

Methods

Monotone Randomized Apportionment

José Correa,^a Paul Gözl,^b Ulrike Schmidt-Kraepelin,^c Jamie Tucker-Foltz,^d Victor Verdugo^{e,f,*}

^aDepartment of Industrial Engineering, Universidad de Chile, Santiago 8370458, Chile; ^bSchool of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14850; ^cDepartment of Mathematics and Computer Science, Eindhoven University of Technology, 5612 AE Eindhoven, Netherlands; ^dYale School of Management, Yale University, New Haven, Connecticut 06511; ^eInstitute for Mathematical and Computational Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile; ^fDepartment of Industrial and Systems Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile

*Corresponding author

Contact: correa@uchile.cl,  <https://orcid.org/0000-0002-3012-7622> (JC); mail@paulgoelz.de,  <https://orcid.org/0000-0002-8101-6818> (PG); u.schmidt.kraepelin@tue.nl,  <https://orcid.org/0000-0002-9213-7746> (US-K); j.tuckerfoltz@yale.edu,  <https://orcid.org/0000-0001-9174-3341> (JT-F); victor.verdugo@uc.cl,  <https://orcid.org/0000-0003-0817-7356> (VV)

Received: September 22, 2024

Revised: February 4, 2026

Accepted: May 12, 2026

Published Online in Articles in Advance:
July 9, 2026

Area of Review: Decision Analysis

<https://doi.org/10.1287/opre.2024.1362>

Copyright: © 2026 INFORMS

Abstract. 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 proven by Balinski and Young shows that no apportionment method can be proportional up to one seat (i.e., quota) while also responding monotonically to changes in the votes (i.e., population monotone). Grimmett 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 nonmonotonicities. When parties or voters care about joint events, such as whether a coalition of parties reaches a majority, these nonmonotonicities can cause paradoxes, including incentives for strategic voting. In this paper, we propose monotonicity axioms ruling out these paradoxes, and we study which of them can be satisfied jointly with Grimmett’s axioms. Essentially, we require that if a set of parties receives 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 owing to Sampford satisfies Grimmett’s axioms and a notion of higher-order correlation monotonicity.

Funding: This work was partially supported by the National Science Foundation [Grant DMS-1928930] and the Alfred P. Sloan Foundation [Grant G-2021-16778] while three of the authors were in residence at the Simons Laufer Mathematical Sciences Institute in Berkeley, California during the fall 2023 semester. Additionally, this work was partially supported by the National Science Foundation [Graduate Research Fellowship Program Grant DGE174] and the National Agency for Research and Development [Grant ACT210005], Center for Mathematical Modeling [Grant FB210005], and Fondo Nacional de Desarrollo Científico y Tecnológico [Grant 1241846].

Keywords: apportionment • randomized rounding • sampling

1. Introduction

Across modern democracies, a long-lasting ideal has been that the legislature should “be an exact portrait, in miniature, of the people” that it represents in the words of John Adams (Adams 1776). 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 differ; federal states divide up seats between the member states in proportion to state populations, whereas 85 countries worldwide divide up seats between political parties in proportion to their vote

shares (ACE Electoral Knowledge Network 2022). 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.

Although 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 of $h = 100$ seats, should it receive 8 or 9 seats or even some other number? Since the eighteenth 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 method should be used has led to intense debate, which often revolved around the methods' mathematical properties and actively involved mathematicians (Szpiro 2010).

Take, for example, *Hamilton's method*, the apportionment method put forward by the first U.S. Secretary of the Treasury Alexander Hamilton. His method first calculates for each party the number of seats that 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 nonmonotonicities 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, whereas party 2 gains one. Such monotonicity violations are not mere mathematical inconveniences but led to the abandonment of Hamilton's method for apportionment over U.S. states in 1901 (Balinski and Young 2001, p. 42) and for the apportionment over parties in Germany in 2008 (German Federal Ministry of Justice 2008). In fact, no apportionment method is immune to such criticism; as Balinski and Young (2001) showed, *any* method violates quota or population monotonicity.

In 2004, the mathematician Geoffrey Grimmett proposed a simple solution to this impasse: allowing apportionment methods to be random (Grimmett 2004). 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 method, first awards each party its lower quota and then, randomly awards some parties one more seat as follows. Let the parties' residues be $p_1, \dots, p_n \in [0, 1)$, 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 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 we 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* because 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 of "award lower quotas, then round" yields a vast space of randomized apportionment methods that satisfy quota and *ex ante proportionality*. To construct such a method, all we need is a *selection lottery* (i.e., a procedure that takes in residues $p_1, \dots, p_n \in [0, 1)$ summing to some integer k and randomly selects a subset $S \subseteq \{1, \dots, n\}$ of size k such that each $1 \leq i \leq k$ is contained in S with probability p_i). This problem was recently considered by Cembrano et al. (2026) and Gözl et al. (2026) with a specific focus on apportionment axioms, but it was also studied more broadly in mathematical statistics as *probability proportional to size (pps) sampling without replacement* (Brewer and Hanif 1983, Tillé 2006). In fact, the seminal book by Brewer and Hanif (1983) lists 50 such methods, although several only apply if $k = 2$ or give inclusion probabilities only approximately equal to the p_i . Among them, *systematic sampling* (Madow 1949) yields exactly Grimmett's method when applied as above. Hedayat et al. (1989) and Deville and Tillé (1998) even developed general frameworks for *pps sampling* that capture infinitely many selection lotteries and hence, randomized apportionment methods. In the past two decades, theoretical computer scientists have searched for the same kind of selection lotteries under the name of *dependent randomized rounding* (Gandhi et al. 2006), which have proven to be powerful tools in algorithm design (e.g., Srinivasan 2001, Charikar and Li 2012, Naor et al. 2025)—frequently reinventing rules previously known in statistics (Tillé 2023).

But, which randomized apportionment method out of this multitude should one choose given that all of these methods satisfy quota and *ex ante proportionality*?¹ Naturally, we define additional desirable axioms and focus on those methods with good axiomatic properties. In this work, we ask the following question. *Which monotonicity axioms are desirable in randomized apportionment? Additionally, 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 1,100 people, Apportia boasts as many as six political parties, three of which lean left and three of which lean 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 mid-

dle columns in Table 1. 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 from party 1, two from party 3, and none from party 5), and the right-wing parties controlled five seats, 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 that we order the parties numerically for Grimmett’s algorithm (see the upper panel of Figure 1), 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 . Because 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 of one. Arguably, this majority is deserved because 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 voters switch to party 4, and 10 voters switch to party 6. The rightmost columns in Table 1 show the new tallies. This shift leaves the lower quotas unchanged but changes the residues. As can be seen in the lower panel of Figure 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 parties to right-wing parties *increased* the chances of a left-wing government from 0 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.²

1.2. Our Approach and Results

Our first contribution is introducing a family of monotonicity axioms that can guide the choice and future development of randomized apportionment methods. Unlike the work of Gözl et al. (2026), 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 that 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.

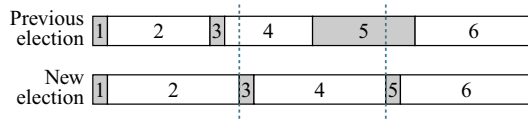
In Section 3, we begin our investigation in the domain of selection lotteries. We find that not only the systematic rounding underlying Grimmett’s method but also, pipage rounding and the selection lottery with maximum entropy (i.e., *conditional Poisson sampling*) violate simple forms of our monotonicity and hence, admit the paradox described in the previous subsection. One selection lottery from the statistics literature, however, bucks this trend; we show that *Sampford sampling* (Sampford 1967) satisfies the axiom of *selection monotonicity*.

Theorem 1 (Sampford Rounding Satisfies Selection Monotonicity). *Let $p, p' \in [0, 1]^n$ 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,*

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

Figure 1. Illustration of Grimmett’s Method (Grimmett 2004) for the Two Elections from Table 1



Notes. Intervals corresponding to left-wing parties are gray. Dashed lines indicate the positions of integers before the random shift.

the probability that Sampford sampling rounds up the set T is at least as high for p' as for 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. However, establishing this inequality is a considerable technical challenge because 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. We formally present Sampford’s method in Section 3, and we defer the full proof of Theorem 1 to Section 4.

Sampford sampling is also the only selection lottery among those that we consider that satisfies a two-coalition version of selection monotonicity, which is logically incomparable with 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* (Brändén and Jonasson 2012)), replacing pipage rounding with Sampford sampling adds our monotonicity axiom for free, which offers a new path for potential applications (e.g., in mechanism design) (see Section 4.1 for more details). Sampford sampling is the first and only selection lottery known to satisfy selection monotonicity, and there are no other obvious contenders. Regarding general results for this axiom, we show that it implies that the selection lottery is Lipschitz continuous.

In Section 5, 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. However, we conjecture that it is indeed satisfied by Sampford sampling. For the restricted case where coalitions consist of only two parties, we show that Grimmett’s method satisfies threshold monotonicity.

1.3. Related Work

Deterministic apportionment has been the subject of a deep axiomatic treatment, much of it developed by Balinski and Young (1978, 2001) (see Balinski and Young 2001 and Robinson and Ullman 2010). Two of the most prominent axioms in this literature, like the ones we develop here, are forms of monotonicity that rule out paradoxes (Balinski and Young 2001, chapter 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 (Balinski and Young 2001, theorem 6.1) and is satisfied exactly by the widely used family of divisor methods (Balinski and Young 2001, theorem 4.2). Divisor methods capture the idea of proportionality by *scaling and rounding* (i.e., proportionality followed by a rounding rule). They offer an alternative to the quota property, which is incompatible with population monotonicity. However, as we show in our work, the quota property can be reconciled with stochastic monotonicity notions provided that randomization is allowed.

Coalitions have mainly been considered in this literature in the context of whether apportionment methods encourage multiple parties to merge into a single party (Balinski and Young 1978); 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 pair-wise 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 π ps sampling without replacement above. A central motivation for this literature is the *Horvitz–Thompson estimator* (Horvitz and Thompson 1952), which we illustrate in a toy example (loosely following Sampford 1967). Suppose we have farms numbered $1, \dots, n$, and suppose each farm i has a known size of land x_i and an unknown yield y_i . Following Horvitz and Thompson (1952), 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 $\sum_{1 \leq i \leq n} x_i/k$.³ The variance of this estimate, which depends on how homogeneous the yield per acre is across farms, can be

estimated from the same sample (Yates and Grundy 1953) as long as the joint sampling probabilities of pairs of farms are known. As a result, the π ps literature scrupulously studied pair-wise inclusion probabilities (e.g., Brewer and Hanif 1983) 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 4.1, typical applications of dependent rounding lean heavily on concentration bounds. In 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* (Panconesi and Srinivasan 1997) over *negative association* (Wajc 2017) to the *strong Rayleigh property* (Borcea et al. 2009). The axioms that we pursue fundamentally differ from those above in that they are *relational* rather than *punctual* (Thomson 2011) (i.e., in that they constrain how the selection lottery's distributions relate across different residue vectors rather than just imposing separate constraints for each input). Whereas we only consider dependent rounding subject to a cardinality constraint (Srinivasan 2001, Gandhi et al. 2006), dependent rounding schemes have since been developed for more general combinatorial constraints, including bipartite matchings (Gandhi et al. 2006) and matroids (Chekuri et al. 2010).

In their work, Gözl et al. (2026), like us, aim to refine the space of randomized apportionment methods by imposing monotonicity axioms in addition to quota and ex ante proportionality. The monotonicity axioms that they consider, however, differ from ours in that they are defined ex post; 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; i.e., they consider the usual monotonicity axioms per realization of the random apportionment method). These axioms can be interpreted as preventing paradoxes but only in artificial scenarios in which the votes or house sizes change *after the randomization of the apportionment has been performed*. Gözl et al. (2026) show that population monotonicity is incompatible with quota, but house monotonicity can be guaranteed alongside quota and ex ante proportionality. In Appendix A, we show that the house monotonicity property of Gözl et al. (2026) is unlikely to be compatible with our axioms; no apportionment methods satisfying full support and none of the methods that we consider in this paper satisfy this property. Finally, we mention the work by Aziz et al. (2019) in the context of strategy-proof peer selection, which also provides an ex ante proportional

randomized apportionment routine based on awarding lower quotas and rounding.

2. Preliminaries

Throughout this paper, let $[n] = \{1, \dots, n\}$ denote a set of *parties*. For a set S and natural number k , we set $\binom{S}{k} = \{T \subseteq S : |T| = k\}$. An *apportionment method* a is a function that takes in the number of *votes* $v_1, \dots, v_n \in \mathbb{R}_{\geq 0}$ for each party and house size h and returns a vector-valued random variable $a(v, h) \in \mathbb{N}^n$ summing to h , which lists the number of seats awarded to each party. For a fixed vector of votes v and a house size h , we denote party i 's *standard quota* (i.e., its proportionally 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 (2004).

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

Ex ante proportionality. For any v, h , the expected number of seats awarded to i equals q_i .

Similarly, a *selection lottery* r is a function that takes in a vector of *target probabilities* $p \in [0, 1]^n$ for each party, which must sum up to an integer k , and returns a set-valued random variable $r(p) \in \binom{[n]}{k}$ (i.e., the values taken on by the random variable are sets of k parties). In addition, r must satisfy for each p and $i \in [n]$ that i 's probability of being included in $r(p)$ equals its target probability p_i . We are overloading p_i by using it both for general target probabilities and for residues induced by a vote profile v . We do so as they are mathematically equivalent objects and will play the same role in Section 3 on selection lotteries and Section 5 on apportionment rules. We say that a selection lottery is *continuous* if it is 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 selection lotteries and vice versa. On the one hand, we can interpret any (quota and ex ante proportional) apportionment method a as a selection lottery; for given target probabilities p summing to $k \in \mathbb{N}$, $a(p, k)$ is a random variable ranging over integer vectors adding to k . Quota ensures that this vector is binary and thus, can be interpreted as a random set of k parties. Ex ante proportionality ensures that the target probabilities are met. Conversely, any selection lottery r induces an *apportionment method* as we have described in Section 1. Given the votes v and house size h , one computes the standard quotas q_i and residues p_i and 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 $r(p)$, one obtains an apportionment method satisfying quota and ex ante proportionality. We define axioms for selection

lotteries and apportionment methods directly in Sections 3 and 5.

3. Monotonicity of Selection Lotteries

We begin our investigation of monotonicity in the subsetting of selection lotteries 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 5 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 π ps sampling and have potential implications for dependent randomized rounding in computer science, such as we explore in Section 4.1.

3.1. Axioms and Selection Lotteries

We rule out the paradox from Section 1 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 1 (Selection Monotonicity). Let $p, p' \in [0, 1]^n$ 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 selection lottery r satisfies selection monotonicity if it always holds that

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

As a warm-up to our proof, we will also consider a variant that 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 2 (Pair-Wise Selection Monotonicity). Let $p, p' \in [0, 1]^n$ be two residue vectors summing up to the same integer k . Let T_1, T_2 be two sets of k parties such that $p'_i \geq p_i$ for all parties $i \in T_1$ and $p'_i \leq p_i$ for all $i \in T_2$. A selection lottery r satisfies pair-wise selection monotonicity if it always holds that $\mathbb{P}_{S \sim r(p')} [S = T_1] \geq \mathbb{P}_{S \sim r(p)} [S = T_1]$ or $\mathbb{P}_{S \sim r(p')} [S = T_2] \leq \mathbb{P}_{S \sim r(p)} [S = T_2]$.

Note that pair-wise selection monotonicity is not implied by (and does not imply) selection monotonicity

because the target probabilities of parties outside of T_1 and T_2 are allowed to vary arbitrarily. We now introduce several selection lotteries from the literature on π ps sampling.

Systematic rounding (Madow 1949). We have already introduced this rule in Section 1 in the context of Grimmett's randomized apportionment method (Grimmett 2004). A random shift $u \sim \text{Uniform}([0, 1])$ 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 (Dewille and Tillé 1998, Srinivasan 2001).⁴ Although 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 exists, select the k parties with target probability of one.

Conditional Poisson rounding (Chen et al. 1994). This selection lottery 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* $\pi_1, \dots, \pi_n \in \mathbb{R}_{\geq 0}$ 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 (Tillé 2006).

Sampford rounding (Sampford 1967). This selection lottery 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 $i_2, \dots, i_k$ with replacement from $[n]$, choosing each party i with probability proportional to $p_i/(1 - p_i)$. If the k drawn parties are distinct, select them; else, start over. Sampford's selection lottery can also be implemented in a nonrejective, polynomial-time manner (Grafström 2009).

It is worth dwelling on how *strange* (or in the words of Tillé 2006, “very ingenious”) Sampford's selection lottery is. If, say, all k units were drawn with probability proportional to $p_i/(1 - p_i)$, each set A of k parties would be drawn with probability proportional to $\prod_{i \in A} (p_i/(1 - p_i))$, which means that this would just be conditional Poisson rounding but for incorrect working probabilities $\pi_i = p_i/(1 - p_i)$. 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).

Although it is not a selection lottery in our model, our analysis will frequently refer to *Poisson sampling* (not to be confused with *conditional Poisson sampling*).

Poisson sampling is not a selection lottery because it samples sets of nondeterministic size; it simply performs for each party I an independent Bernoulli trial with success probability p_i , and it contains i in its output if this trial is a success. Clearly, the probability of sampling a set A is just $\prod_{i \in A} p_i \prod_{i \notin A} (1 - p_i)$.

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 (Hájek 1981, theorem 7.3), they are not exactly met.⁵

Even though these *almost selection lotteries* easily satisfy selection monotonicity, most proper selection lotteries fail the axiom. Our example from Section 1 already proved that systematic rounding (even with a randomized ordering of parties) fails selection monotonicity, and the following proposition offers a similar counterexample for pipage rounding; we defer its proof to Appendix B.

Proposition 1. *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 selection lotteries 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 a few k -subsets of parties. Sharing none of these properties and being the “most natural” selection lottery 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 $p_1, \dots, p_6$, which is negative exactly for the vectors of target probabilities that would certify a violation of selection monotonicity. 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 2. *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 selection lottery. 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 let party 4's target decrease to 1%. Because 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

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

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

In our analysis, we will be working with the following expression, which we obtain by reformulating the probability-mass term:

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

where the second equality expands the fraction with $\prod_{i \in [n]} (1 - p_i)$ and where

$$f(A) = \sum_{i \in A} (1 - p_i) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j). \quad (3)$$

We introduced the function $A \mapsto f(A)$ because it will appear extensively throughout our analysis.

3.3.1. Warm-up: Pair-Wise Selection Monotonicity. As an easy warm-up and as an indication that Sampford

rounding is indeed well suited for monotonicity, we first show that it satisfies the pair-wise variant of selection monotonicity.

Theorem 2. *Sampford rounding satisfies pair-wise selection monotonicity.*

Proof. Fix the target probabilities p, p' , the sample size k , and the sets T_1, T_2 . Considering the numerator of the probability-mass term in Equation (1), it is easy to see that the numerator of $\mathbb{P}[S = T_1]$ weakly increases when going from p to 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 p to p' or not. If it increases, then because the numerator for T_2 decreases, the probability of sampling T_2 must decrease, and we are done. Else, because the numerator for T_1 increases, the probability of sampling T_1 increases, concluding the proof. $\square$

Whereas Sampford rounding satisfies pair-wise selection monotonicity quite handily, we show in Appendix B that the three other selection lotteries violate this axiom. Key to the previous proof is that we barely had to engage with the denominator term, a sum over exponentially many terms that indicates the probability of nonrejection for the sampling procedure. Engaging with how the probability of nonrejection changes with the target probabilities will be the major obstacle in proving our main result: that Sampford satisfies selection monotonicity.

3.3.2. Ex Ante Proportionality. Although the full proof of Theorem 1 is deferred to Section 4, we give here a new proof for the ex ante proportionality of Sampford rounding. Our approach is mainly combinatorial, providing an alternative to the original proof given by Sampford (1967). It also serves to introduce some of the ideas that will be useful to show Theorem 1. We want to show that for every party ℓ , the probability that S contains ℓ equals p_ℓ ; that is,

$$\frac{\sum_{A \in \binom{[n]}{k}: \ell \in A} f(A)}{\sum_{B \in \binom{[n]}{k}} f(B)} = p_\ell.$$

We start by simply rearranging the previous expression as follows:

$$\begin{aligned} & \frac{\sum_{A \in \binom{[n]}{k}: \ell \in A} f(A)}{\sum_{B \in \binom{[n]}{k}} f(B)} \\ &= \frac{\sum_{A \in \binom{[n]}{k}: \ell \in A} f(A)}{\sum_{B \in \binom{[n]}{k}: \ell \in B} f(B) + \sum_{B \in \binom{[n]}{k}: \ell \notin B} f(B)} \\ &= \frac{Q}{Q + Q'}, \end{aligned} \quad (4)$$

where $Q = \sum_{A \in \binom{[n]}{k}: \ell \in A} f(A)$ and $Q' = \sum_{B \in \binom{[n]}{k}: \ell \notin B} f(B)$. Next, we show that $Q' = (1 - p_\ell)Q/p_\ell$, which is

sufficient to conclude ex ante proportionality because then, $Q/(Q + Q') = p_\ell$. To show the previous identity between Q and Q' , we use the following combinatorial identity that we summarize as the following observation.

Observation 1. For every $U \subseteq [n]$ and $k \geq 1$, we have

$$\begin{aligned} & \sum_{B \in \binom{U}{k}} \sum_{i \in B} (1 - p_i) \prod_{j \in B} p_j \prod_{j \in U \setminus B} (1 - p_j) \\ &= \sum_{C \in \binom{U}{k-1}} \sum_{i \in U \setminus C} p_i \prod_{j \in C} p_j \prod_{j \in U \setminus C} (1 - p_j). \end{aligned}$$

Before arguing how to derive the previous combinatorial identity, we show how to finish the ex ante proportionality using Observation 1:

$$\begin{aligned} Q' &= \sum_{B \in \binom{[n]}{k}: \ell \notin B} f(B) \\ &= \sum_{B \in \binom{[n]}{k}: \ell \notin B} \sum_{i \in B} (1 - p_i) \prod_{j \in B} p_j \prod_{j \in [n] \setminus B} (1 - p_j) \\ &= (1 - p_\ell) \sum_{B \in \binom{[n]}{k}: \ell \notin B} \sum_{i \in B} (1 - p_i) \prod_{j \in B} p_j \prod_{j \in ([n] \setminus \{\ell\}) \setminus B} (1 - p_j) \\ &= (1 - p_\ell) \sum_{B \in \binom{[n] \setminus \{\ell\}}{k}} \sum_{i \in B} (1 - p_i) \prod_{j \in B} p_j \prod_{j \in ([n] \setminus \{\ell\}) \setminus B} (1 - p_j) \\ &= (1 - p_\ell) \sum_{C \in \binom{[n] \setminus \{\ell\}}{k-1}} \sum_{i \in ([n] \setminus \{\ell\}) \setminus C} p_i \prod_{j \in C} p_j \prod_{j \in ([n] \setminus \{\ell\}) \setminus C} (1 - p_j) \\ &= \frac{1 - p_\ell}{p_\ell} \cdot Q, \end{aligned}$$

where the second equality holds by replacing with the definition of $f(B)$, the third equality holds by taking out the $(1 - p_\ell)$ factor from the last product, and the fifth equality is a consequence of Observation 1 applied with $U = [n] \setminus \{\ell\}$.

Coming back to Observation 1, the identity holds by rearranging the sum and matching terms in an appropriate way. Indeed, observe that

$$\begin{aligned} & \sum_{B \in \binom{U}{k}} \sum_{i \in B} (1 - p_i) \prod_{j \in B} p_j \prod_{j \in U \setminus B} (1 - p_j) \\ &= \sum_{B \in \binom{U}{k}} \sum_{i \in B} p_i \prod_{j \in B \setminus \{i\}} p_j \prod_{j \in U \setminus (B \setminus \{i\})} (1 - p_j). \\ &= \sum_{C \in \binom{U}{k-1}} \sum_{i \in U \setminus C} p_i \prod_{j \in C} p_j \prod_{j \in U \setminus C} (1 - p_j), \end{aligned}$$

where the first equality holds by taking p_i out of the first product and by sending $(1 - p_i)$ to the last product and the second equality follows by applying a bijection from the set $\{(B, i) : B \in \binom{U}{k}, i \in B\}$ to the set $\{(C, i) : C \in \binom{U}{k-1}, i \in U \setminus C\}$, which maps (A, i) to $(A \setminus \{i\}, i)$. We remark

that Observation 1 will be used in our monotonicity analysis of Section 4.

3.4. 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. We defer its proof to Appendix D.2.

Theorem 3. *Any selection monotone selection lottery is Lipschitz continuous. Specifically, let p, p' be two residue profiles, and let T be a set of k parties. Then, $|\mathbb{P}[r(p') = T] - \mathbb{P}[r(p) = T]| \leq \|p - p'\|_1$.*

4. Sampford Rounding Is Monotone

In the remainder of this section, we will prove Theorem 1 (i.e., that Sampford rounding is selection monotone). By symmetry, we will assume that $A = [k]$. A simple argument shows that it suffices to consider the special case where the target probability of one party in A (without loss of generality (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. Because Sampford's probability-mass function is differentiable, proving the following proposition will be sufficient to conclude Theorem 1 (in Appendix C, we present a formal proof for this reduction).

Proposition 3. *Let $p \in [0, 1]^n$ such that $\sum_{i \in [n]} p_i = k$ for some integer $k \in \{1, \dots, n-1\}$, $p_1 > 0$, and $S = r_{\text{Sampford}}(p)$.*

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

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 one. 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 zero.

Clearly, to bound the partial derivative from Proposition 3, we will eventually need to apply the quotient rule, which requires us to understand and bound the numerator and denominator in Equation (2) as well as their derivatives. Although 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 2, 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 3–5 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. For the remainder of the section (including all lemma statements), we fix $p \in [0, 1]^n$ as defined in Proposition 3. To prove our key Lemma 2, we require the following auxiliary lemma.

Lemma 1. *For all $\ell \in \{0, 1, \dots, n-1\}$, 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 have

$$\begin{aligned} \sum_{A \in \binom{[n]}{\ell+1}} f(A) &= \sum_{A \in \binom{[n]}{\ell+1}} \sum_{i \in A} (1 - p_i) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \\ &= \sum_{A \in \binom{[n]}{\ell}} \sum_{i \notin A} p_i \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \\ &= \sum_{A \in \binom{[n]}{\ell}} \left(k - \ell + \ell - \sum_{i \in A} p_i \right) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \\ &= \sum_{A \in \binom{[n]}{\ell}} (k - \ell) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \\ &\quad + \sum_{A \in \binom{[n]}{\ell}} \left(\sum_{i \in A} 1 - p_i \right) \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \\ &= (k - \ell) \sum_{A \in \binom{[n]}{\ell}} \prod_{j \in A} p_j \prod_{j \notin A} (1 - p_j) \\ &\quad + \sum_{A \in \binom{[n]}{\ell}} f(A), \end{aligned}$$

where the second equality holds from Observation 1, and in the third equality, we make use that $\sum_{i \in [n]} p_i = k$. This concludes the proof. $\square$

Lemma 1 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 toward the mean.

Lemma 2. *Let $B \subseteq [n]$ be a random variable distributed according to a Poisson trial (i.e., B contains each $i \in [n]$ independently with probability p_i). Then,*

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

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

$$\begin{aligned}
& \sum_{A \in \binom{[n]}{k}} f(A) \\
&= \sum_{A \in \binom{[n]}{k}} f(A) - \sum_{A \in \binom{[n]}{0}} f(A) \quad (\text{second term is zero}) \\
&= \sum_{0 \leq \ell \leq k-1} \left(\sum_{A \in \binom{[n]}{\ell+1}} f(A) - \sum_{A \in \binom{[n]}{\ell}} f(A) \right) \\
&= \sum_{0 \leq \ell \leq k-1} (k-\ell) \cdot \sum_{A \in \binom{[n]}{\ell}} \prod_{j \in A} p_j \prod_{j \notin A} (1-p_j) \quad (\text{by Lemma 1}) \\
&= \sum_{0 \leq \ell \leq k-1} (k-\ell) \cdot \mathbb{P}[|B| = \ell] \\
&= \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)]
\end{aligned}$$

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 $\mathbb{E}[|B|] = \sum_{i \in [n]} p_i = k$ yields that $\mathbb{E}[|B| - k] = 0$. Because $\mathbb{E}[|B| - k] = \mathbb{E}[\mathbf{1}\{|B| - k \geq 0\} \cdot (|B| - k)] + \mathbb{E}[\mathbf{1}\{|B| - k < 0\} \cdot (|B| - k)]$, this implies that $\mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] = -\mathbb{E}[\mathbf{1}\{|B| - k < 0\} \cdot (|B| - k)] = \mathbb{E}[\mathbf{1}\{|B| - k \geq 0\} \cdot (|B| - k)]$. The desired equality follows by observing that

$$\begin{aligned}
\mathbb{E}[|B| - k] &= \mathbb{E}[\mathbf{1}\{|B| - k \geq 0\} \cdot (|B| - k)] \\
&\quad + \mathbb{E}[\mathbf{1}\{|B| - k < 0\} \cdot (k - |B|)] \\
&= 2\mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)],
\end{aligned}$$

which concludes the proof. $\square$

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. Let $B \subseteq [n]$ be a random variable distributed according to a Poisson trial (i.e., B contains each $i \in [n]$ independently with probability p_i). Then, $\mathbb{E}[|B| - k] \leq 2s$.

Proof. We start by upper bounding the relevant term by the variance of $|B|$. That is, $\mathbb{E}[|B| - k] \leq \mathbb{E}[(|B| - k)^2] = \text{Var}(|B|)$, 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, $\text{Var}(|B|) = \sum_{i \in [n]} p_i(1-p_i) \leq \sum_{i \in [k]} (1-p_i) + \sum_{i \notin [k]} p_i = 2s$, where the last equality follows from $s = \sum_{i \in [k]} (1-p_i) = \sum_{i \notin [k]} p_i$. $\square$

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 ℓ of B that are relevant in the term of the expectation (i.e., $\ell \in \{0, 1, \dots, k-1\}$) and shows that all terms besides one term for $\ell = k-1$ cancel out.

Lemma 4. Let $B \subseteq [n]$ be a random variable distributed according to a Poisson trial (i.e., B contains each $i \in [n]$ independently with probability p_i). Let $\hat{B} \subseteq \{2, 3, \dots, n-1\}$ be distributed according to a Poisson trial as above but omitting parties 1 and n . Then,

$$\begin{aligned}
& \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\
&= (p_n - p_1) \mathbb{P}[|\hat{B}| = k-1].
\end{aligned}$$

The last step before turning toward proving Proposition 3 is to provide a lower bound for the derivative of the denominator, which we just calculated in Lemma 4. 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 $\{2, \dots, k\}$, 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 zero). However, we will only need this bound in the extreme case that $s < p_1$.

Lemma 5. Let $\hat{B}$ be as in Lemma 4. Then, $\mathbb{P}[|\hat{B}| = k-1] \geq \frac{1-2s}{p_1(1-p_n)}$.

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

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

We lower bound $\mathbb{P}[B = [k]]$ 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, $\mathbb{P}[B = [k]] \geq 1 - \sum_{1 \leq j \leq k} (1-p_j) - \sum_{k+1 \leq j \leq n} p_j = 1 - 2s$ because $s = \sum_{i \in [k]} (1-p_i) = \sum_{i \notin [k]} p_i$. The claim follows by combining both inequalities. $\square$

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, we apply the quotient rule, and finally, we distinguish the two cases $s \geq p_1$ and $s < p_1$. The former case will require little additional effort, whereas the latter requires the application of Lemma 3 and Lemma 5.

Proof of Proposition 3. Recall that we want to bound the partial derivatives of the ratio 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 $f([k])$:

$$\begin{aligned} \frac{\partial}{\partial p_1} f([k]) &= \frac{\partial}{\partial p_1} \sum_{i \in [k] \setminus 1} (1 - p_i) \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]} (1 - p_j) \\ &\quad + \frac{\partial}{\partial p_1} 1 \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]} (1 - p_j) \\ &\quad - \frac{\partial}{\partial p_1} p_1 \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]} (1 - p_j) \\ &= \sum_{i \in [k] \setminus 1} (1 - p_i) \cdot \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]} (1 - p_j) \\ &\quad + \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]} (1 - p_j) \\ &\quad - 2p_1 \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]} (1 - p_j) \\ &= \sum_{i \in [k]} (1 - p_i) \cdot \prod_{j \in [k] \setminus 1} p_j \prod_{j \notin [k]} (1 - p_j) \\ &\quad - \prod_{j \in [k]} p_j \prod_{j \notin [k]} (1 - p_j) \\ &= \left(\frac{1}{p_1} - \frac{1}{\sum_{i \in [k]} (1 - p_i)} \right) f([k]) = \left(\frac{1}{p_1} - \frac{1}{s} \right) f([k]). \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{\partial}{\partial p_n} f([k]) &= - \sum_{i \in [k]} (1 - p_i) \cdot \prod_{j \in [k]} p_j \prod_{j \notin [k]} (1 - p_j) \\ &= - \frac{f([k])}{1 - p_n}. \end{aligned} \quad (6)$$

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

$$\begin{aligned} &\mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)]^2 \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \frac{f([k])}{\sum_{A' \in \binom{[n]}{k}} f(A')} \\ &= \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)]^2 \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \frac{f([k])}{\mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)]} \quad (\text{by Lemma 2}) \\ &= \left(\left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) f([k]) \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &\quad - f([k]) \cdot \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &= \left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n} \right) f([k]) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &\quad - f([k]) (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1] \\ &= f([k]) \left(\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \right. \\ &\quad \left. - (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1] \right) \\ &\quad (\text{by Equations (5) and (6) and Lemma 4}) \end{aligned}$$

Thus, it suffices to show the nonnegativity of the factor

$$\begin{aligned} &\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &\quad - (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1]. \end{aligned} \quad (7)$$

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 seems more common because the residues of all parties outside of $[k]$ outweigh that of the single party 1—is straightforward. The hard case is when $s < p_1$ (i.e., when the residues in $[k]$ are almost one and the residues outside $[k]$ are almost zero).

Case $s \geq p_1$. Under this assumption, clearly $1/p_1 - 1/s \geq 0$, which allows us to ignore these terms in Equation (7). Next, we observe that $\mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \geq \mathbb{P}[|B| = k - 1]$. Moreover, because one way for the number of successes $|B|$ in a Poisson trial to be equal to $k - 1$ is if there are $k - 1$ successes among the Bernoulli variables indexed $2, \dots, n - 1$ and additionally, the Bernoulli variables indexed one and n both fail, it must hold that $\mathbb{P}[|B| = k - 1] \geq (1 - p_1)(1 - p_n) \mathbb{P}[|\hat{B}| = k - 1]$. These two inequalities together imply that $\mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] / (1 - p_n) \geq (1 - p_1) \mathbb{P}[|\hat{B}| = k - 1] \geq (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1]$, which shows the claimed nonnegativity of the term in Equation (7).

Case $s < p_1$. Note that $p_1 > s = \sum_{i \notin [k]} p_i \geq p_n$. Thus, the term $(p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1]$ that we subtract in Equation (7) 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, Equation (7) 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 that we have previously derived:

$$\begin{aligned} &\left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &\quad - (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1] \\ &\geq \left(\frac{1}{p_1} - \frac{1}{s} + \frac{1}{1 - p_n} \right) s - (p_n - p_1) \frac{1 - 2s}{p_1(1 - p_n)} \quad (\text{by Lemmas 2, 3, and 5}) \\ &= \frac{1}{p_1(1 - p_n)} ((1 - p_n)s - p_1(1 - p_n) + p_1s \\ &\quad - (p_n - p_1)(1 - 2s)) \\ &= \frac{1}{p_1(1 - p_n)} (s - p_1s + p_1p_n - p_n + p_ns) \\ &\geq \frac{1}{p_1(1 - p_n)} (p_n(p_1 - 1 + s)) \\ &\geq \frac{p_n}{p_1(1 - p_n)} (p_1 - 1 + (1 - p_1)) = 0 \quad (\text{note } s - sp_1 \geq 0) \end{aligned}$$

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. $\square$

Because we have established (informally in this section and formally in Appendix C) that Proposition 3 implies selection monotonicity, the above proof implies our main result, which is Theorem 1.

4.1. 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 (Srinivasan 2001)—are equally satisfied by Sampford rounding (Brändén and Jonasson 2012). Hence, in all algorithms that use dependent rounding (subject to a cardinality constraint) as a black-box subroutine (e.g., Charikar and Li 2012, Byrka et al. 2018, Cheng et al. 2020), swapping out pipage rounding for Sampford rounding would preserve the algorithm's guarantees while gaining Sampford's monotonicity properties. Although we have yet to explore leveraging 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 (2012) 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 1) might allow the design of 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 proven to be influential in other applications (e.g., Steiner tree problems (Srinivasan 2001), k -median (Charikar and Li 2012), committee selection (Cheng et al. 2020), and online algorithms (Naor et al. 2025)).

Concretely, consider an approximation algorithm that consists of the two steps mentioned above; (1) determine a point x inside P , and 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 overall monotonicity; if agent i increases their report, X_i will be more likely to be sampled. Hence, by charging the agents Myerson payments (Myerson 1981), 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 5.2), this blueprint extends to richer utilities; for example, agent i could care about a subset T_i of matroid elements and have a monotone nondecreasing function $f_i : \mathbb{N} \rightarrow \mathbb{R}$ that expresses diminishing returns or returns to scale, and the agent's utility would be $\tau_i \cdot \mathbb{E}[f_i(|X \cap T_i|)]$.

5. Monotonicity of Apportionment Methods

Returning to the apportionment setting (i.e., where the input is given by vote counts $v = (v_1, \dots, v_n)$ and the house size H), what does the (pair-wise) selection monotonicity of Sampford rounding win us? Fix an “old” election and a “new” election with respective votes v, 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. 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 Section 1, 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 zero to one. Because this shift also decreases the sum of all residues from three to two, no apportionment method 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., European Parliament 2019, rules 33 and 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 we show that Grimmett's method satisfies it for coalitions of size $T = 2$.

5.1. Axioms and Apportionment Methods

The apportionment methods that we study are induced by the selection lotteries 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*, 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 (Grimmett 2004). Recall that all of these apportionment methods satisfy quota and ex ante proportionality. The main way that we strengthen monotonicity for apportionment is *threshold monotonicity*.

Definition 3 (Threshold Monotonicity). Let $v, v' \in \mathbb{R}_{\geq 0}^n$ 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 $i \in T$ and $q'_i \leq q_i$ for all $i \notin T$.⁶ An apportionment method a satisfies threshold monotonicity if it always holds for all thresholds $\vartheta \in \mathbb{N}$ that

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

In other words, the axiom requires the random variable $\sum_{i \in T} a(v', h)_i$, which describes the total number of seats

awarded to T for the votes v' , to first-order stochastically dominate the corresponding random variable for the votes v . We also define a pair-wise analog to the axiom.

Definition 4 (Pair-Wise Threshold Monotonicity). Let $v, v' \in \mathbb{R}_{\geq 0}^n$ 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 pair-wise threshold monotonicity if at least one of the following two statements always holds; for all $\vartheta \in \mathbb{N}$, $\mathbb{P}_{\alpha \sim a(v', h)} [\sum_{i \in T_1} \alpha_i \geq \vartheta] \geq \mathbb{P}_{\alpha \sim a(v, h)} [\sum_{i \in T_1} \alpha_i \geq \vartheta]$, or for all $\vartheta \in \mathbb{N}$, $\mathbb{P}_{\alpha \sim a(v', h)} [\sum_{i \in T_2} \alpha_i \geq \vartheta] \leq \mathbb{P}_{\alpha \sim a(v, h)} [\sum_{i \in T_2} \alpha_i \geq \vartheta]$.

Although these axioms (for apportionment methods) cannot syntactically be compared with our previous monotonicity axioms (for selection lotteries), they should be thought of as a strengthening of their selection equivalents. This relationship can be made formal using the embedding from Section 2. Any apportionment method satisfying (pair-wise) threshold monotonicity, when translated into a selection lottery, satisfies (pair-wise) selection monotonicity by selecting the threshold so that all parties in the coalition must be selected to meet the threshold. We will show that pair-wise threshold monotonicity is incompatible with a natural condition on apportionment methods:

Definition 5 (Full Support). An apportionment method a satisfies the full support property if for all vote vectors v , house sizes h , and apportionment vectors α that satisfy quota, $a(v, h)$ will take on the value α with positive probability.

The definitions of Sampford rounding immediately imply that the Sampford method has full support. Similarly, the conditional Poisson method has full support because if the residue of any party i is nonzero, its working probability π_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 *pair-wise vote-count threshold monotonicity* like its nonvote-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 does not change the apportionment), (pair-wise) vote-count threshold monotonicity implies (pair-wise) threshold monotonicity.

5.2. Conjecture and 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 1. *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 ϑ 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.

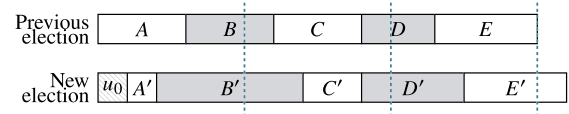
Still, we do have some limited good news. Although Grimmett’s method does not satisfy threshold monotonicity for coalitions of three and more parties (not even selection monotonicity as we saw in Section 1.1), it satisfies the axiom for coalitions of size 2.

Theorem 4. *Grimmett’s method satisfies threshold monotonicity for coalitions T of size ≤ 2 .*

Proof. We prove threshold monotonicity of the variant of Grimmett’s method that uses an arbitrary fixed ordering of the parties, 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; because extending the interval of each party i from length p_i to length $q_i = p_i + [q_i]$ adds an integral length $[q_i]$, the party’s interval will contain exactly $[q_i]$ more integers than previously, and the shift of other parties remains the same modulo 1. Second, we arbitrarily shift all of the intervals to the right by some amount u_0 . Again, this has no bearing on the probability distribution of the numbers of seats awarded because 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 (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

Figure 2. Illustration of the Proof of Theorem 4 Stating That Grimmett’s Method (Grimmett 2004) Satisfies Threshold Monotonicity for $|T| = 2$



Note. Intervals corresponding to growing parties are gray.

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 . Although we know that the sizes of the intervals corresponding to i and j increase when going from (v, h) to (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.2 in Section D.3 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 (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 (v, h) , the same value of u will award at least t seats to the parties in T when running $G(u_0)$ on (v', h) because the same t integers that were included in B and D will be contained within B' and D' . Because $G(0)$ and $G(u_0)$ are both equivalent to Grimmett’s method, we have shown that it satisfies threshold monotonicity. $\square$

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 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.

5.3. Impossibility of Pair-Wise Threshold Monotonicity

When we considered selection monotonicity and its pair-wise variant in Section 3, the most promising selection lotteries seemed to be 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 pair-wise variant of selection monotonicity.

In contrast, when moving to pair-wise 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 5. *No (ex ante proportional and quota) apportionment method that has full support can satisfy pair-wise 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

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

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 $\{1, \dots, 10\}$, we get

$$\begin{aligned} & \sum_{R \in \binom{\{1, \dots, 10\}}{2}} \mathbb{P}[S_m \supseteq R] \\ & \geq \mathbb{P}[|S_m \cap \{1, \dots, 10\}| \geq 2] \\ & = \mathbb{P}[|S_m \cap \{11, 12\}| \leq 1] \\ & = 1 - \mathbb{P}[S_m \supseteq \{11, 12\}] \\ & \geq 1 - \mathbb{P}[11 \in S_m] \\ & = 1 - \frac{1}{m} \geq \frac{1}{2}. \end{aligned}$$

Therefore, by the averaging principle, there must exist some $R \subseteq \{1, \dots, 10\}$ of size 2 such that $\mathbb{P}[S_m \supseteq R] \geq 1/(2 \cdot \binom{10}{2}) = 1/90$. Thus, for each $m \in \{2, 3, 4, \dots\}$, choose some subset $R_m \subseteq \{1, \dots, 10\}$ of size 2 such that $\mathbb{P}[S_m \supseteq R_m] \geq 1/90$. Because the sequence $R_2, R_3, R_4, \dots$ only takes finitely many different values, by the pigeonhole principle, there must be some infinite subsequence of indices $m_1, m_2, m_3, \dots$ such that R_{m_i} is the same for each positive integer i . Call this set T_2 , and choose an arbitrary set $T_1 \subseteq \{1, \dots, 10\} \setminus T_2$ of size 3. Without loss of generality, we relabel the parties so that $T_2 = \{1, 2\}$ and $T_1 = \{3, 4, 5\}$. Now, consider the alternative residue vector

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

and let S' be the random set that a rounds up on input 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 $(p', p(m_i))$ violates pair-wise 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 p' to $p(m_i)$, residues within T_2 have decreased from $\frac{1}{4}$ to $\frac{1}{5}$. However, because selecting both parties in T_2 entails excluding at least one of parties 11 and 12, we have

$$\begin{aligned} & \mathbb{P}[S' \supseteq T_2] \\ & \leq \mathbb{P}[11 \notin S' \text{ or } 12 \notin S'] \\ & \leq \mathbb{P}[11 \notin S'] + \mathbb{P}[12 \notin S'] \quad (\text{by the union bound}) \\ & = 2 - \mathbb{P}[11 \in S'] - \mathbb{P}[12 \in S'] \\ & = 2 - \left(1 - \frac{1}{200}\right) - \left(1 - \frac{1}{200}\right) \\ & = \frac{1}{100} < \frac{1}{90} \leq \mathbb{P}[S_{m_i} \supseteq R_{m_i}] = \mathbb{P}[S_{m_i} \supseteq T_2] \\ & \quad (\text{by proportionality}) \end{aligned}$$

so the probability of selecting all parties in T_2 has increased. Next, consider the set T_1 . In passing from p' to $p(m_i)$, residues in T_1 increased from $1/6$ to $1/5$, but

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

where in the final inequality, we have used the fact that selecting T_1 entails not selecting any other party as T_1 has size 3. Thus, the probability of selecting all parties in T_1 has decreased, and so, we have shown that pair-wise threshold monotonicity is violated. $\square$

Although this theorem only rules out methods having full support, we see that 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 pair-wise selection monotonicity implies that their induced methods fail pair-wise threshold monotonicity as described in Section 5.1.

We also remark that our proof leans heavily on the fact that pair-wise 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.

5.4. 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, although this latter perspective is arguably more prominent (e.g., Balinski and Young 2001, pp. 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 (pair-wise) 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 it often is 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 6. *No (ex ante proportional and quota) apportionment method can satisfy vote-count threshold monotonicity or pair-wise vote-count threshold monotonicity.*

Proof. Let a be any proportional apportionment method. We fix the house size at $h = 8$ and consider the initial vote vector $v = (380, 140, 140, 140)$. Let S be the random variable defined as the set of parties that a rounded up on input (v, h) . Observe that the total number of votes is 800, and thus, the lower quotas and residues are given by $\lfloor q \rfloor = (3, 1, 1, 1)$ and $p = (0.8, 0.4, 0.4, 0.4)$. 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 $0.2/3 = 0.0\overline{6}$. Because 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 $v' = (376, 142, 142, 100)$. The total population is now 760, and the lower quotas remain unchanged. However, the residues are now $p' = (0.96, 0.05, 0.49, 0.49)$ (rounded to two decimal places). Let S' be the random set of size 2 that a rounds up on input (v', h) . Because the residues still sum to two, rounding up both parties in T_1 precludes rounding up party 1, so $\mathbb{P}[S' = T_1] \leq \mathbb{P}[1 \notin S'] = 1 - \mathbb{P}[1 \in S'] = 1 - 0.96 = 0.04$ by proportionality. We claim that a violates pair-wise vote-count threshold monotonicity for the pair of vote vectors (v, 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 four seats if and only if it is rounded up because the lower quota

for party 1 (the only party in T_2) is three. In passing from v to 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 four seats if and only if both parties 2 and 3 are rounded up because the lower quotas for these parties are each one. In passing from v to 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 four seats has decreased. Thus, a violates pair-wise vote-count threshold monotonicity. Because $v'_i \leq v_i$ for all $i \notin T_1$, this also shows that a violates vote-count threshold monotonicity for T_1 . $\square$

Note that the proof requires $n = 4$ parties. This is tight because for $n = 3$, there is only one proportional selection lottery to use, and it is easy to see that it satisfies all monotonicity properties.

6. Discussion and Final Remarks

From our results, a landscape of monotonicity notions in randomized apportionment and rounding is taking shape. On one side are the (pair-wise) selection monotonicity axioms, which can be attained by Sampford sampling. On the other hand, we have the vote-count variants and pair-wise threshold monotonicity, which are overly strong axioms that cannot be satisfied (at least for the latter by natural apportionment methods). The next step in exploring monotonicity will focus on 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. Consider a voter who approves a subset of the parties in the sense that they want as many seats as possible to be filled from this subset.⁷ If 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 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 (Thomson 2011) 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? Where in algorithms, mechanism design, and statistics could they prove to be useful?

Acknowledgments

The authors are grateful to Ariel Procaccia and Dominik Peters for many productive discussions and useful ideas. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix A. Relationship with Prior Work

Observation A.1. Neither Grimmer’s method, the pipage method, the conditional Poisson method, the Sampford method, nor any apportionment method satisfying full support satisfy the house monotonicity axiom of Gözl et al. (2026).

Proof. As Gözl et al. (2026, section 4.1) show 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 method and the Sampford method. Gözl et al. (2026) show in the same section that Grimmer’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. $\square$

Appendix B. Selection Monotonicity Counterexamples for Specific Rules

Here, we prove Propositions 1 and 2 by way of counterexamples, showing that none of the selection lotteries mentioned in this paper satisfy selection monotonicity except for Sampford sampling. We then complement this result with Proposition B.1, which states that these rules also all violate pair-wise selection monotonicity.

Proposition 1. *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

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

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 (Table B.1).

Table B.1. Updates Sequence Obtained by 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

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

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 one and p_3 is fixed to zero or vice versa. Otherwise, p_1 is set to zero 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

$$p = (0.07, 0.9, 0.57, 0.37, 0.99, 0.1).$$

Computing the probabilities by a straightforward 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

$$(p' = (0.06, 0.91, 0.57, 0.37, 0.99, 0.1))$$

yields

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

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

For systematic rounding with fixed order, the counterexample for Grimmer’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. $\square$

Proposition 2. *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 that 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

$k = 3:$
$$\begin{aligned}\pi &= (99620001435175085845613951348591, \\ &\quad 33206667145059699577734936400435, \\ &\quad 33206667145059699577734936400435, \\ &\quad 23244667001544291253373835102276586, \\ &\quad 23244667001544291253373835102276586, \\ &\quad 1660333357252963458777541885429371) \\ \pi' &= (99620001435175193801835755646020, \\ &\quad 33206667145059681577227243883092, \\ &\quad 33206667145059681577227243883092, \\ &\quad 23244667001544299141767505142336500, \\ &\quad 23244667001544299141767505142336500, \\ &\quad 1660333357252962147206216649823732).\end{aligned}$$

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:

[illegible]

One may verify that these correspond to a pair of residue vectors p and p' such that $p_1 \leq p'_1$, $p_2 \leq p'_2$, $p_3 \leq p'_3$, $p_4 \geq p'_4$, $p_5 \geq p'_5$, $p_6 \geq p'_6$; yet,

$$\frac{\pi_1 \pi_2 \pi_3}{(N)^3} \geq \frac{\pi'_1 \pi'_2 \pi'_3}{(N')^3},$$

and that is, the set $\{1, 2, 3\}$ is more likely under p than p' . $\square$

Proposition B.1. *Systematic rounding, pipage rounding, and conditional Poisson rounding violate pair-wise selection monotonicity.*

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

$$p = (0.03, 0.59, 0.07, 0.42, 0.47, 0.42),$$

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:

$$p' = (0.03, 0.59, 0.16, 0.33, 0.47, 0.42).$$

Note that although 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 prove the violation of pair-wise selection monotonicity.

For systematic rounding, we let $n=6$ and $k=2$, and suppose $p=(0.1,0.1,0.2,0.2,0.5,0.9)$. 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 party quotas), the probability of selecting T_1 decreases (to zero), whereas the probability of selecting T_2 increases, which violates pair-wise selection monotonicity.

Finally, for conditional Poisson rounding, we must resort to computational verification as in the proof of Proposition 2. We unfortunately do not have a closed form for the working probabilities, but the pair of residue profiles is

$$\begin{aligned} p &:= (0.0618342562928861, 0.0207176796116814, \\ &\quad 0.0207176796116814, 0.9933997806603289, \\ &\quad 0.9933997806603289, 0.9099308231630933) \\ p' &:= (0.0618342562928862, 0.0207176796116814, \\ &\quad 0.0207176796116814, 0.9933997806603289, \\ &\quad 0.9933997806603289, 0.9099308231630932). \end{aligned}$$

Observe that the only residues that changed in going from p to 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 5×10^{-26} , whereas $T_2 := \{3, 4, 5\}$ (which contains only residues that weakly decreased because they all stayed the same) increased in probability by roughly 2×10^{-18} . Thus, pair-wise selection monotonicity is violated. $\square$

Appendix C. Sufficient Condition for Selection Monotonicity

We call a selection lottery neutral if reordering of the input vector 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 selection lottery. If for any $p \in [0, 1]^n$ with $\sum_{1 \leq i \leq n} p_i = k$ for some integer $k \in \{1, \dots, n-1\}$ and $p_1 > 0$, it holds that $\left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n}\right) \mathbb{P}_{S \sim r(p)}[S = [k]] \geq 0$, then r satisfies selection monotonicity.*

Proof. Let $p, p' \in [0, 1]^n$ be two vectors summing up to the same integer k . Let T be a set of k parties such that $p'_i \geq p_i$ for all parties $i \in T$ and $p'_i \leq p_i$ for all $i \notin T$. By neutrality, we can reorder p and 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 $\mathbb{P}_{S \sim r(p')} [S = [k]] \geq \mathbb{P}_{S \sim r(p)} [S = [k]]$. We prove the statement by induction over

the number of indices in $[n]$ for which $p_i \neq p'_i$ holds. Because p and 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 p and p' . That is, for all $\lambda \in [0, 1]$, we define $f(\lambda) = p + \lambda(p' - p)$. Note that $f_1(\lambda) = p_1 + \lambda\epsilon$, $f_n(\lambda) = p_n - \lambda\epsilon$, and $f_i(\lambda) = p_i = p'_i$ for all $i \in \{2, \dots, n-1\}$. We get that

$$\begin{aligned} & \frac{\partial}{\partial \lambda} \mathbb{P}_{S \sim r(f(\lambda))}[S = [k]] \quad (\text{C.1}) \\ &= \frac{\partial}{\partial f_1} \mathbb{P}_{S \sim r(f)}[S = [k]] \frac{\partial f_1}{\partial \lambda} + \frac{\partial}{\partial f_n} \mathbb{P}_{S \sim r(f)}[S = [k]] \frac{\partial f_n}{\partial \lambda} \\ & \quad (\text{chain rule}) \\ &= \epsilon \cdot \left(\frac{\partial}{\partial f_1} \mathbb{P}_{S \sim r(f)}[S = [k]] - \frac{\partial}{\partial f_n} \mathbb{P}_{S \sim r(f)}[S = [k]] \right) \\ & \geq 0 \quad (\text{precondition of the lemma}) \end{aligned}$$

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

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

For the induction step, consider the case when p and p' differ in more than two indices. Let i be the index minimizing $|p'_i - p_i|$ among all $i \in [n]$ with $p'_i \neq p_i$, and assume without loss of generality that $p'_i > p_i$. Then, let $j \in [n]$ be such that $p_j < p'_j$. Let 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 $w \in [0, 1]^n$, w sums to k and that p and w differ in two indices, namely $w_i > p_i$ and $w_j < p_j$, where $i \in [k]$ and $j \in [n] \setminus [k]$ hold. Hence, using the same argument as in the base case, we can show that

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

Moreover, note that w and p' differ in one less index than p and p' , and it also holds that $p'_\ell \geq w_\ell$ for all $\ell \in [k]$ and $p'_\ell \leq w_\ell$ for all $\ell \in [n] \setminus [k]$. Thus, by the induction hypothesis, it follows that

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

which concludes the proof. $\square$

Appendix D. Deferred Proofs

D.1. Proof of Lemma 4

Lemma 4. Let $B \subseteq [n]$ be a random variable distributed according to a Poisson trial (i.e., B contains each $i \in [n]$ independently with probability p_i). Let $\hat{B} \subseteq \{2, 3, \dots, n-1\}$ be distributed according to a Poisson trial as above but omitting parties 1 and n .

Then,

$$\begin{aligned} & \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &= (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1]. \end{aligned}$$

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

$$\begin{aligned} & \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{E}[\mathbf{1}\{|B| < k\} \cdot (k - |B|)] \\ &= \sum_{0 \leq \ell \leq k-1} (k - \ell) \cdot \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) \mathbb{P}[|B| = \ell] \\ &= \sum_{0 \leq \ell \leq k-1} (k - \ell) \cdot \left(\frac{\partial}{\partial p_1} - \frac{\partial}{\partial p_n} \right) ((1 - p_1)(1 - p_n) \mathbb{P}[|\hat{B}| = \ell] \\ & \quad + (p_1(1 - p_n) + (1 - p_1)p_n) \mathbb{P}[|\hat{B}| = \ell - 1] \\ & \quad + p_1 p_n \mathbb{P}[|\hat{B}| = \ell - 2]) \\ &= \sum_{0 \leq \ell \leq k-1} (k - \ell) ((p_n - p_1) \mathbb{P}[|\hat{B}| = \ell] + 2(p_1 - p_n) \mathbb{P}[|\hat{B}| = \ell - 1] \\ & \quad + (p_n - p_1) \mathbb{P}[|\hat{B}| = \ell - 2]) \\ &= (p_n - p_1) \left(\sum_{0 \leq \ell \leq k-1} (k - \ell) (\mathbb{P}[|\hat{B}| = \ell] - \mathbb{P}[|\hat{B}| = \ell - 1]) \right. \\ & \quad \left. - \sum_{0 \leq \ell \leq k-1} (k - \ell) (\mathbb{P}[|\hat{B}| = \ell - 1] - \mathbb{P}[|\hat{B}| = \ell - 2]) \right) \\ &= (p_n - p_1) \left(\sum_{0 \leq \ell \leq k-1} \mathbb{P}[|\hat{B}| = \ell] - \sum_{0 \leq \ell \leq k-2} \mathbb{P}[|\hat{B}| = \ell] \right) \\ &= (p_n - p_1) \mathbb{P}[|\hat{B}| = k - 1], \end{aligned}$$

(telescoping sums, $\mathbb{P}[|\hat{B}| < 0] = 0$)

which concludes the proof. $\square$

D.2. Proof of Theorem 3 As in the proof of Theorem 1, we first consider the special case when only two residues are changing at a time.

Lemma D.1. Consider a selection monotone selection lottery r , a residue profile p , and $0 < \delta < \min(p_2, 1 - p_1)$. Define a residue profile p' , in which $p'_1 = p_1 + \delta$, $p'_2 = p_2 - \delta$, and all other $p'_i = p_i$. Then, for any set T of k parties, it holds that

$$|\mathbb{P}[r(p') = T] - \mathbb{P}[r(p) = T]| \leq 2\delta. \quad (\text{D.1})$$

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

Case 1. Sets T such that $1 \in T$ and $2 \notin T$. By ex ante proportionality, $\sum_{T \in \binom{[n]}{k}, 1 \in T} \mathbb{P}[r(p) = T] = p_1$ and $\sum_{T \in \binom{[n]}{k}, 1 \in T} \mathbb{P}[r(p') = T] = p_1 + \delta$. For all T 's that contain one but not two, selection

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

$$\begin{aligned}
 \delta &= (p_1 + \delta) - p_1 \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T}} \mathbb{P}[r(p') = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T}} \mathbb{P}[r(p) = T] \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T \nexists 2}} (\mathbb{P}[r(p') = T] - \mathbb{P}[r(p) = T]) \\
 &\quad + \sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(p') = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(p) = T] \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T \nexists 2}} (\mathbb{P}[r(p') = T] - \mathbb{P}[r(p) = T]) \\
 &\quad + \mathbb{P}[r(p') \supseteq \{1,2\}] - \mathbb{P}[r(p) \supseteq \{1,2\}] \\
 &\geq \sum_{\substack{T \in \binom{[n]}{k} \\ 1 \in T \nexists 2}} (\mathbb{P}[r(p') = T] - \mathbb{P}[r(p) = T]) \geq 0.
 \end{aligned}$$

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

Case 2. Sets T such that $2 \in T$ and $T \nexists 1$. By the ex ante proportionality property, $\sum_{T \in \binom{[n]}{k}, 2 \in T} \mathbb{P}[r(p) = T] = p_2$ and $\sum_{T \in \binom{[n]}{k}, 2 \in T} \mathbb{P}[r(p') = T] = p_2 - \delta$. For all T that contain two but not one, selection monotonicity ensures that $\mathbb{P}[r(p') = T] \leq \mathbb{P}[r(p) = T]$. Hence,

$$\begin{aligned}
 \delta &= p_2 - (p_2 - \delta) \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T}} \mathbb{P}[r(p) = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T}} \mathbb{P}[r(p') = T] \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T \nexists 1}} (\mathbb{P}[r(p) = T] - \mathbb{P}[r(p') = T]) \\
 &\quad + \sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(p) = T] - \sum_{\substack{T \in \binom{[n]}{k} \\ T \supseteq \{1,2\}}} \mathbb{P}[r(p') = T], \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T \nexists 1}} (\mathbb{P}[r(p) = T] - \mathbb{P}[r(p') = T]) \\
 &\quad + \mathbb{P}[r(p) \supseteq \{1,2\}] - \mathbb{P}[r(p') \supseteq \{1,2\}], \\
 &= \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T \nexists 1}} (\mathbb{P}[r(p) = T] - \mathbb{P}[r(p') = T]) - \delta,
 \end{aligned}$$

which therefore, implies Equation (D.1) for this case because

$$2\delta \geq \sum_{\substack{T \in \binom{[n]}{k} \\ 2 \in T \nexists 1}} (\mathbb{P}[r(p) = T] - \mathbb{P}[r(p') = T]) \geq 0.$$

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 one and two (i.e., if $1, 2 \in T$, then choose a party in T , and if $1, 2 \notin T$, 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 δ is less than $1 - p_3$ (which is positive because residues are less than one by definition). Else (i.e., if δ is larger), we can split up the increase of p_1 by δ and the decrease of p_2 by δ 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 δ then follows by applying the triangle inequality to the bounds for the smaller steps.

Let p'' be a new residue profile, where $p''_1 = p_1 - \delta$, $p''_2 = p_2 + \delta = p'_2$, and $p''_3 = p_3 + \delta = p'_3 + \delta$. Note that this is a valid residue profile by how we chose ℓ . The key observation is that we already know that the bound in Equation (D.1) holds going from p to p'' and from p'' to p' because the residues changing involve one party in T and one party outside of T . Thus,

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

Naïvely, this would yield a bound of 4δ , 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 ≤ 0 , and the second is ≥ 0 ; however, if $1, 2 \in T$ and $3 \notin T$, then the first term is ≥ 0 , and the second term is ≤ 0 . Either way, the sum of these terms is at most 2δ , so Equation (D.1) holds. $\square$

Proof of Theorem 3. Let p, p' be an arbitrary pair of residue profiles. Define the following sets of parties: $S_+ = \{j \in [n] : p'_j > p_j\}$ and $S_- = \{j \in [n] : p'_j < p_j\}$. Consider the following procedure, which transforms p into p' .

1. $p^1 \leftarrow p$ and $i \leftarrow 1$.
2. **while** $p^i \neq p'$ **do**
3. $j_1 \leftarrow$ any element of S_+ such that $p'_{j_1} > p^i_{j_1}$
4. $j_2 \leftarrow$ any element of S_- such that $p'_{j_2} < p^i_{j_2}$
5. $\delta_i \leftarrow \min\{p'_{j_1} - p^i_{j_1}, p^i_{j_2} - p'_{j_2}\}$
6. $p^{i+1} \leftarrow$ copy of p^i except $p^{i+1}_{j_1} = p^i_{j_1} + \delta_i$ and $p^{i+1}_{j_2} = p^i_{j_2} - \delta_i$
7. $i \leftarrow i + 1$

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

$$\begin{aligned}
 \sum_{i=1}^{\ell-1} \delta_i &= \sum_{j \in S_+} (p'_j - p_j) \leq \max_{S \subseteq [n]} \sum_{j \in S} (p'_j - p_j) = d_{TV}(p, p') \\
 &= \frac{1}{2} \|p' - p\|_1,
 \end{aligned} \tag{D.2}$$

where we have applied the well-known characterization for the total variation distance d_{TV} .⁸

Note that Lemma D.1 applies to each consecutive pair of residue profiles p^i and p^{i+1} because they differ in only two

coordinates by δ_i . Hence, we have

$$\begin{aligned} & |\mathbb{P}[r(p') = T] - \mathbb{P}[r(p) = T]| \\ & \leq \sum_{i=1}^{\ell-1} |\mathbb{P}[r(p^{i+1}) = T] - \mathbb{P}[r(p^i) = T]| \quad (\text{triangle inequality}) \\ & \leq \sum_{i=1}^{\ell-1} 2\delta_i \quad (\text{from Lemma D.1}) \\ & \leq 2 \cdot \frac{1}{2} \|p - p'\|_1 \quad (\text{from Equation (D.2)}) \\ & = \|p - p'\|_1 \end{aligned}$$

which concludes the proof. $\square$

D.3. Auxiliary Lemma for the Proof of Theorem 4

Lemma D.2. Let $a, b, c, d, e \in \mathbb{R}$ be such that $a, c, e \leq 0$, $b, d \geq 0$, and $a + b + c + d + e = 0$. Then, there exists $u_0 \in \mathbb{R}$ such that the following all hold.

1. $u_0 + a \leq 0$
2. $u_0 + a + b \geq 0$
3. $u_0 + a + b + c \leq 0$
4. $u_0 + a + b + c + d \geq 0$

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

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

On the other hand, if $b \geq -c$, 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. $\square$

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.

Endnotes

¹ Grimmett (2004) himself offers “no justification for [his] scheme apart from fairness and ease of implementation.”

² 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 (2004) 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.

³ 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 unbiased, but it can have higher variance because it often misses large farms with high yields.

⁴ This method was first introduced in statistics as the *pivotal method* (Deville and Tillé 1998), but we use the term common in computer science.

⁵ 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.

⁶ Or, equivalently, we can require that i ’s share of the vote $v_i / \sum_{j \in [n]} v_j = q_i / h$ increases rather than its quota.

⁷ Such a utility is natural in a setting of proportional representation. For example, the approved parties could support a new policy that the voter hopes to see adopted.

⁸ This equality still holds even when probabilities sum to some constant other than one.

References

- ACE Electoral Knowledge Network (2022) Electoral system (chamber 1). Accessed June 30, 2026, <https://aceproject.org/epic-en>.
- Adams J (1776) Letter to John Penn. Accessed June 30, 2026, <https://founders.archives.gov/documents/Adams/06-04-02-0026-0003>.
- Aziz H, Lev O, Mattei N, Rosenschein JS, Walsh T (2019) Strategy-proof peer selection using randomization, partitioning, and apportionment. *Artificial Intelligence* 275:295–309.
- Balinski ML, Young HP (1978) Stability, coalitions and schisms in proportional representation systems. *Amer. Political Sci. Rev.* 72(3):848–858.
- Balinski ML, Young HP (2001) *Fair Representation: Meeting the Ideal of One Man, One Vote*, 2nd ed. (Brookings Institution Press, Washington, DC).
- Borcea J, Brändén P, Liggett TM (2009) Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.* 22(2):521–567.
- Brändén P, Jonasson J (2012) Negative dependence in sampling. *Scandinavian J. Statist.* 39(4):830–838.
- Brewer KRW, Hanif M (1983) *An Introduction to Sampling with Unequal Probabilities*, vol. 15 (Springer, New York).
- Byrka J, Skowron P, Sornat K (2018) Proportional approval voting, harmonic k -median, and negative association. *45th Internat. Colloquium Automata Languages Programming* (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Wadern, Germany).
- Cembrano J, Correa J, Griesbach SM, Verdugo V (2026) Online proportional apportionment. *Proc. 2026 Annual ACM-SIAM Sympos. Discrete Algorithms (SODA)* (SIAM, Philadelphia), 4846–4860.
- Charikar M, Li S (2012) A dependent LP-rounding approach for the k -median problem. *Proc. 39th Internat. Colloquium on Automata, Languages, and Programming (ICALP 2012)*, Lecture Notes in Computer Science, vol. 7391 (Springer, Berlin, Heidelberg), 194–205.
- Chekuri C, Vondrák J, Zenklus R (2010) Dependent randomized rounding via exchange properties of combinatorial structures. *Proc. 51st Annual IEEE Sympos. Foundations of Comput. Sci. (FOCS 2010)* (IEEE, Piscataway, NJ), 575–584.
- Chen XH, Dempster AP, Liu JS (1994) Weighted finite population sampling to maximize entropy. *Biometrika* 81(3):457–469.
- Cheng Y, Jiang Z, Munagala K, Wang K (2020) Group fairness in committee selection. *ACM Trans. Econom. Comput.* 8(4):1–18.
- Deville J-C, Tillé Y (1998) Unequal probability sampling without replacement through a splitting method. *Biometrika* 85(1):89–101.
- European Parliament (2019) Rule of procedure. Accessed June 30, 2026, https://www.europarl.europa.eu/doceo/document/RULES-9-2019-07-02_EN.pdf.
- Gandhi R, Khuller S, Parthasarathy S, Srinivasan A (2006) Dependent rounding and its applications to approximation algorithms. *J. ACM* 53(3):324–360.
- German Federal Ministry of Justice (2008) Gesetz zur Änderung des Wahl- und Abgeordnetenrechts vom 17. Accessed June 30, 2026, http://www.bgb1.de/xaver/bgb1/start.xav?startbk=Bundesanzeiger_BGB1&jumpTo=bgb1108010s0394.pdf.
- Gölz P, Peters D, Procaccia AD (2026) In this apportionment lottery, the house always wins. *Oper. Res.* 74(1):390–407.
- Grafström A (2009) Non-rejective implementations of the Sampford sampling design. *J. Statist. Planning Inference* 139(6):2111–2114.

- Grimmett G (2004) Stochastic apportionment. *Amer. Math. Monthly* 111(4):299–307.
- Hájek J (1981) *Sampling from a Finite Population*, Number 37 in Statistics (Dekker, New York).
- Hedayat A, Lin BY, Stufken J (1989) The construction of IIPS sampling designs through a method of emptying boxes. *Ann. Statist.* 17(4):1886–1905.
- Horvitz DG, Thompson DJ (1952) A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* 47(260):663–685.
- Madow WG (1949) On the theory of systematic sampling, II. *Ann. Math. Statist.* 20(3):333–354.
- Myerson RB (1981) Optimal auction design. *Math. Oper. Res.* 6(1):58–73.
- Naor J, Srinivasan A, Wajc D (2025) Online dependent rounding schemes for bipartite matchings, with applications. *Proc. 2025 Annual ACM-SIAM Sympos. Discrete Algorithms (SODA)* (SIAM, Philadelphia), 3116–3154.
- Panconesi A, Srinivasan A (1997) Randomized distributed edge coloring via an extension of the Chernoff–Hoeffding bounds. *SIAM J. Comput.* 26(2):350–368.
- Robinson EA, Ullman D (2010) *A Mathematical Look at Politics* (CRC Press, Boca Raton, FL).
- Sampford MR (1967) On sampling without replacement with unequal probabilities of selection. *Biometrika* 54(3–4):499–513.
- Srinivasan A (2001) Distributions on levelsets with applications to approximation algorithms. *Proc. 42nd IEEE Sympos. Foundations Comput. Sci.* (IEEE Computer Society, Washington, DC), 588–597.
- Szpiro G (2010) *Numbers Rule: The Vexing Mathematics of Democracy, from Plato to the Present* (Princeton University Press, Princeton, NJ).
- Thomson W (2011) Fair allocation rules. Arrow KJ, Sen A, Suzumura K, eds. *Handbook of Social Choice and Welfare*, vol. 2 (Elsevier, Amsterdam), 393–506.
- Tillé Y (2006) *Sampling Algorithms*, Springer Series in Statistics (Springer, New York).
- Tillé Y (2023) Remarks on some misconceptions about unequal probability sampling without replacement. *Comput. Sci. Rev.* 47:100533.
- Wajc D (2017) Negative association—Definition, properties, and applications. Accessed June 30, 2026, <https://www.cs.cmu.edu/~dwajc/notes/Negative%20Association.pdf>.
- Yates F, Grundy PM (1953) Selection without replacement from within strata with probability proportional to size. *J. Roy. Statist. Soc. Ser. B Methodological* 15(2):253–261.

José Correa is a professor in the Department of Industrial Engineering at Universidad de Chile. His research is in operations research, market design, algorithmic game theory, and mechanism design, with applications to transportation, pricing, and online decision making.

Paul Gözl is an assistant professor in the School of Operations Research and Information Engineering at Cornell University. His research uses tools from algorithms, optimization, artificial intelligence (AI), and game theory to study democratic decision making, fair resource allocation, computational social choice, and pluralistic AI alignment.

Ulrike Schmidt-Kraepelin is an assistant professor in the Algorithms Cluster at Eindhoven University of Technology. Her research is motivated by questions in social choice and uses tools from combinatorial optimization and game theory. Her work spans multiwinner elections, participatory budgeting, apportionment, liquid democracy, and fair division.

Jamie Tucker-Foltz is an assistant professor of operations at Yale School of Management with a secondary appointment in Computer Science. His research focuses on the design of fair political and economic institutions using tools from operations research, economics, and theoretical computer science. His work spans fair division, social choice, algorithmic game theory, and algorithms for fair redistricting and gerrymandering detection.

Victor Verdugo is an assistant professor at the Institute for Mathematical and Computational Engineering and the Department of Industrial and Systems Engineering at Pontificia Universidad Católica de Chile. His research lies at the intersection of algorithms, optimization, and economics, with a focus on online decision making, market design, discrete optimization, and fair representation.