[n]}{\ell}} f(A)\right)$$

$$= \sum_{0 \le \ell \le k-1} (k - \ell) \cdot \sum_{A \in \binom{[n]}{\ell}} \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \qquad \text{(by Lemma 3.7)}$$

$$= \sum_{0 \le \ell \le k-1} (k - \ell) \cdot \mathbb{P}\big[|B| = \ell\big] = \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right].$$

To show the statement's second equality, recall that the first central moment of any random variable is zero. Applying this to the variable $|B|$ together with the fact that $\mathbb{E}\big[|B|\big] = \sum_{i \in [n]} p_i = k$ yields that $\mathbb{E}\big[|B| - k\big] = 0$. Since $\mathbb{E}\big[|B| - k\big] = \mathbb{E}\big[\mathbb{1}\{|B| - k \ge 0\} \cdot (|B| - k)\big] + \mathbb{E}\big[\mathbb{1}\{|B| - k < 0\} \cdot (|B| - k)\big]$, this implies that

$$\mathbb{E}\big[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\big] = -\mathbb{E}\big[\mathbb{1}\{|B| - k < 0\} \cdot (|B| - k)\big]$$
$$= \mathbb{E}\big[\mathbb{1}\{|B| - k \ge 0\} \cdot (|B| - k)\big].$$

The desired equality follows by observing that

$$\mathbb{E}\big[\big||B| - k\big|\big] = \mathbb{E}\big[\mathbb{1}\{|B| - k \ge 0\} \cdot (|B| - k)\big] + \mathbb{E}\big[\mathbb{1}\{|B| - k < 0\} \cdot (k - |B|)\big]$$
$$= 2\,\mathbb{E}\big[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\big],$$

which concludes the proof.                                                                                       □

Below, we provide a concise upper bound for the denominator in terms of our parameter $s$, building upon the close connection to the variance of the random variable $|B|$ pointed out above.

LEMMA 3.9. *Let $B$ be as in Lemma 3.8. Then, $\mathbb{E}\big[\big||B| - k\big|\big] \le 2s$.*

PROOF. We start by upper bounding the relevant term by the variance of $|B|$. That is,

$$\mathbb{E}\big[\big||B| - k\big|\big] \le \mathbb{E}\big[\big(|B| - k\big)^2\big] = \mathrm{Var}\big(|B|\big),$$

where the first equality uses that $|B| - k$ is integer-valued and the second equality uses that the expected value of $|B|$ is $k$. By bounding our term by the variance of $|B|$, we can exploit that $|B|$ is the sum of $n$ independent Bernoulli variables, where the variance of the Bernoulli variable with bias $p_i$ is $p_i\,(1 - p_i)$. Therefore,

$$\mathrm{Var}\big(|B|\big) = \sum_{i \in [n]} p_i\,(1 - p_i) \le \sum_{i \in [k]} (1 - p_i) + \sum_{i \notin [k]} p_i = 2s,$$

where the last equality follows from $s = \sum_{i \in [k]} (1 - p_i) = \sum_{i \notin [k]} p_i$.                    □

We now express the derivative of the denominator and link it to the probability that a set sampled from a Poisson trial is of size $k - 1$. The proof, which we defer to Appendix D.1, takes the derivative for all sizes $\ell$ of $B$ that are relevant in the term of the expectation, i.e., $0 \le \ell \le k - 1$, and shows that all terms besides one term for $\ell = k - 1$ cancel out.

LEMMA 3.10. *Let $B$ be as in Lemma 3.8. Let $\hat{B} \subseteq \{2, 3, \ldots, n-1\}$ be distributed according to a Poisson trial as above, but omitting parties 1 and n. Then,*

$$\left(\tfrac{\partial}{\partial p_1} - \tfrac{\partial}{\partial p_n}\right) \mathbb{E}\big[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\big] = (p_n - p_1)\, \mathbb{P}\big[|\hat{B}| = k - 1\big].$$

The last step before turning towards proving Proposition 3.6 is to provide a lower bound for the derivative of the denominator, which we just calculated in Lemma 3.10. The key idea is to lower bound the probability that $|\hat{B}|$ is of size $k - 1$ by the probability that $\hat{B}$ equals exactly $\{2, \ldots, k\}$, which we can then use to connect this probability to our parameter $s$. At first sight, the bound appears to be rather weak, e.g., when $s = \sum_{i \notin [k]} p_i > \frac{1}{2}$, the bound is dominated by the trivial lower bound of 0. However, we will only need this bound in the extreme case that $s < p_1$.

LEMMA 3.11. *Let $\hat{B}$ be as in Lemma 3.10. Then,*

$$\mathbb{P}\big[|\hat{B}| = k - 1\big] \geq \frac{1 - 2s}{p_1(1 - p_n)}.$$

PROOF. One way in which $|\hat{B}|$ could equal $k - 1$ is if $\hat{B}$ is exactly $\{2, \ldots, k\}$. Thus,

$$\mathbb{P}\big[|\hat{B}| = k - 1\big] \geq \prod_{2 \leq j \leq k} p_j \prod_{k+1 \leq j \leq n-1} (1 - p_j) = \frac{\prod_{j \in [k]} p_j \prod_{k+1 \leq j \leq n}(1 - p_j)}{p_1(1 - p_n)} = \frac{\mathbb{P}\big[B = [k]\big]}{p_1(1 - p_n)}.$$

We lower-bound $\mathbb{P}\big[B = [k]\big]$ using a union bound. Specifically, we need to rule out that any of the Bernoulli trials indexed $1 \leq j \leq k$ fails or that any of the Bernoulli trials indexed $k + 1 \leq j \leq n$ succeeds. Thus,

$$\mathbb{P}\big[B = [k]\big] \geq 1 - \sum_{1 \leq j \leq k}(1 - p_j) - \sum_{k+1 \leq j \leq n} p_j = 1 - 2s,$$

since $s = \sum_{i \in [k]}(1 - p_i) = \sum_{i \notin [k]} p_i$. The claim follows by combining both inequalities. $\quad\square$

We are now ready to prove that the partial derivative of $\mathbb{P}[S = [k]]$ is nonnegative. We start by expressing the partial derivatives of the numerator of Equation (2), then apply the quotient rule and finally distinguish the two cases $s \geq p_1$ and $s < p_1$. The former case will require little additional effort, while the latter requires the application of Lemma 3.9 and Lemma 3.11.

PROOF OF PROPOSITION 3.6. Recall that we want to bound the partial derivatives of the term $f([k])/\sum_{A' \in \binom{[n]}{k}} f(A')$, where $f(A) = \sum_{i \in A}(1 - p_i) \prod_{j \in A} p_j \prod_{j \notin A}(1 - p_j)$. We begin by computing partial derivatives for the numerator $f([k])$:

$$\begin{aligned}
\tfrac{\partial}{\partial p_1} f([k]) &= \tfrac{\partial}{\partial p_1} \sum_{i \in [k] \setminus 1}(1 - p_i) \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]}(1 - p_j) \qquad\qquad\qquad\qquad\qquad\qquad (3) \\
&\quad + \tfrac{\partial}{\partial p_1} 1 \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]}(1 - p_j) - \tfrac{\partial}{\partial p_1} p_1 \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]}(1 - p_j) \\
&= \sum_{i \in [k] \setminus 1}(1 - p_i) \cdot \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]}(1 - p_j) \\
&\quad + \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]}(1 - p_j) - 2 p_1 \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]}(1 - p_j) \\
&= \sum_{i \in [k]}(1 - p_i) \cdot \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]}(1 - p_j) - \prod_{j \in [k]} p_j \prod_{j \notin [k]}(1 - p_j) \\
&= \left(\frac{1}{p_1} - \frac{1}{\sum_{i \in [k]}(1 - p_i)}\right) f([k]) = \left(\frac{1}{p_1} - \frac{1}{s}\right) f([k]).
\end{aligned}$$

$$\tfrac{\partial}{\partial p_n} f([k]) = -\sum_{i \in [k]}(1 - p_i) \cdot \prod_{j \in [k]} p_j \prod_{\substack{j \notin [k] \\ j \neq n}}(1 - p_j) = -\frac{f([k])}{1 - p_n}. \qquad\qquad\qquad (4)$$

We now apply the quotient rule and plug in equalities we have derived so far:

$$\left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n}\right) \frac{f([k])}{\sum_{A' \in \binom{[n]}{k}} f(A')} = \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n}\right) \frac{f([k])}{\mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right]} \qquad \text{(by Lemma 3.8)}$$

$$= \frac{\left(\left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n}\right) f([k])\right) \cdot \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] - f([k]) \cdot \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n}\right) \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right]}{\mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right]^2}$$

$$= \frac{\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n}\right) \cdot f([k]) \cdot \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] - f([k]) \cdot (p_n - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right]}{\mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right]^2}$$

$$\text{(by Eqs. (3) and (4) and Lemma 3.10)}$$

$$= \underbrace{\frac{f([k])}{\mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right]^2}}_{\geq 0} \left(\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n}\right) \cdot \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] - (p_n - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right]\right)$$

Thus, it suffices to show the nonnegativity of the factor

$$\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n}\right) \cdot \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] - (p_n - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right]. \qquad (5)$$

We will prove the nonnegativity by distinguishing two cases, depending on whether $s \geq p_1$ or not. Surprisingly, the case $s \geq p_1$ — which intuitively seem more common because the residues of all parties outside of $[k]$ outweigh that of the single party 1 — is straight-forward. The hard case of the proof is when $s < p_1$, i.e., when the residues in $[k]$ are almost 1 and the residues outside $[k]$ almost 0.

**Case $s \geq p_1$.** Under this assumption, clearly $1/p_1 - 1/s \geq 0$, which allows us to ignore these terms in Eq. (5). Next, we observe that $\mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] \geq \mathbb{P}\left[|B| = k - 1\right]$. Moreover, since one way for the number of successes $|B|$ in a Poisson trial being equal to $k - 1$ is if there are $k - 1$ successes among the Bernoulli variables indexed $2, \ldots, n - 1$ and additionally the Bernoulli variables indexed 1 and $n$ both fail, it must hold that $\mathbb{P}\left[|B| = k - 1\right] \geq (1 - p_1)(1 - p_n) \mathbb{P}\left[|\hat{B}| = k - 1\right]$. These two inequalities together imply that $\mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right]/(1 - p_n) \geq (1 - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right] \geq (p_n - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right]$, which shows the claimed nonnegativity of the term in Eq. (5).

**Case $s < p_1$.** Note that $p_1 > s = \sum_{i \notin [k]} p_i \geq p_n$. Thus, the term $(p_n - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right]$ we subtract in Eq. (5) is nonpositive, i.e., it can only help us in proving nonnegativity. If the coefficient $\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n}$ was nonnegative, Eq. (5) would be trivially nonnegative; hence, we will assume that it is negative instead. Knowing the signs of these coefficients allows us to apply the bounds we have previously derived:

$$\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n}\right) \cdot \mathbb{E}\left[\mathbb{1}\{|B| < k\} \cdot (k - |B|)\right] - (p_n - p_1) \mathbb{P}\left[|\hat{B}| = k - 1\right]$$

$$\geq \left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n}\right) \cdot s - (p_n - p_1) \frac{1 - 2s}{p_1(1 - p_n)} \qquad \text{(by Lemmas 3.8, 3.9, and 3.11)}$$

$$= \frac{1}{p_1(1 - p_n)} \cdot \left((1 - p_n)s - p_1(1 - p_n) + p_1 s - (p_n - p_1)(1 - 2s)\right)$$

$$= \frac{1}{p_1(1 - p_n)} \cdot \left(s - p_1 s + p_1 p_n - p_n + p_n s\right) \qquad \text{(note } s - sp_1 \geq 0\text{)}$$

$$\geq \frac{1}{p_1 \left(1 - p_n\right)} \cdot \left(p_n \left(p_1 - 1 + s\right)\right) \geq \frac{p_n}{p_1 \left(1 - p_n\right)} \cdot \left(p_1 - 1 + \left(1 - p_1\right)\right) = 0.$$

This establishes that, in both cases, $\left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n}\right) \mathbb{P}[S = [k]]$ is nonnegative as claimed. □

Since we have established (informally in this section and formally in Appendix C) that Proposition 3.6 implies selection monotonicity, the above proof implies our main result, Theorem 1.1.

### 3.4 Potential Implications for Dependent Randomized Rounding

The three properties that turned pipage rounding into such a successful tool — ex-ante proportionality, selecting exactly $k$ units without replacement, and notions of negative correlation [30] — are equally satisfied by Sampford rounding [6]. Hence, in all algorithms that use dependent rounding (subject to a cardinality constraint) as a black-box subroutine [e.g., 8, 9, 12], swapping out pipage rounding for Sampford rounding would preserve the algorithm's guarantees, while gaining Sampford's monotonicity properties "for free". Though we have yet to explore how to leverage this monotonicity in specific applications, we sketch below one path in which Sampford's monotonicity could add incentive guarantees to algorithms using dependent rounding.

Dependent randomized rounding is commonly used in computer science to develop algorithms, in particular approximation algorithms, for combinatorial problems. Dependent rounding can be seen as randomly rounding a point $x$ within an integral polytope $P$ to a vertex $X$ of $P$, ensuring that the expected value of $X$ equals $x$. This is a useful subroutine in many algorithms for NP-hard optimization problems: First, locate a point in the space of the polytope, typically achieved by solving a fractional relaxation of the problem. Then, this point is rounded in a way that ensures that a linear function (or a continuous extension of a submodular function) satisfies Chernoff-style concentration bounds. These guarantees can then be used to bound the objective value with high probability, thus delivering approximation guarantees. As Brändén and Jonasson [6] showed, Sampford sampling, like pipage rounding, satisfies the *strong Rayleigh* property, which implies extremely general concentration properties.

In such applications of dependent rounding, selection monotonicity (and similar monotonicity properties that Sampford rounding might possess, see Conjecture 4.4) might for example allow to construct approximation algorithms that can be implemented as truthful mechanisms. In our case, the polytope $P$ is simply the base polytope of the uniform matroid, which has proved influential in applications including Steiner tree problems [30], $k$-median [9], committee selection [12], and, recently, online algorithms [25].

Concretely, consider an approximation algorithm that consists of the two steps mentioned above: *(1)* determine a point $x$ inside $P$, then *(2)* randomly round $x$ to select a vertex $X$ of $P$. Let each agent $i$ have a preferred vertex $X_i$ of $P$ with utility $\tau_i \cdot \mathbb{P}[X = X_i]$, which depends on their private type $\tau_i \in \mathbb{R}$. These types are inputs to the approximation algorithm, whose objective need not be related to the utilities. Suppose that step *(1)* of the algorithm is monotone in the reported types, i.e., whenever agent $i$ increases their reported type, this causes $x_j$ to weakly increase for the $k$ many matroid elements $j$ that correspond to the vertex $X_i$ and $x_j$ to weakly decrease for all other $j$. Then, the selection monotonicity of Sampford rounding implies an overall monotonicity: if agent $i$ increases their report, $X_i$ will be more likely to be sampled. Hence, by charging the agents Myerson payments [24], we obtain a mechanism that is incentive compatible in expectation (and hence the approximation algorithm receives truthful inputs). Should Sampford sampling even satisfy stronger notions of proportionality (as we conjecture in Section 4.2), this blueprint extends to richer utilities; for example, agent $i$ could care about a subset $T_i$ of matroid elements, have a monotone

nondecreasing function $f_i : \mathbb{N} \to \mathbb{R}$ that expresses diminishing returns or returns to scale, and the agent's utility would be $\tau_i \cdot \mathbb{E}\left[f_i(|X \cap T_i|)\right]$.

## 3.5 Selection Monotonicity Implies Continuity

We note that all rules discussed in this paper are continuous: small changes to the input residues do not result in large changes to the probability that any set is selected. As the following result shows, this is a necessary requirement for selection monotonicity.

THEOREM 3.12. *Any selection monotone rounding rule is Lipschitz continuous. Specifically, let $\vec{p}, \vec{p}'$ be two residue profiles, and let $T$ be a set of $k$ parties. Then,*

$$\left|\mathbb{P}\left[r(\vec{p}') = T\right] - \mathbb{P}\left[r(\vec{p}) = T\right]\right| \le \|\vec{p} - \vec{p}'\|_1.$$

As in the proof of Theorem 1.1, we first consider the special case when only two residues are changing at a time.

LEMMA 3.13. *Consider a selection monotone rounding rule $r$, a residue profile $\vec{p}$, and $0 < \delta < \min(p_2, 1 - p_1)$. Define a residue profile $\vec{p}'$, in which $p_1' \coloneqq p_1 + \delta$, $p_2' \coloneqq p_2 - \delta$, and all other $p_i' \coloneqq p_i$. Then, for any set $T$ of $k$ parties, it holds that*

$$\left|\mathbb{P}\left[r(\vec{p}') = T\right] - \mathbb{P}\left[r(\vec{p}) = T\right]\right| \le 2\,\delta. \tag{6}$$

PROOF. We may assume w.l.o.g. that $\mathbb{P}\left[r(\vec{p}') \supseteq \{1, 2\}\right] \ge \mathbb{P}\left[r(\vec{p}) \supseteq \{1, 2\}\right]$; otherwise, simply change the roles of parties 1 and 2, and the roles of $\vec{p}, \vec{p}'$. We break the proof up into three cases, depending on how $T$ intersects $\{1, 2\}$.

**Case 1: Sets $T$ such that $1 \in T \not\ni 2$.** By ex-ante proportionality, $\sum_{T \in \binom{[n]}{k}, 1 \in T} \mathbb{P}[r(\vec{p}) = T] = p_1$ and $\sum_{T \in \binom{[n]}{k}, 1 \in T} \mathbb{P}[r(\vec{p}') = T] = p_1 + \delta$. For all $T$ that contain 1 but not 2, selection monotonicity ensures that $\mathbb{P}[r(\vec{p}') = T] \ge \mathbb{P}[r(\vec{p}) = T]$. Hence,

$$\begin{aligned}
\delta &= (p_1 + \delta) - p_1 \\
&= \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T}} \mathbb{P}[r(\vec{p}') = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T}} \mathbb{P}[r(\vec{p}) = T] \\
&= \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T \not\ni 2}} \left(\mathbb{P}[r(\vec{p}') = T] - \mathbb{P}[r(\vec{p}) = T]\right) + \underbrace{\sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(\vec{p}') = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(\vec{p}) = T]}_{=\mathbb{P}\left[r(\vec{p}') \supseteq \{1,2\}\right] - \mathbb{P}\left[r(\vec{p}) \supseteq \{1,2\}\right] \ge 0} \\
&\ge \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T \not\ni 2}} \underbrace{\left(\mathbb{P}[r(\vec{p}') = T] - \mathbb{P}[r(\vec{p}) = T]\right)}_{\ge 0}.
\end{aligned}$$

Hence, for each set $T$ in the sum, $0 \le \mathbb{P}[r(\vec{p}') = T] - \mathbb{P}[r(\vec{p}) = T] \le \delta$; i.e., Equation (6) holds for all sets $T$ containing 1 but not 2. Note that this chain of inequalities also shows that $\mathbb{P}\left[r(\vec{p}') \supseteq \{1, 2\}\right] - \mathbb{P}\left[r(\vec{p}) \supseteq \{1, 2\}\right] \le \delta$, which we will use in the second case.

**Case 2: Sets $T$ such that $2 \in T \not\ni 1$.** By ex-ante proportionality, $\sum_{T \in \binom{[n]}{k}, 2 \in T} \mathbb{P}[r(\vec{p}) = T] = p_2$ and $\sum_{T \in \binom{[n]}{k}, 2 \in T} \mathbb{P}[r(\vec{p}') = T] = p_2 - \delta$. For all $T$ that contain 2 but not 1, selection monotonicity

ensures that $\mathbb{P}[r(\vec{p}') = T] \leq \mathbb{P}[r(\vec{p}) = T]$. Hence,

$$
\begin{aligned}
\delta &= p_2 - (p_2 - \delta) \\
&= \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T}} \mathbb{P}[r(\vec{p}) = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T}} \mathbb{P}[r(\vec{p}') = T] \\
&= \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T \not\ni 1}} \left( \mathbb{P}[r(\vec{p}) = T] - \mathbb{P}[r(\vec{p}') = T] \right) + \underbrace{\sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(\vec{p}) = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(\vec{p}') = T]}_{= \mathbb{P}\left[r(\vec{p}) \supseteq \{1,2\}\right] - \mathbb{P}\left[r(\vec{p}') \supseteq \{1,2\}\right] \geq -\delta},
\end{aligned}
$$

which implies Eq. (6) for this case:

$$
2\delta \geq \sum_{2 \in T \not\ni 1} \underbrace{\left( \mathbb{P}[r(\vec{p}) = T] - \mathbb{P}[r(\vec{p}') = T] \right)}_{\geq 0}.
$$

**Case 3: Sets $T$ such that** $1, 2 \notin T$ **or** $1, 2 \in T$. Pick a different party whose membership in $T$ is opposite that of 1 and 2, i.e., if $1, 2 \in T$, then choose a party in $T$, and if $1, 2 \in T$, choose a party outside of $T$. Without loss of generality, let this party be party 3. We may also assume without loss of generality that $\delta$ is less than $1 - p_3$ (which is positive since residues are less than 1, by definition). Else, i.e., if $\delta$ is larger, we can split up the increase of $p_1$ by $\delta$ and decrease of $p_2$ by $\delta$ into finitely many steps in which these residues successively increase or decrease, by some $\delta' < 1 - p_3$ in each step. Our claim for the original $\delta$ then follows by applying the triangle inequality to the bounds for the smaller steps.

Let $\vec{p}''$ be a new residue profile, where $p_1'' := p_1 = p_1' - \delta$, $p_2'' := p_2 + \delta = p_2'$ and $p_3'' = p_3 + \delta = p_3' + \delta$. Note that this is a valid residue profile by how we chose $\ell$. The key observation is that we already know the bound in Equation (6) holds going from $\vec{p}$ to $\vec{p}''$ and from $\vec{p}''$ to $\vec{p}'$, since the residues changing involve one party in $T$ and one party outside of $T$. Thus,

$$
\begin{aligned}
&\left| \mathbb{P}\left[r(\vec{p}') = T\right] - \mathbb{P}\left[r(\vec{p}) = T\right] \right| \\
&= \left| \underbrace{\left( \mathbb{P}\left[r(\vec{p}') = T\right] - \mathbb{P}\left[r(\vec{p}'') = T\right] \right)}_{\in [-2\delta, 2\delta]} + \underbrace{\left( \mathbb{P}\left[r(\vec{p}'') = T\right] - \mathbb{P}\left[r(\vec{p}) = T\right] \right)}_{\in [-2\delta, 2\delta]} \right|.
\end{aligned}
$$

Naïvely, this would yield a bound of $4\delta$, but we can sharpen it by observing that selection monotonicity ensures that both of the terms above have opposite signs: If $1, 2 \notin T$ and $3 \in T$, then the first of these two terms is $\leq 0$ and the second $\geq 0$, whereas if $1, 2 \in T$ and $3 \notin T$, then the first term is $\geq 0$ and the second term is $\leq 0$. Either way, the sum of these terms is at most $2\delta$, so Equation (6) holds. □

PROOF OF THEOREM 3.12. Let $\vec{p}, \vec{p}'$ be an arbitrary pair of residue profiles. Define the following sets of parties:

$$
\begin{aligned}
S_+ &:= \{j \in [n] \mid p_j' > p_j\}, \\
S_- &:= \{j \in [n] \mid p_j' < p_j\}.
\end{aligned}
$$

Consider the following algorithm, which transforms $\vec{p}$ into $\vec{p}'$:

$\vec{p}^1 \leftarrow \vec{p}$;
$i \leftarrow 1$;
**while** $\vec{p}^i \neq \vec{p}'$ **do**
  $j_1 \leftarrow$ any element of $S_+$ such that $p'_{j_1} > p^i_{j_1}$;
  $j_2 \leftarrow$ any element of $S_-$ such that $p'_{j_2} < p^i_{j_2}$;
  $\delta_i \leftarrow \min\{p'_{j_1} - p^i_{j_1}, p^i_{j_2} - p'_{j_2}\}$;
  $\vec{p}^{i+1} \leftarrow$ copy of $\vec{p}^i$, except $p^{i+1}_{j_1} := p^i_{j_1} + \delta_i$ and $p^{i+1}_{j_2} := p^i_{j_2} - \delta_i$;
  $i \leftarrow i + 1$;
**end**

Observe that, in each round of this algorithm, we make a new component of $\vec{p}^i$ equal to that of $\vec{p}'$. Thus, the algorithm terminates after at most $n - 1$ iterations, producing a sequence of residue profiles $\vec{p} = \vec{p}^1, \vec{p}^2, \vec{p}^3, \ldots, \vec{p}^\ell = \vec{p}'$ for some $\ell \geq n$. It is easy to see that the total sum of the changes $\delta_i$ at each iteration is precisely the sum of differences between $\vec{p}$ and $\vec{p}'$ across $S_+$ (or, alternatively, across $S_-$). Thus,

$$\sum_{i=1}^{\ell-1} \delta_i = \sum_{j \in S^+} (p'_j - p_j) \leq \max_{S \subseteq [n]} \sum_{j \in S} (p'_j - p_j) = d_{\mathrm{TV}}(\vec{p}, \vec{p}') = \frac{1}{2} \|\vec{p}' - \vec{p}\|_1, \qquad (7)$$

where we have applied the well-known formula for the total variation distance $d_{\mathrm{TV}}$.[7]

Note that Lemma 3.13 applies to each consecutive pair of residue profiles $\vec{p}^i$ and $\vec{p}^{i+1}$, since they differ in only two coordinates by $\delta_i$. Hence, we have

$$\left| \mathbb{P}\left[r(\vec{p}') = T\right] - \mathbb{P}\left[r(\vec{p}) = T\right] \right| \leq \sum_{i=1}^{\ell-1} \left| \mathbb{P}\left[r(\vec{p}^{i+1}) = T\right] - \mathbb{P}\left[r(\vec{p}^i) = T\right] \right| \qquad \text{(triangle inequality)}$$

$$\leq \sum_{i=1}^{\ell-1} 2\delta_i \qquad \text{(from Lemma 3.13)}$$

$$\leq 2 \cdot \frac{1}{2} \|\vec{p} - \vec{p}'\|_1 \qquad \text{(from Equation (7))}$$

$$= \|\vec{p} - \vec{p}'\|_1. \qquad \square$$

## 4 MONOTONICITY OF APPORTIONMENT METHODS

Returning to our full setting of randomized apportionment, what does the (pairwise) selection monotonicity of Sampford rounding win us? Fix an "old" and a "new" election with respective votes $\vec{v}, \vec{v}'$ and the same house size $h$, as well as a coalition $T$ of parties. Suppose that each party $i \in T$ has increased its share of the vote from the old election to the new one, whereas each party $i \notin T$ has lost in vote share. Then, if no party's vote share has shifted by so much that its lower quota changed and the sum of residues equals $|T|$, the apportionment method induced by Sampford rounding gives $T$ a higher chance of jointly being rounded up, by Theorem 1.1. As a result, the paradox from Section 1.1 cannot happen for this method.

When lower quotas change between elections, however, the entire concept of "$T$ being jointly rounded up" stops making much sense. Returning to the example from the introduction, suppose that, instead of the exodus of voters from the left-wing party 5, 40 voters from the right-wing party 4 had moved to party 5, which among other things would increase party 5's lower quota from 0 to 1. Since this shift also decreases the sum of all residues from 3 to 2, no apportionment method

---

[7]This formula still holds even when probabilities sum to some constant other than 1. Also, we really have equality in the middle transition as well, but that is not necessary to argue for this proof.

satisfying quota can "round up" all three left-wing parties. But this line of thinking muddles that the increase of party 5's lower quota is good news for the left-wing coalition: party 5 receiving one seat has turned from a possibility into certainty, and party 5 might even receive a second seat!

In our view, the right generalization of the paradox to changing lower quotas is by considering how likely the coalition is to exceed certain *thresholds* of house seats. In our example, the relevant threshold was that of a majority of seats; broadly speaking, party 5 increasing its lower quota ought to increase the prospects of the coalition gaining such a majority. Besides the simple majority, a coalition might care about its probability of exceeding other seat thresholds, which might turn it into a supermajority, a blocking minority, or a parliamentary group with certain privileges [e.g., 14, rules 33–34]. Ideally, an apportionment method should have the property that the probability of $T$ exceeding a threshold is monotone, for all thresholds simultaneously.

In this section, we capture a monotonicity in the probability of crossing thresholds in four axioms. Two of them will not be achievable for any randomized apportionment method, and one more axiom is at least impossible if the apportionment method satisfies a natural condition. The satisfiability of the fourth monotonicity axiom is the most tantalizing open question of our work; we conjecture that it is satisfied by the Sampford apportionment method, and show that Grimmett's method satisfies it for coalitions of size $T = 2$.

## 4.1 Axioms and Apportionment Methods

The apportionment methods we study are induced by the rounding rules from the last section, through the construction defined in Section 2. We refer to the apportionment method induced by conditional Poisson rounding as the *conditional Poisson method*, to the method induced by Sampford rounding as the *Sampford method*, and so forth. An exception is the method induced by systematic rounding, which we refer to as *Grimmett's method* [19]. Recall that all of these apportionment methods satisfy quota and ex-ante proportionality.

The main way in which we strengthen monotonicity for apportionment is *threshold monotonicity*:

*Definition 4.1 (Threshold monotonicity).* Let $\vec{v}, \vec{v}' \in \mathbb{R}^n_{\geq 0}$ be two vote vectors, and let $h \in \mathbb{N}$. Let $T$ be a set of parties such that $q'_i \geq q_i$ for all $i \in T$[8] and $q'_i \leq q_i$ for all $i \notin T$. An apportionment method $a$ satisfies threshold monotonicity if it always holds, for all thresholds $\vartheta \in \mathbb{N}$, that

$$\mathbb{P}_{\vec{\alpha} \sim a(\vec{v}', h)}\left[\textstyle\sum_{i \in T} \alpha_i \geq \vartheta\right] \geq \mathbb{P}_{\vec{\alpha} \sim a(\vec{v}, h)}\left[\textstyle\sum_{i \in T} \alpha_i \geq \vartheta\right].$$

In other words, the axiom requires the random variable $\sum_{i \in T} a(\vec{v}', h)_i$, which describes the total number of seats awarded to $T$ for the votes $\vec{v}'$, to first-order stochastically dominate the corresponding random variable for the votes $\vec{v}$. We also define a pairwise analogue to the axiom:

*Definition 4.2 (Pairwise threshold monotonicity).* Let $\vec{v}, \vec{v}' \in \mathbb{R}^n_{\geq 0}$ be two vote vectors, and let $h \in \mathbb{N}$. Let $T_1, T_2$ be two sets of parties such that $q'_i \geq q_i$ for all $i \in T_1$ and $q'_i \leq q_i$ for all $i \in T_2$. An apportionment method $a$ satisfies pairwise threshold monotonicity if always at least one of the following two statements holds:

$$\forall \vartheta \in \mathbb{N}, \mathbb{P}_{\vec{\alpha} \sim a(\vec{v}', h)}\left[\textstyle\sum_{i \in T_1} \alpha_i \geq \vartheta\right] \geq \mathbb{P}_{\vec{\alpha} \sim a(\vec{v}, h)}\left[\textstyle\sum_{i \in T_1} \alpha_i \geq \vartheta\right] \quad \text{or}$$

$$\forall \vartheta \in \mathbb{N}, \mathbb{P}_{\vec{\alpha} \sim a(\vec{v}', h)}\left[\textstyle\sum_{i \in T_2} \alpha_i \geq \vartheta\right] \leq \mathbb{P}_{\vec{\alpha} \sim a(\vec{v}, h)}\left[\textstyle\sum_{i \in T_2} \alpha_i \geq \vartheta\right].$$

Though these axioms (for apportionment methods) cannot syntactically be compared to our previous monotonicity axioms (for rounding rules), they should be thought of as strengthenings of their selection equivalents. This relationship can be made formal using the embedding from Section 2: Any apportionment method satisfying (pairwise) threshold monotonicity, when translated

---

[8]Or, equivalently, we can require that $i$'s share of the vote $\frac{v_i}{\sum_{j \in [n]} v_j} = q_i/h$ increases rather than its quota.

into a rounding rule, satisfies (pairwise) selection monotonicity, by selecting the threshold so that all parties in the coalition must be selected to meet the threshold.

We will show that pairwise threshold monotonicity is incompatible with a natural condition on apportionment methods:

*Definition 4.3 (Full support).* An apportionment method $a$ satisfies full support if, for all vote vectors $\vec{v}$, house sizes $h$, and apportionment vectors $\vec{\alpha}$ that satisfy quota (i.e., $\vec{\alpha} \in \mathbb{N}^n$, $\sum_{i \in [n]} \alpha_i = h$, and $\alpha_i \in \{\lfloor q_i \rfloor, \lceil q_i \rceil\}$ for all $i \in [n]$), $a(\vec{v}, h)$ will take on the value $\vec{\alpha}$ with positive probability.

The definitions of Sampford rounding immediately implies that the Sampford method has full support. Similarly, the conditional Poisson method has full support, since, if the residue of any party $i$ is non-zero, its working probability $\pi_i$ must also be positive by ex-ante proportionality.

Finally, we define variants of the two preceding axioms whose precondition argues about changes not to the standard quotas or vote shares, but to changes in the raw vote counts: *Vote-count threshold monotonicity* is defined just like threshold monotonicity, except requiring that $v'_i \geq v_i$ for all $i \in T$ and $v'_i \leq v_i$ for all $i \notin T$, instead of the matching assumptions for the $q_i$ and $q'_i$. In the same way, we define *pairwise vote-count threshold monotonicity* like its non-vote-count variant, where we now require that $v'_i \geq v_i$ for all $i \in T_1$ and $v'_i \leq v_i$ for all $i \in T_2$. Under the mild assumption that an apportionment method satisfies homogeneity, i.e., that scaling the vote vector by a constant doesn't change the apportionment, (pairwise) vote-count threshold monotonicity implies (pairwise) threshold monotonicity.

## 4.2 Conjecture and $|T| = 2$ Possibility for Threshold Monotonicity

We leave the most pressing question, finding an apportionment method that satisfies threshold monotonicity, unresolved. Based on extensive computational search for counterexamples, however, we conjecture that the Sampford method satisfies this axiom:

CONJECTURE 4.4. *The Sampford method satisfies threshold monotonicity.*

Proving this conjecture, of course, appears technically very difficult. A major complication relative to our proof of selection monotonicity is that, in many of the apportionment outcomes forming the event "$T$ receives at least $\vartheta$ seats", some of the rounded seats go to parties outside of $T$ whose vote share is shrinking, which means that the probability of many of these outcomes will decrease for the Sampford method. Thus, a proof for our conjecture will require a way to charge the decreases in probability of some outcomes to increases in probability of other outcomes.

Still, we do have some limited good news: Though Grimmett's method does not satisfy threshold monotonicity for coalitions of three and more parties (not even selection monotonicity, as we saw in Section 1.1), it satisfies the axiom for coalitions of size 2.

THEOREM 4.5. *Grimmett's method satisfies threshold monotonicity for coalitions $T$ of size $\leq 2$.*

PROOF OF PROPOSITION 4.5. We prove threshold monotonicity of the variant of Grimmett's method that uses an arbitrary fixed ordering of the states, clearly implying threshold monotonicity for a randomized order as well. For our analysis, we will modify our presentation of Grimmett's method in two ways. First, instead of giving all parties their lower quotas and then lining up their residues as intervals on the number line, we directly line up their standard quotas as intervals on the number line. Then each party gets a number of seats equal to the number of integers their interval contains. This is clearly equivalent: since extending the interval of each party $i$ from length $p_i$ to length $q_i = p_i + \lfloor q_i \rfloor$ adds an integral length $\lfloor q_i \rfloor$, the party's interval will contain exactly $\lfloor q_i \rfloor$ more integers than previously, and the shift of other parties remains the same modulo 1. Second, we arbitrarily shift all the intervals to the right by some amount $u_0$. Again, this has no bearing on
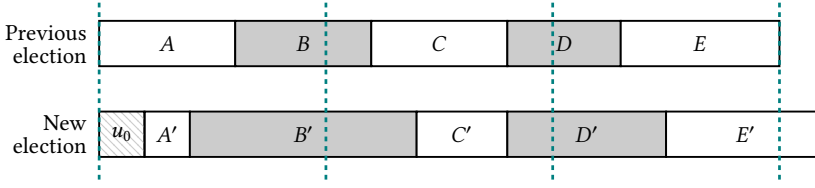
Fig. 2. Illustration of the proof of Theorem 4.5 stating that Grimmett's method satisfies threshold monotonicity for $|T| = 2$. Intervals corresponding to growing parties are gray.

the probability distribution of numbers of seats awarded, since all intervals are additionally shifted by a random $u$ drawn uniformly from $[0, 1)$, and the distribution of $u + u_0$ modulo 1 is identical to that of $u$ modulo 1. We call this equivalent implementation of Grimmett's method $G(u_0)$.

If $|T| = 1$, there is nothing to show: proportionality alone implies threshold monotonicity. So suppose $|T| = \{i, j\}$, where $1 \leq i < j \leq n$. First consider running $G(0)$ on an arbitrary input $(\vec{v}, h)$. Before adding the random shift $u$, let $A \cup B \cup C \cup D \cup E$ be the partition of $[0, h)$ where $A$ is the union of all intervals for parties numbered lower than $i$, $B$ is the interval for party $i$, $C$ is the union of all intervals numbered between $i$ and $j$, $D$ is the interval for party $j$, and $E$ is the union of all intervals for parties numbered higher than $j$. Though we know that the sizes of the intervals corresponding to $i$ and $j$ increase when going from $(\vec{v}, h)$ to $(\vec{v}', h)$ for $G(0)$, it need not be true that the new intervals contain the previous ones. To overcome this issue, we show in Lemma D.1 in Appendix D.2 that there exists a shift $u_0 \in \mathbb{R}$ such that, when $A'$, $B'$, $C'$, $D'$, and $E'$ are the new respective intervals in running $G(u_0)$ on input $(\vec{v}', h)$, it holds that $B \subseteq B'$ and $D \subseteq D'$. See Figure 2 for an illustration of an example. Thus, for any realization of the random shift $u$ that awarded a given threshold of $t$ seats to the parties in $T$ when running $G(0)$ on $(\vec{v}, h)$, the same value of $u$ will award at least $t$ seats to the parties in $T$ when running $G(u_0)$ on $(\vec{v}', h)$, since the same $t$ integers that were included in $B$ and $D$ will be contained within $B'$ and $D'$. Since $G(0)$ and $G(u_0)$ are both equivalent to Grimmett's method, we have shown that it satisfies threshold monotonicity. □

As discussed, the fact that systematic rounding violates selection monotonicity for coalitions of size 3 implies that Grimmett's method violates threshold monotonicity for $|T| = 3$. Our example in Figure 1 also illustrates why the approach applied in the proof of Theorem 4.5 fails for larger coalitions: For the old interval of party 3 to be included in its new interval, we would need to have a positive shift $u_0 > 0$. However, any positive shift would violate the same constraint for party 1.

### 4.3 Impossibility of Pairwise Threshold Monotonicity

When we considered selection monotonicity and its pairwise variant in Section 3, the most promising rounding rules seemed those whose probability distributions spread their mass smoothly and broadly over the possible outcomes (like conditional Poisson and Sampford rounding, and unlike systematic and pipage rounding). Furthermore, an easy argument showed that Sampford rounding satisfied the pairwise variant of selection monotonicity.

In contrast, when moving to pairwise *threshold* monotonicity, we show that no apportionment method whose distributions have such a full support, and thus neither the conditional Poisson method nor the Sampford method, can satisfy the axiom.

THEOREM 4.6. *No (ex-ante proportional and quota) apportionment method that has full support can satisfy pairwise threshold monotonicity.*

PROOF. Let $a$ be any proportional and optimistic apportionment algorithm. We fix $h = 3$ and consider a collection of vote vectors of size $n = 12$ where all lower quotas are zero. Thus, we may

think of the input to $a$ as just being the residue vector (which is the same as the vector of standard quotas). For any integer $m \geq 2$, we define the residues

$$\vec{p}(m) := \left( \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{m}, 1 - \frac{1}{m} \right)$$

and let $S_m$ be the random variable defined as the set of parties that $a$ rounds up on input $p(m)$. By applying union bound over all events in which $S_m$ contains at least two elements from $\{1, \ldots, 10\}$, we get

$$\sum_{R \in \binom{\{1,\ldots,10\}}{2}} \mathbb{P}[S_m \supseteq R] \geq \mathbb{P}[|S_m \cap \{1, \ldots, 10\}| \geq 2] = \mathbb{P}[|S_m \cap \{11, 12\}| \leq 1]$$

$$= 1 - \mathbb{P}[S_m \supseteq \{11, 12\}] \geq 1 - \mathbb{P}[11 \in S_m]$$

$$= 1 - \frac{1}{m} \geq \frac{1}{2}.$$

Therefore, by the averaging principle there must exist some $R \subseteq \{1, \ldots, 10\}$ of size 2 such that

$$\mathbb{P}[S_m \supseteq R] \geq \frac{1}{2 \cdot \binom{10}{2}} = \frac{1}{90}.$$

Thus, for each $m \in \{2, 3, 4, \ldots\}$, choose some $R_m \subseteq \{1, \ldots, 10\}$ of size 2 such that

$$\mathbb{P}[S_m \supseteq R_m] \geq \frac{1}{90}.$$

Since the sequence $R_2, R_3, R_4, \ldots$ only takes finitely-many different values, by the pigeonhole principle there must be some infinite subsequence of indices $m_1, m_2, m_3, \ldots$ such that $R_{m_i}$ is the same for each positive integer $i$. Call this set $T_2$ and choose an arbitrary set $T_1 \subseteq \{1, \ldots, 10\} \setminus T_2$ of size 3. Without loss of generality, we re-label the states so that $T_2 = \{1, 2\}$ and $T_1 = \{3, 4, 5\}$. Now consider the alternative residue vector

$$\vec{p}' := \left( \frac{1}{4}, \frac{1}{4}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{500}, \frac{1}{500}, \frac{1}{500}, \frac{1}{500}, \frac{1}{500}, 1 - \frac{1}{200}, 1 - \frac{1}{200} \right)$$

and let $S'$ be the random set that $a$ rounds up on input $\vec{p}'$. By full support, we know that

$$\mathbb{P}[S' \supseteq T_1] =: \varepsilon > 0.$$

Choose $i$ large enough so that $\frac{1}{m_i} < \varepsilon$. We claim that the pair of residue vectors $(\vec{p}', \vec{p}(m_i))$ violates pairwise threshold monotonicity for the sets $T_1$ and $T_2$, with respective thresholds $\vartheta_1 = 3$ and $\vartheta_2 = 2$ (so we are interested in the probabilities that *all* elements of $T_1$ or $T_2$ are rounded up).

First consider the set $T_2$. Observe that, in passing from $\vec{p}'$ to $\vec{p}(m_i)$, residues within $T_2$ have decreased from $\frac{1}{4}$ to $\frac{1}{5}$. However, since selecting both states in $T_2$ entails excluding at least one of states 11 and 12, we have

$$\mathbb{P}[S' \supseteq T_2] \leq \mathbb{P}[11 \notin S' \text{ or } 12 \notin S'] \leq \mathbb{P}[11 \notin S'] + \mathbb{P}[12 \notin S'] \qquad \text{(by the union bound)}$$

$$= 2 - \mathbb{P}[11 \in S'] - \mathbb{P}[12 \in S'] = 2 - \left(1 - \frac{1}{200}\right) - \left(1 - \frac{1}{200}\right) \qquad \text{(by proportionality)}$$

$$= \frac{1}{100} < \frac{1}{90} \leq \mathbb{P}[S_{m_i} \supseteq R_{m_i}] = \mathbb{P}[S_{m_i} \supseteq T_2],$$

so the probability of selecting all states in $T_2$ has increased.

Next consider the set $T_1$. In passing from $\vec{p}'$ to $\vec{p}(m_i)$, residues in $T_1$ increased from $\frac{1}{6}$ to $\frac{1}{5}$, but

$$\mathbb{P}[S' \supseteq T_1] = \varepsilon > \frac{1}{m_i} = 1 - \left(1 - \frac{1}{m_i}\right) = 1 - \mathbb{P}[12 \in S_{m_i}] \qquad \text{(by proportionality)}$$
$$= \mathbb{P}[12 \notin S_{m_i}] \geq \mathbb{P}[S_{m_i} \supseteq T_1],$$

where in the final inequality we have used the fact that selecting $T_1$ entails not selecting any other state, as $T_1$ has size 3. Thus, the probability of selecting all states in $T_1$ has decreased, and so we have shown that pairwise threshold monotonicity is violated.                                                  □

Though this theorem only rules out methods having full support, we see this as a consequence of the difficulty of making a uniform argument over many partially degenerate apportionment methods, not as an invitation to explore apportionment methods without full support. In particular, the failure of systematic and pipage rounding to satisfy pairwise selection monotonicity implies that their induced methods fail pairwise threshold monotonicity, as described in Section 4.1.

We also remark that our proof leans heavily on the fact that pairwise threshold monotonicity allows vote shares outside of the coalitions to both grow and shrink, which is why we see no indication that threshold monotonicity would be subject to a similar impossibility.

## 4.4 Impossibility of Monotonicity in Raw Vote Counts

In the previous two axioms for apportionment, what entitled a coalition to more seats was if its constituent parties increased their *share* of the votes. This seems especially natural in the party-apportionment setting — for example, when elections are covered in the media, the vote percentages are typically reported as the central statistics of interest. Alternatively, one could consider increases in the raw vote counts as what entitles a coalition to more joint representation. (There is precedent for both perspectives in classical apportionment theory, though this latter perspective is arguably more prominent [e.g., 4, p. 106–108].) The two axioms coincide whenever the total population does not change. When the overall number of votes shifts, however, axioms based on the vote counts may grant a coalition increases in joint representation in additional scenarios.

Unfortunately, formulating (pairwise) threshold monotonicity in terms of vote counts is too much to ask, i.e., the resulting axioms cannot be satisfied by any method. The proof is (as often in apportionment theory) based on the fact that, when we rescale the populations after a decrease in the total number of votes, this rescaling increases the standard quota of large parties by a larger absolute amount than the standard quota of small parties, which can bring the residue of such a large party close to one and force us to select this party almost always.

THEOREM 4.7. *No (ex-ante proportional and quota) apportionment method can satisfy vote-count threshold monotonicity or pairwise vote-count threshold monotonicity.*

PROOF. Let $a$ be any proportional apportionment method. We fix the house size be $h = 8$ and consider the initial vote vector

$$\vec{v} := (380, 140, 140, 140).$$

Let $S$ be the random variable defined as the set of parties that $a$ rounded up on input $(\vec{v}, h)$. Observe that the total number of votes is 800, and thus the lower quotas and residues are given by

$$\lfloor \vec{q} \rfloor = (3, 1, 1, 1) \quad \text{and} \quad \vec{p} = (0.8, 0.4, 0.4, 0.4).$$

Hence, $S$ must be of size 2. By proportionality,

$$0.8 = \mathbb{P}[1 \in S] = 1 - \mathbb{P}[1 \notin S]$$
$$= 1 - (\mathbb{P}[S = \{2, 3\}] + \mathbb{P}[S = \{2, 4\}] + \mathbb{P}[S = \{3, 4\}]).$$

Rearranging, we get

$$\mathbb{P}[S = \{2, 3\}] + \mathbb{P}[S = \{2, 4\}] + \mathbb{P}[S = \{3, 4\}] = 0.2.$$

Thus, by the averaging principle, at least one of these three probabilities must be at least $0.2/3 = 0.0\overline{6}$. Since $\vec{v}$ is symmetric with respect to parties 2, 3, and 4, we may assume without loss of generality that this holds for the first term. That is, letting $T_1 := \{2, 3\}$, we have

$$\mathbb{P}[S = T_1] \geq 0.0\overline{6}.$$

Now consider the alternative vote vector

$$\vec{v}' := (376, 142, 142, 100).$$

The total population is now 760, and the lower quotas remain unchanged. However, the residues are now

$$\vec{p}' = (0.96, 0.05, 0.49, 0.49)$$

(rounded to two decimal places). Let $S'$ be the random set of size 2 that $a$ rounds up on input $(\vec{v}', h)$. Since the residues still sum to 2, rounding up both parties in $T_1$ precludes rounding up party 1, so

$$\mathbb{P}[S' = T_1] \leq \mathbb{P}[1 \notin S'] = 1 - \mathbb{P}[1 \in S']$$
$$= 1 - 0.96 = 0.04. \qquad \text{(by proportionality)}$$

We claim that $a$ violates pairwise vote-count threshold monotonicity for the pair of vote vectors $(\vec{v}, \vec{v}')$, with $T_1 = \{2, 3\}$ and $T_2 := \{1\}$, with thresholds $\vartheta_1 = \vartheta_2 = 4$.

First consider $T_2$, which controls at least 4 seats if and only if it is rounded up, since the lower quota for party 1 (the only party in $T_2$) is 3. In passing from $\vec{v}$ to $\vec{v}'$, the number of votes for this party has decreased from 380 to 376, yet, by proportionality, the probability that $T_2$ is selected must have increased from 0.8 to 0.96.

Next consider $T_1$, which controls at least 4 seats if and only if both of parties 2 and 3 are rounded up, since the lower quotas for these parties are each 1. In passing from $\vec{v}$ to $\vec{v}'$, the numbers of votes for both parties in $T_1$ have increased from 140 to 142, yet we have shown that

$$\mathbb{P}[S = T_1] = 0.0\overline{6} > 0.04 \geq \mathbb{P}[S' \supseteq T_1],$$

so the probability that $T_1$ controls 4 seats has decreased. Thus, $a$ violates pairwise vote-count threshold monotonicity. Since $v_i' \leq v_i$ for all $i \notin T_1$, this also shows that $a$ violates vote-count threshold monotonicity for $T_1$. □

Note that the proof requires $n = 4$ states. This is tight since for $n = 3$ there is only one proportional rounding rule to use, and it is easy to see that it satisfies all monotonicity properties.

## 5 CONCLUSION

From our results, a landscape of monotonicity notions in randomized apportionment and rounding is taking shape. On one side lie the (pairwise) selection monotonicity axioms, which can be attained by Sampford sampling. On the other side, we have the vote-count variants and pairwise threshold monotonicity, overly strong axioms that cannot be satisfied (at least, for the latter, by natural apportionment methods). The next step in the exploration of monotonicity will focus on the axiom in their middle: threshold monotonicity, which we conjecture to be feasible but could not yet show for general-size coalitions.

One motivation to pursue threshold monotonicity is that it would imply a general notion of strategy-proofness for randomized apportionment. Consider a voter who approves a subset of the

parties, in the sense that they want as many legislative seats as possible to be filled from this subset.[9] Suppose that the voter maximizes an expected utility that is a monotone function in the number of seats — a very general assumption, which can in particular capture the value of certain seat thresholds, as well as risk-averse or risk-seeking preferences. A simple consequence of threshold monotonicity is that this voter can never profit from strategically voting for a disapproved party, and this even holds for coalitions of such voters.

Beyond the setting of apportionment, we are broadly intrigued by the prospect of relational axioms [32] on higher-order correlations, i.e., axioms that capture how the higher-order correlations of a random process vary as a function of changes in the process's parameters. Going beyond rounding subject to cardinality constraints, where else can we obtain such guarantees, and where in algorithms, mechanism design, and statistics could they prove useful?

## 6 ACKNOWLEDGMENTS

## REFERENCES

[1] ACE Electoral Knowledge Network. 2022. *Electoral System (Chamber 1)*. https://aceproject.org/epic-en/CDTable?question=ES005&view=country&set_language=en

[2] John Adams. 1776. Letter to John Penn. https://founders.archives.gov/documents/Adams/06-04-02-0026-0003

[3] Michel L. Balinski and H. Peyton Young. 1978. Stability, Coalitions and Schisms in Proportional Representation Systems. *American Political Science Review* 72, 3 (1978), 848–858. https://doi.org/10.2307/1955106

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

[5] Julius Borcea, Petter Brändén, and Thomas Liggett. 2009. Negative Dependence and the Geometry of Polynomials. *Journal of the American Mathematical Society* 22, 2 (2009), 521–567. https://www.ams.org/jams/2009-22-02/S0894-0347-08-00618-8/

[6] Petter Brändén and Johan Jonasson. 2012. Negative Dependence in Sampling. *Scandinavian Journal of Statistics* 39, 4 (2012), 830–838. https://doi.org/10.1111/j.1467-9469.2011.00766.x

[7] K. R. W. Brewer and Muhammad Hanif. 1983. *An Introduction to Sampling with Unequal Probabilities*. Vol. 15. Springer, New York, NY. https://doi.org/10.1007/978-1-4684-9407-5_1

[8] Jaroslaw Byrka, Piotr Skowron, and Krzysztof Sornat. 2018. Proportional Approval Voting, Harmonic k-Median, and Negative Association. (2018), 14 pages, 566143 bytes. https://doi.org/10.4230/LIPICS.ICALP.2018.26

[9] Moses Charikar and Shi Li. 2012. A Dependent LP-Rounding Approach for the k-Median Problem. In *Automata, Languages, and Programming*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Artur Czumaj, Kurt Mehlhorn, Andrew Pitts, and Roger Wattenhofer (Eds.). Vol. 7391. Springer, Berlin, Heidelberg, 194–205. https://doi.org/10.1007/978-3-642-31594-7_17

---

[9]Such a utility is natural in a setting of proportional representation. For example, the approved parties could be those supporting a new policy the voter hopes to see adopted.

[10] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. 2010. Dependent Randomized Rounding via Exchange Properties of Combinatorial Structures. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*. 575–584. https://doi.org/10/ffbj8s

[11] Xiang-Hui Chen, Arthur P. Dempster, and Jun S. Liu. 1994. Weighted Finite Population Sampling to Maximize Entropy. *Biometrika* 81, 3 (1994), 457–469. https://doi.org/10.1093/biomet/81.3.457

[12] Yu Cheng, Zhihao Jiang, Kamesh Munagala, and Kangning Wang. 2020. Group Fairness in Committee Selection. *ACM Transactions on Economics and Computation* 8, 4 (2020), 1–18. https://doi.org/10.1145/3417750

[13] Jean-Claude Deville and Yves Tille. 1998. Unequal Probability Sampling without Replacement through a Splitting Method. *Biometrika* 85, 1 (1998), 89–101. https://doi.org/10.1093/biomet/85.1.89

[14] European Parliament. 2019. Rule of Procedure. https://www.europarl.europa.eu/doceo/document/RULES-9-2019-07-02_EN.pdf

[15] Rajiv Gandhi, Samir Khuller, Srinivasan Parthasarathy, and Aravind Srinivasan. 2006. Dependent Rounding and Its Applications to Approximation Algorithms. *Journal of the ACM (JACM)* 53, 3 (2006), 324–360. https://doi.org/10/fbbkqc

[16] German Federal Ministry of Justice. 2008. Gesetz zur Änderung des Wahl- und Abgeordnetenrechts vom 17. März 2008. http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl108010s0394.pdf

[17] Paul Gölz, Dominik Peters, and Ariel D. Procaccia. 2022. In This Apportionment Lottery, the House Always Wins. In *Proceedings of the ACM Conference on Economics and Computation (EC)*. https://doi.org/10.1145/3490486.3538299

[18] Anton Grafström. 2009. Non-Rejective Implementations of the Sampford Sampling Design. *Journal of Statistical Planning and Inference* 139, 6 (2009), 2111–2114. https://doi.org/10/cnbjjm

[19] Geoffrey Grimmett. 2004. Stochastic Apportionment. *The American Mathematical Monthly* 111, 4 (2004), 299–307. https://doi.org/10/gqcq6h

[20] Jaroslav Hájek. 1981. *Sampling from a Finite Population.* Number 37 in Statistics. Dekker, New York, NY.

[21] A. Hedayat, Bing-Ying Lin, and John Stufken. 1989. The Construction of ΠPS Sampling Designs Through a Method of Emptying Boxes. *The Annals of Statistics* 17, 4 (1989), 1886–1905. https://doi.org/10.1214/aos/1176347400

[22] D. G. Horvitz and D. J. Thompson. 1952. A Generalization of Sampling Without Replacement From a Finite Universe. *J. Amer. Statist. Assoc.* 47, 260 (1952), 663. https://doi.org/10/c97fzh

[23] William G. Madow. 1949. On the Theory of Systematic Sampling, II. *The Annals of Mathematical Statistics* 20, 3 (1949), 333–354. https://doi.org/10.1214/aoms/1177729988

[24] Roger B Myerson. 1981. Optimal auction design. *Mathematics of operations research* 6, 1 (1981), 58–73.

[25] Joseph Naor, Aravind Srinivasan, and David Wajc. 2023. Online Dependent Rounding Schemes. *arXiv:2301.08680 [cs.DS]* (2023).

[26] Alessandro Panconesi and Aravind Srinivasan. 1997. Randomized Distributed Edge Coloring via an Extension of the Chernoff–Hoeffding Bounds. *SIAM J. Comput.* 26, 2 (1997), 350–368. https://doi.org/10/dz9czg

[27] Yuval Peres, Mohit Singh, and Nisheeth K. Vishnoi. 2017. Random Walks in Polytopes and Negative Dependence. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*.

[28] E. Arthur Robinson and Daniel Ullman. 2010. *A Mathematical Look at Politics.* CRC Press.

[29] Michael R. Sampford. 1967. On Sampling without Replacement with Unequal Probabilities of Selection. *Biometrika* 54, 3-4 (1967), 499–513. https://doi.org/10/ccb3kz

[30] Aravind Srinivasan. 2001. Distributions on Level-Sets with Applications to Approximation Algorithms. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*. 588–597. https://doi.org/10.1109/SFCS.2001.959935

[31] George Szpiro. 2010. *Numbers Rule: The Vexing Mathematics of Democracy, from Plato to the Present.* Princeton University Press, Princeton, N.J.

[32] William Thomson. 2011. Fair Allocation Rules. In *Handbook of Social Choice and Welfare.* Vol. 2. Elsevier, 393–506.

[33] Yves Tillé. 2006. *Sampling Algorithms.* Springer, New York.

[34] Yves Tillé. 2023. Remarks on Some Misconceptions about Unequal Probability Sampling without Replacement. *Computer Science Review* 47 (2023), 100533. https://doi.org/10.1016/j.cosrev.2022.100533

[35] David Wajc. 2017. Negative Association - Definition, Properties, and Applications. (2017). https://www.cs.cmu.edu/~dwajc/notes/Negative%20Association.pdf

[36] F. Yates and P. M. Grundy. 1953. Selection Without Replacement from Within Strata with Probability Proportional to Size. *Journal of the Royal Statistical Society: Series B (Methodological)* 15, 2 (1953), 253–261. https://doi.org/10.1111/j.2517-6161.1953.tb00140.x

# APPENDIX

## A  RELATIONSHIP WITH PRIOR WORK

OBSERVATION A.1. *Neither Grimmett's method, the pipage method, the conditional Poisson method, nor the Sampford method, nor any apportionment method satisfying full support satisfies the house monotonicity axiom of Gölz et al. [17].*

PROOF. As Gölz et al. [17] show in Section 4.1 of their paper that, in a scenario with four parties, vote vector $v_1 = 1, v_2 = 2, v_3 = 1, v_4 = 2$, and house size 2, no method satisfying quota and house monotonicity may give both seats to parties 1 and 3 with positive probability. By definition, any method satisfying full support would do so, and thus cannot satisfy house monotonicity, which rules out the conditional Poisson and the Sampford method. Gölz et al. [17] show in the same section that Grimmett's method (with or without a randomized order) fails this test and thus fails house monotonicity. Finally, one easily verifies that the pipage method, when presented with the parties in the above worst-case order, also has a positive probability of awarding both seats to parties 1 and 3. When the order of parties is chosen uniformly at random, the order will with nonzero probability be the one above, showing again that the method fails house monotonicity. □

## B  SELECTION MONOTONICITY COUNTEREXAMPLES FOR SPECIFIC RULES

Here we prove Propositions 3.3 and 3.4 by way of counterexamples, showing that none of the rounding rules mentioned in this paper satisfy selection monotonicity except for Sampford sampling. We then compliment this result with Proposition B.1, which states that these rules also all violate pairwise selection monotonicity.

PROPOSITION 3.3. *Systematic rounding and pipage rounding violate selection monotonicity, even if the order of parties is uniformly shuffled.*

PROOF. We start by showing the statement for pipage rounding. For the sake of illustration, we start with a counterexample for a fixed order corresponding to the order of the indices. Let

$$\vec{p} = \left( \frac{1}{3}, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{1}{2} \right)$$

and $T = \{1, 3, 6\}$. We claim that

$$\mathbb{P}[S = T] > 0.$$

To see this, consider the following update sequence which happens with positive probability by the definition of pipage rounding:

| Step | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_6$ |
|------|-------|-------|-------|-------|-------|-------|
| 1 | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{2}{3}$ | $\frac{2}{3}$ | $\frac{1}{2}$ |
| 2 | $\frac{5}{6}$ | 0 | $\frac{1}{3}$ | $\frac{2}{3}$ | $\frac{2}{3}$ | $\frac{1}{2}$ |
| 3 | $\frac{1}{6}$ | 0 | 1 | $\frac{2}{3}$ | $\frac{2}{3}$ | $\frac{1}{2}$ |
| 4 | $\frac{5}{6}$ | 0 | 1 | 0 | $\frac{2}{3}$ | $\frac{1}{2}$ |
| 5 | 1 | 0 | 1 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| 6 | 1 | 0 | 1 | 0 | 0 | 1 |

Now, consider the new residue vector

$$\vec{p}' = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3} \right).$$

We claim that for this profile $\mathbb{P}[\{1,3\} \subseteq S] = 0$. To see this, first consider the case that $p_1$ is set to $\frac{2}{3}$ in the first step. Then, in the second step, either $p_1$ is fixed to 1 and $p_3$ to 0 or vice versa. Otherwise, $p_1$ is set to 0 already in the first step. Thus $\mathbb{P}[S = T] = 0$.

We also give the following counterexample that works for the version of pipage rounding in which we randomize over all possible orders. Let $T = \{2, 3, 4\}$ and

$$\vec{p} = \Big(0.07, 0.9, 0.57, 0.37, 0.99, 0.1\Big).$$

Computing the probabilities by a straight-forward brute-force implementation of the method yields

$$\mathbb{P}[S = T] \geq 0.0043.$$

However, decreasing the residue for party 1 and increasing it for party 2 as follows

$$\vec{p}' = \Big(0.06, 0.91, 0.57, 0.37, 0.99, 0.1\Big)$$

yields

$$\mathbb{P}[S = T] \leq 0.0042,$$

showing that also the random order version of pipage rounding violates selection monotonicity.

For systematic rounding with fixed order, the counterexample for Grimmett's apportionment method, given in Section 1 directly provides a counterexample when using the residues given in Table 1 as input for systematic rounding. For random order, a straightforward brute-force check shows that in the first election there exists no order for which the left-wing parties get all three additional seats with positive probability. □

PROPOSITION 3.4. *Conditional Poisson rounding violates selection monotonicity.*

PROOF. The following counterexample was discovered and verified with considerable computational effort, using symbolic and rational arithmetic to ensure there were no numerical errors. It involves such orders of magnitude that it is not detectable using standard Python packages with floating-point arithmetic. Consider the following pair of scaled working probabilities for $n = 6$ and $k = 3$:

$$\vec{\pi} := (9962000143517508584561395134859, 3320666714505969957773493640043,$$
$$3320666714505969957773493640435, 2324466700154429125337383510227586,$$
$$2324466700154429125337383510227686, 1660333357252963458777541885429371)$$
$$\vec{\pi}' := (99620001435175193801835755646020, 3320666714505968157722724388092,$$
$$3320666714505968157722724388092, 2324466700154429914176750514233600,$$
$$2324466700154429914176750514233600, 1660333357252962147206216649823732)$$

Note that, to get the true working probabilities, these must be scaled down by the following respective normalization constants, which do not fit on one line:

$$N = 99999999999999999999999999999999997683569039326925979$$
$$6111908263941097160679374150252278225009624541592641^{1/3}$$
$$N' = 99999999999999999999999999999999999031428737331175393$$
$$5635009441949575197038130365899121713973587636026881^{1/3}$$

One may verify that these correspond to a pair of residue vectors $\vec{p}$ and $\vec{p}'$ such that $p_1 \leq p_1'$, $p_2 \leq p_2'$, $p_3 \leq p_3'$, $p_4 \geq p_4'$, $p_5 \geq p_5'$, $p_6 \geq p_6'$, and yet

$$\frac{\pi_1 \pi_2 \pi_3}{N} \geq \frac{\pi_1' \pi_2' \pi_3'}{N'},$$

i.e., the set $\{1, 2, 3\}$ is more likely under $\vec{p}$ than $\vec{p}'$.                    $\square$

PROPOSITION B.1. *Systematic rounding, pipage rounding, and conditional Poisson rounding violate pairwise selection monotonicity.*

PROOF. For pipage rounding (with random order), consider the following vector of residues

$$\vec{p} = (0.03, 0.59, 0.07, 0.42, 0.47, 0.42),$$

summing up to $k = 2$. We consider the two coalitions $T_1 = \{1, 2\}$ and $T_2 = \{5, 6\}$. A straightforward implementation of pipage rounding (with randomized order) yields the following selection probabilities for the two coalitions:

$$\mathbb{P}[S = T_1] \geq 0.011 \text{ and } \mathbb{P}[S = T_2] \leq 0.12.$$

Now, consider the new vector of residues

$$\vec{p}' = (0.03, 0.59, 0.16, 0.33, 0.47, 0.42).$$

Note that, while residues for $T_1$ and $T_2$ remained unchanged, the residues of parties 3 and 4 did change. This leads to coalition $T_1$ losing selection probability and coalition $T_2$ gaining selection probability. Precisely,

$$\mathbb{P}[S = T_1] \leq 0.01 \text{ and } \mathbb{P}[S = T_2] \geq 0.1204,$$

which proves the violation of pairwise selection monotonicity.

For systematic rounding, we let $n = 6$ and $k = 2$, and suppose $\vec{p} = (0.1, 0.1, 0.2, 0.2, 0.5, 0.9)$. Let $T_1$ be the set of parties with residue 0.1 and $T_2$ be the set of parties with residue 0.2. Then one can check that, if we increase $0.5 \rightarrow 0.6$ and decrease $0.9 \rightarrow 0.8$ (without changing the total population or state quotas), the probability of selecting $T_1$ decreases (to zero) while the probability of selecting $T_2$ increases, which violates pairwise selection monotonicity.

Finally, for conditional Poisson rounding, we must resort to computational verification as in the proof of Proposition 3.4. We unfortunately do not have a closed form for the working probabilities, but the pair of residue profiles is

$$\vec{p} := (0.0618342562928861, 0.0207176796116814,$$
$$0.0207176796116814, 0.9933997806603289,$$
$$0.9933997806603289, 0.9099308231630933)$$

$$\vec{p}' := (0.0618342562928862, 0.0207176796116814,$$
$$0.0207176796116814, 0.9933997806603289,$$
$$0.9933997806603289, 0.9099308231630932)$$

Observe that the only residues that changed in going from $\vec{p}$ to $\vec{p}'$ were $p_1$, which increased, and $p_6$, which decreased. The unique working probabilities are algebraic numbers given by complicated expressions. Using Mathematica, we computed that the set $T_1 := \{1, 2, 3\}$ (which contains only residues that weakly increased) decreased in probability by roughly $5 \times 10^{-26}$ while $T_2 := \{3, 4, 5\}$ (which contains only residues that weakly decreased, because they all stayed the same) increased in probability by roughly $2 \times 10^{-18}$. Thus, pairwise selection monotonicity is violated.                    $\square$

## C   SUFFICIENT CONDITION FOR SELECTION MONOTONICITY

We call a rounding rule neutral, if reordering of the input vector $\vec{p}$ does not change the selection probabilities. Clearly, Sampford rounding is neutral. Under neutrality, we can assume without loss of generality that the set of parties of interest in the definition of selection monotonicity is $T = [k]$.

LEMMA C.1. *Let $r$ be a neutral rounding rule. If, for any $\vec{p} \in [0, 1)^n$ with $\sum_{1 \le i \le n} p_i = k$ for some integer $1 \le k \le n - 1$ and $p_1 > 0$ it holds that*

$$\left( \frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{P}_{S \sim r(\vec{p})}[S = [k]] \ge 0,$$

*then $r$ satisfies selection monotonicity.*

PROOF. Let $\vec{p}, \vec{p}' \in [0, 1)^n$ be two vectors summing up to the same integer $k$. Let $T$ be a set of $k$ parties such that $p_i' \ge p_i$ for all parties $i \in T$ and $p_i' \le p_i$ for all $i \notin T$. By neutrality we can reorder $\vec{p}$ and $\vec{p}'$ such that $T = [k]$ and do not change the selection probabilities of $T$. To prove that $r$ satisfies selection monotonicity, we aim to show that

$$\mathbb{P}_{S \sim r(\vec{p}')}[S = [k]] \ge \mathbb{P}_{S \sim r(\vec{p})}[S = [k]].$$

We prove the statement by induction over the number of indices in $[n]$ for which $p_i \ne p_i'$ holds. Since $\vec{p}$ and $\vec{p}'$ are normalized, they differ in at least two indices (unless they are equal). Thus, in our base case, we assume without loss of generality that $p_1' = p_1 + \epsilon$ and $p_n' = p_n - \epsilon$ for some $\epsilon \in (0, 1)$, where the specific form again follows from normalization. We define the following function, which linearly interpolates between $\vec{p}$ and $\vec{p}'$. That is, for all $\lambda \in [0, 1]$ we define

$$\vec{f}(\lambda) = \vec{p} + \lambda (\vec{p}' - \vec{p}).$$

Note that $f_1(\lambda) = p_1 + \lambda\epsilon$ and $f_n(\lambda) = p_n - \lambda\epsilon$ and $f_i(\lambda) = p_i = p_i'$ for all $i \in \{2, \dots, n - 1\}$. We get that

$$\frac{\partial}{\partial \lambda} \mathbb{P}_{S \sim r(\vec{f}(\lambda))}[S = [k]] = \frac{\partial}{\partial f_1} \mathbb{P}_{S \sim r(\vec{f})}[S = [k]] \frac{\partial f_1}{\partial \lambda} + \frac{\partial}{\partial f_n} \mathbb{P}_{S \sim r(\vec{f})}[S = [k]] \frac{\partial f_n}{\partial \lambda} \quad \text{(chain rule)}$$

$$= \epsilon \cdot \left( \frac{\partial}{\partial f_1} \mathbb{P}_{S \sim r(\vec{f})}[S = [k]] - \frac{\partial}{\partial f_n} \cdot \mathbb{P}_{S \sim r(\vec{f})}[S = [k]] \right)$$

$$\ge 0 \quad \text{(precondition of lemma)}$$

Thus, by the fundamental theorem of calculus and the fact that $\vec{f}(0) = \vec{p}$ and $\vec{f}(1) = \vec{p}'$, we get that

$$\mathbb{P}_{S \sim r(\vec{p}')}[S = [k]] - \mathbb{P}_{S \sim r(\vec{p})}[S = [k]] = \int_0^1 \frac{\partial}{\partial \lambda} \mathbb{P}_{S \sim r(\vec{f}(\lambda))}[S = [k]] d\lambda \ge 0.$$

For the induction step, consider the case when $\vec{p}$ and $\vec{p}'$ differ in more than two indices. Let $i$ be the index minimizing $|p_i' - p_i|$ among all $i \in [n]$ with $p_i' \ne p_i$ and assume without loss of generality that $p_i' > p_i$. Then, let $j \in [n]$ be such that $p_j < p_j'$. Let $\vec{w}$ be the vector with $w_\ell = p_\ell$ for all $\ell \in [n] \setminus \{i, j\}$, $w_i = p_i'$ and $w_j = p_j - (p_i' - p_i)$. Note that $\vec{w} \in [0, 1)^n$, $\vec{w}$ sums to $k$ and $\vec{p}$ and $\vec{w}$ differ in two indices, namely $w_i > p_i$ and $w_j < p_j$, where $i \in [k]$ and $j \in [n] \setminus [k]$ holds. Hence, using the same argument as in the base case, we can show that

$$\mathbb{P}_{S \sim r(\vec{w})}[S = [k]] \ge \mathbb{P}_{S \sim r(\vec{p})}[S = [k]].$$

Moreover, note that $\vec{w}$ and $\vec{p}'$ differ in one less index than $\vec{p}$ and $\vec{p}'$, and it also holds that $p_\ell' \ge w_\ell$ for all $\ell \in [k]$ and $p_\ell' \le w_\ell$ for all $\ell \in [n] \setminus [k]$. Thus, by induction hypothesis it follows that

$$\mathbb{P}_{S \sim r(\vec{p}')}[S = [k]] \ge \mathbb{P}_{S \sim r(\vec{w})}[S = [k]] \ge \mathbb{P}_{S \sim r(\vec{p})}[S = [k]],$$

which concludes the proof.                                                                          □

# D DEFERRED PROOFS

## D.1 Formula from Section 3

LEMMA 3.10. *Let $B$ be as in Lemma 3.8. Let $\hat{B} \subseteq \{2, 3, \ldots, n-1\}$ be distributed according to a Poisson trial as above, but omitting parties 1 and $n$. Then,*

$$\left(\tfrac{\partial}{\partial p_1} - \tfrac{\partial}{\partial p_n}\right) \mathbb{E}\big[\mathbb{1}\{|B| < k\}\cdot(k - |B|)\big] = (p_n - p_1)\,\mathbb{P}\big[|\hat{B}| = k - 1\big].$$

PROOF. We take the derivative for each $\ell \in [k-1]$ individually. That is,

$$\left(\tfrac{\partial}{\partial p_1} - \tfrac{\partial}{\partial p_n}\right) \mathbb{E}\big[\mathbb{1}\{|B| < k\}\cdot(k - |B|)\big]$$

$$= \sum_{0 \le \ell \le k-1} (k - \ell) \cdot \left(\tfrac{\partial}{\partial p_1} - \tfrac{\partial}{\partial p_n}\right) \mathbb{P}\big[|B| = \ell\big]$$

$$= \sum_{0 \le \ell \le k-1} (k - \ell) \cdot \left(\tfrac{\partial}{\partial p_1} - \tfrac{\partial}{\partial p_n}\right) \Big((1 - p_1)\,(1 - p_n)\,\mathbb{P}\big[|\hat{B}| = \ell\big]$$

$$+ \big(p_1\,(1 - p_n) + (1 - p_1)\,p_n\big)\,\mathbb{P}\big[|\hat{B}| = \ell - 1\big]$$

$$+ p_1\,p_n\,\mathbb{P}\big[|\hat{B}| = \ell - 2\big]\Big)$$

$$= \sum_{0 \le \ell \le k-1} (k - \ell) \big((p_n - p_1)\,\mathbb{P}\big[|\hat{B}| = \ell\big] + 2\,(p_1 - p_n)\,\mathbb{P}\big[|\hat{B}| = \ell - 1\big] + (p_n - p_1)\,\mathbb{P}\big[|\hat{B}| = \ell - 2\big]\big)$$

$$= (p_n - p_1) \Bigg( \sum_{0 \le \ell \le k-1} (k - \ell) \big(\mathbb{P}\big[|\hat{B}| = \ell\big] - \mathbb{P}\big[|\hat{B}| = \ell - 1\big]\big)$$

$$- \sum_{0 \le \ell \le k-1} (k - \ell) \big(\mathbb{P}\big[|\hat{B}| = \ell - 1\big] - \mathbb{P}\big[|\hat{B}| = \ell - 2\big]\big)\Bigg)$$

$$= (p_n - p_1) \Bigg( \sum_{0 \le \ell \le k-1} \mathbb{P}\big[|\hat{B}| = \ell\big] - \sum_{0 \le \ell \le k-2} \mathbb{P}\big[|\hat{B}| = \ell\big]\Bigg) \qquad \text{(telescoping sums, } \mathbb{P}\big[|\hat{B}| < 0\big] = 0\text{)}$$

$$= (p_n - p_1)\,\mathbb{P}\big[|\hat{B}| = k - 1\big],$$

which concludes the proof. $\qquad\qquad\square$

## D.2 Auxiliary Lemma for the Proof of Theorem 4.5

LEMMA D.1. *Let $a, b, c, d, e \in \mathbb{R}$ be such that $a, c, e \le 0$, $b, d \ge 0$, and $a + b + c + d + e = 0$. Then there exists $u_0 \in \mathbb{R}$ such that the following all hold:*

(1) $u_0 + a \le 0$
(2) $u_0 + a + b \ge 0$
(3) $u_0 + a + b + c \le 0$
(4) $u_0 + a + b + c + d \ge 0$

PROOF. There are two cases to consider. First suppose $b \le -c$. Then we let $u_0 := -a$, which satisfies all four inequalities:

(1) $u_0 + a = 0$
(2) $u_0 + a + b = b \ge 0$
(3) $u_0 + a + b + c = b + c \le -c + c = 0$
(4) $u_0 + a + b + c + d = b + c + d = -a - e \ge 0$

On the other hand, if $b \ge -c$, we let $u_0 := -a - b - c$, which then also satisfies all four inequalities:

(1) $u_0 + a = -b - c \leq -b + b = 0$
(2) $u_0 + a + b = -c \geq 0$
(3) $u_0 + a + b + c = 0$
(4) $u_0 + a + b + c + d = d \geq 0$

This concludes the proof. □

By letting $a := ||A'| - |A||$, $b := ||B'| - |B||$, and so on, this lemma implies that $B \subseteq B'$ and $D \subseteq D'$ in the proof of Theorem 4.5.