Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 378 (2004) 71–91
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Surface volumes of rounding polytopes

Mathias Drton [a,*], Udo Schwingenschlögl [b]

[a]*Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195-4322, USA*
[b]*Institut für Physik, Universität Augsburg, 86135 Augsburg, Germany*

Received 24 June 2003; accepted 10 September 2003

Submitted by R.A. Brualdi

Abstract

The paper determines the vertices and surface volumes of all rounding polytopes for commonly used rounding methods: the quota method of greatest remainders, and the divisor methods. These methods are used to round continuous non-negative weights summing to one to non-negative integers summing to a predetermined accuracy, e.g. to 100 when rounding to percentages. Our results are of interest when average properties of rounding methods are investigated, and an example from political science is included.

Keywords: Apportionment method; Probability simplex; Proportional representation; Rounding method; Rounding rule; Seat allocation; Seat bias

1. Introduction

Consider a vector $w = (w_1, \dots, w_\ell)$ of $\ell \geqslant 2$ non-negative continuous weights that sum to one. These weights could, for example, be a set of probabilities. The rounding problem consists of rounding each weight w_i to a non-negative integer m_i such that the rounding result $m = (m_1, \dots, m_\ell)$ sums to a given integer accuracy M, i.e. the (continuous) weight w_i is approximated by the (rational) proportion m_i/M. It is well known that rounding the weights w_i individually may leave a discrepancy between the sum of the rounding results m_i and the desired accuracy M (cf. [12, Section 1]. However, such a discrepancy is often infeasible, and rounding methods

* Corresponding author. Tel.: +1-206-616-6592; fax: +1-206-685-7419.
E-mail addresses: drton@stat.washington.edu (M. Drton), udo.schwingenschloegl@physik.uni-augsburg.de (U. Schwingenschlögl).

0024-3795/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2003.09.005

are needed that yield rounding results summing to the predetermined accuracy M. An example is the apportionment of seats in a parliament with the fixed house site M, by rounding proportions of votes. Other examples can be found in Statistics [16,17].

This paper develops new mathematical insight into traditional rounding methods by characterizing the sets of weight vectors w that get rounded to a fixed integer vector m. These sets are polytopes, for all methods considered here. Since the weights are constrained to sum to one, the rounding polytopes are of dimension $\ell - 1$, where ℓ is the number of weights to be rounded. For a given rounding method, our main results determine the vertices and surface volumes of all rounding polytopes. Our work is based on the monographs by Balinski and Young [5], and by Kopfermann [13], as well as the original work by Pòlya [15].

Our results on the surface volumes of rounding polytopes are of importance for the comparison of different methods in terms of their average behavior. For such an average behavior it is common to assume uniformly distributed weights, so that the probability of a rounding polytope is proportional to its surface volume. For examples of papers dealing with uniformly distributed weights see [1–4,6,7,12,18,20].

The paper is organized as follows. In Section 2 we introduce the rounding methods dealt with in the sequel. In Sections 3 and 4 we derive our results on the vertices and surface volumes of rounding polytopes. In Section 5 we illustrate the use of our results in a political science application [18].

2. Rounding methods

Let the *probability simplex* S be the set of all non-negative weight vectors summing to one

$$S := \left\{ w \in [0,1]^\ell \;\middle|\; \sum_{i=1}^{\ell} w_i = 1 \right\}.$$

Rounding the weight vector $w \in S$ to a given integer accuracy M means that w is mapped to a vector of non-negative integers m with components summing to M. Hence a *rounding method* is a mapping $R : S \to G(M)$, where

$$G(M) := \left\{ m \in \mathbb{N}_0^\ell \;\middle|\; \sum_{i=1}^{\ell} m_i = M \right\}.$$

Throughout this paper, we will consider accuracies $M > \ell$. Details on the more pathological case $M \leqslant \ell$ can be found in [10].

The *quota method of greatest remainders* operates in two stages. First, the proportion $w_i M$ is rounded down to its integer part $\tilde{m}_i = \lfloor w_i M \rfloor$. In the (unlikely) case that all $w_i M$ are integers the discrepancy vanishes, i.e. $M - \sum_{i=1}^{\ell} \tilde{m}_i = 0$, and we set $m_i = \tilde{m}_i$. Otherwise, there is a positive discrepancy $\delta = M - \sum_i \tilde{m}_i \geqslant 1$, and the fractional parts $\delta_i = w_i M - \tilde{m}_i$ are ranked to obtain $\delta_{(1)} \geqslant \delta_{(2)} \geqslant \cdots \geqslant \delta_{(\ell)}$ (where ties are broken arbitrarily). The vector m is obtained by setting $m_{(i)} = \tilde{m}_{(i)} + 1$ for all $i \leqslant \delta$ and $m_{(i)} = \tilde{m}_{(i)}$ for all $i > \delta$. That is, the δ largest remainders are rounded up to one, the $\ell - \delta$ smallest remainders are rounded down to zero.

All other rounding methods considered are divisor methods. Following Balinski and Young [5, p. 99], the definition of a divisor method is based on a strictly isotonic sequence of reals such that $k \leqslant s(k) \leqslant k + 1$. This *sign-post sequence* $s = (s(k))_{k \geqslant 0}$ defines a *rounding function*

$$r : [0, \infty) \to \mathbb{N}_0, \quad x \mapsto r(x) := \begin{cases} k & \text{if } x \in [k, s(k)), \\ k + 1 & \text{if } x \in [s(k), k + 1). \end{cases}$$

(Ties $x = s(k)$ may be broken in a different way than setting $r(x) = k + 1$ without affecting our future results.) The *divisor method* with sign-post sequence s maps a weight vector w into the integer vector $R(w) = m = (m_1, \ldots, m_\ell) \in G(M)$ such that there exists a divisor $D \in (0, \infty)$ with $m_i = r(w_i/D)$ for all i.

Important sub-classes are the *q-stationary divisor methods* with parameter $q \in [0, 1]$ based on the sign-post sequences $s(k) = k + q$, and the *p-power mean divisor methods* with parameter $p \in \mathbb{R}$ based on $s(k) = [(k^p + (k + 1)^p)/2]^{1/p}$. There are five "traditional" divisor methods (cf. [5, p. 61]):

- Adams: $s(k) = k$ (rounding up, $q = 0$, $p = -\infty$),
- Dean: $s(k) = k(k + 1)/(k + 0.5)$ (harmonic rounding, $p = -1$),
- Hill/Huntington: $s(k) = \sqrt{k(k + 1)}$ (geometric rounding, $p = 0$),
- Webster/Sainte-Laguë: $s(k) = k + 0.5$ (standard rounding, $q = 0.5$, $p = 1$),
- Jefferson/d'Hondt: $s(k) = k + 1$ (rounding down, $q = 1$, $p = \infty$).

Marshall et al. [14] give a comparison of these five methods in terms of majorization. An implementation of divisor methods following Dorfleitner and Klein [9] is provided by the computer program BAZI.[1]

All methods presented map a weight vector with permuted entries to the permuted integer vector m, i.e. $R(w_{\sigma(1)}, \ldots, w_{\sigma(\ell)}) = (m_{\sigma(1)}, \ldots, m_{\sigma(\ell)})$ for any permutation σ. This property will be tacitly used in some of the subsequent proofs.

In the sequel we study the sets $\{w \in S \mid R(w) = m\}$ of weight vectors w that are rounded to a given integer vector $m \in G(M)$. For both the quota method of greatest remainders and the divisor methods ties were broken arbitrarily. For example, if $w = (0.5, 0.5)$ and $M = 3$ then the rounding results $m = (2, 1)$ and $\bar{m} = (1, 2)$ are possible. Thus we will consider the sets

$$P_R(m) := \mathrm{cl}\{w \in S \mid R(w) = m\}, \quad m \in G(M),$$

where cl denotes set closure. Then $P_R(m)$ contains all weight vectors that can be rounded to m under R if ties are broken arbitrarily.

[1] See `http://www.uni-augsburg.de/bazi`.

Lemma 2.1 states that the methods mentioned can be described by linear inequalities.

Lemma 2.1. *Let $m \in G(M)$ be a rounding result and let $w \in S$ be a weight vector.*

(a) *Let R be the quota method of greatest remainders. Then $w \in P_R(m)$ if and only if*

$$Mw_i - m_i \leqslant Mw_j - m_j + 1 \quad \forall i, j = 1, \ldots, \ell : i \neq j. \tag{1}$$

(b) *Let R be the divisor method with sign-post sequence s. Then $w \in P_R(m)$ if and only if*

$$w_i s(m_j - 1) \leqslant w_j s(m_i) \quad \forall i, j = 1, \ldots, \ell : i \neq j. \tag{2}$$

Proof. See [13, pp. 196, 202] and [5, p. 100]. □

The inequalities of Lemma 2.1 describe $P_R(m)$ as a polyhedron. Since $P_R(m) \subseteq S$ and S is bounded, $P_R(m)$ is a polytope. We call $P_R(m)$ the *rounding polytope* of the rounding result m under the rounding method R. Fig. 1 illustrates rounding polytopes for four methods, in the case of $\ell = 3$ weights and accuracy $M = 5$ in barycentric coordinates, i.e. a point w in one of the triangles represents the vector of the three shortest distances from w to each one of the three triangle edges. Note that a divisor method with $s(0) = 0$ rounds exclusively to interior lattice points; compare the case of rounding up in Fig. 1(b).

We characterize $P_R(m)$ in terms of its vertices. Then we compute the surface volume of $P_R(m)$. In the special case that R is a q-stationary divisor method and that $m_i \geqslant 1$ for all i, our results were already obtained by Kopfermann [13, Section 6.2]. The boundary cases, with $m_i = 0$ for some i, need particular attention, see Fig. 1. Going beyond the work of Kopfermann our considerations comprise all boundary cases for divisor methods as well as a full treatment of the quota method of greatest remainders.

In our (and Kopfermann's) approach to the computation of surface volumes, $P_R(m)$ is decomposed into simplices whose surface volumes can be computed via determinant formulas [19, p. 278]. Recall that a d-dimensional simplex is a d-dimensional polytope with $d + 1$ vertices $v_0, \ldots, v_d$, and by the determinant formula, the d-dimensional volume of this simplex equals $1/d!$ times the modulus of the determinant of the $d \times d$ matrix with columns $v_i - v_0$, $i = 1, \ldots, d$.

A surface volume is defined by means of full-dimensional volume after a projection (cf. [11, Section V.4]). Here, if

$$\pi : S \to \left\{ w \in [0, 1]^{\ell-1} \;\middle|\; \sum_{i=1}^{\ell-1} w_i \leqslant 1 \right\}, \quad w \mapsto (w_1, \ldots, w_{\ell-1}),$$

is the projection on the first $\ell - 1$ components, then for any measurable set $A \subseteq S$:

$$\mathrm{vol}_{\ell-1}(A) = \sqrt{\ell} \times \mathrm{vol}_{\ell-1}(\pi(A)). \tag{3}$$

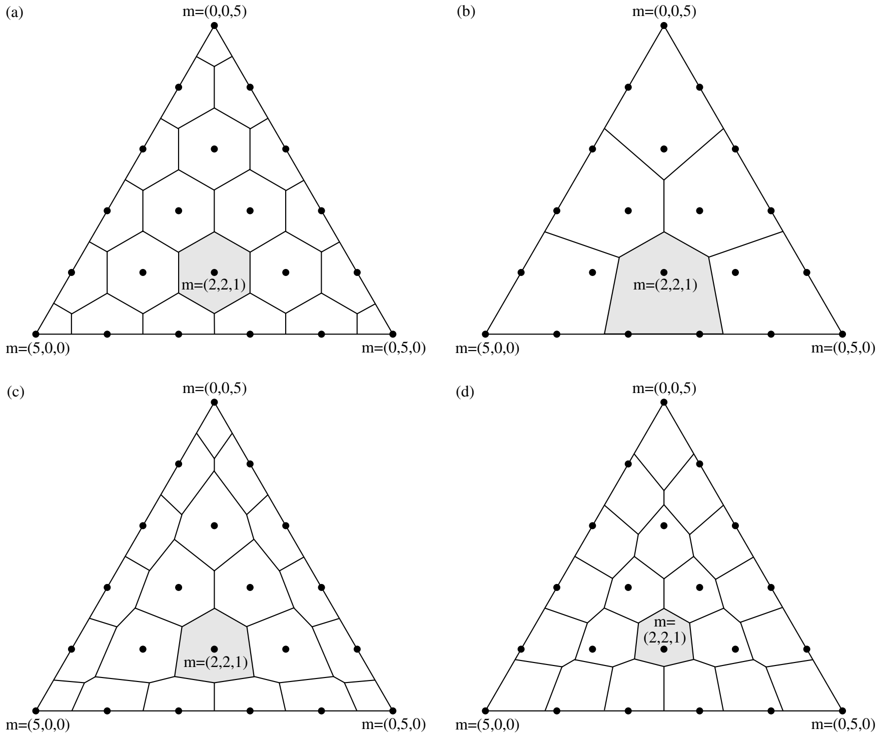

Fig. 1. Rounding polytopes for $\ell = 3$ weights and accuracy $M = 5$. The highlighted polytopes serve as examples in Sections 3.3 and 4.4. (a) Quota method of greatest remainders; (b) divisor method: rounding up; (c) divisor method: standard rounding; (d) divisor method: rounding down.

Note that the volume on the left hand side of (3) is a surface volume whereas the volume on the right hand side is full-dimensional. In particular, the simplex S has volume

$$\mathrm{vol}_{\ell-1}(S) = \frac{\sqrt{\ell}}{(\ell-1)!}. \tag{4}$$

Since in the following no confusion is possible we will refer to surface volumes simply as volumes.

3. Rounding polytopes for the quota method of greatest remainders

Let R be the quota method of greatest remainders and $m \in G(M)$ a possible rounding result. Let $N(m) = \{i \mid m_i = 0\}$ be the set of indices of zero components of m, and let $n(m) = |N(m)|$ be its cardinality. In Section 3.1 we study the vertices of $P(m) := P_R(m)$. The volume of $P(m)$ is calculated in Section 3.2. Section 3.3 illustrates our results.

3.1. Vertices

By Lemma 2.1, the translation $T : w \mapsto x = w - m/M$ maps any rounding polytope $P(m)$ that lies in the interior of S, i.e. for $n(m) = 0$, into the *standard polytope*

$$P_0 := \left\{ x \in \mathbb{R}^\ell \,\middle|\, \sum_{i=1}^{\ell} x_i = 0,\ x_i \leqslant x_j + 1/M\ \forall i \neq j \right\}. \tag{5}$$

If $m_i = 0$ then the constraint $w_i \geqslant 0$ remains invariant under T, i.e. it is translated into the constraint $x_i \geqslant 0$. Therefore, the rounding polytope $P(m)$ with $n(m) \geqslant 1$ is translated into the restricted standard polytope

$$P_0 \cap \bigcap_{i \in N(m)} \{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}. \tag{6}$$

In particular, $P(m)$ and $P(\tilde{m})$ are congruent whenever $n(m) = n(\tilde{m})$.

Theorem 3.1 yields the vertices of $P_0 \cap \bigcap_{i \in N(m)} \{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}$, and adding m/M yields the vertices of $P(m)$. We denote the row vectors in $\mathbb{R}^\ell$ with all components equal to 1 or 0 by 1_ℓ and 0_ℓ, respectively.

Theorem 3.1. *The polytope $P_0 \cap \bigcap_{i \in N(m)} \{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}$ has $2^\ell - 2^{n(m)} - 1$ vertices $v^{(\lambda)}$, which are induced by $\lambda \in \{0,1\}^\ell \setminus \{0_\ell, 1_\ell\}$ with $\lambda_j = 0$ for some index $j \notin N(m)$. The components of $v^{(\lambda)}$ are*

$$v_i^{(\lambda)} = \begin{cases} \frac{1}{M} \times \frac{\ell - z(\lambda) - e(\lambda)}{\ell - z(\lambda)} & \text{if } \lambda_i = 1, \\ \frac{1}{M} \times \frac{-e(\lambda)}{\ell - z(\lambda)} & \text{if } \lambda_i = 0 \text{ and } i \notin N(m), \\ 0 & \text{if } \lambda_i = 0 \text{ and } i \in N(m), \end{cases} \quad i = 1, \ldots, \ell, \tag{7}$$

where $z(\lambda) := |\{i \in N(m) \mid \lambda_i = 0\}|$ and $e(\lambda) := |\{1 \leqslant i \leqslant \ell \mid \lambda_i = 1\}|$.

If $n(m) = \ell - 1$, then $v^{(0_\ell)} := 0_\ell$ is also a vertex and the restricted standard polytope has $2^\ell - 2^{n(m)} = 2^{\ell-1}$ vertices.

There are no other vertices than the indicated $v^{(\lambda)}$.

In order to prove Theorem 3.1 we study the standard polytope P_0 from (5), and later the restricted standard polytope from (6). Lemma 3.2 provides a parallelotope decomposition of P_0. Lemma 3.3 gives the vertices of P_0.

Let the vector $u^{(i)} \in \mathbb{R}^\ell$ have component i equal to $(\ell - 1)/\ell$ and all other components equal to $-1/\ell$.

Lemma 3.2. *Define the parallelotopes*

$$L_i = \left\{ \sum_{j \neq i} \mu_j u^{(j)} : \mu_j \in [0,1] \right\} \subseteq \mathbb{R}^\ell, \quad i = 1, \ldots, \ell.$$

Then $\operatorname{int}(L_i)$ *and* $\operatorname{int}(L_j)$ *are disjoint if $i \neq j$, and $P_0 = \bigcup_{i=1}^{\ell} L_i$.*

Proof. A vector $x \in \mathrm{int}(L_i) \cap \mathrm{int}(L_j)$ can be expressed as

$$x = \sum_{k \neq i} \mu_k u^{(k)} = \sum_{k \neq j} \delta_k u^{(k)}$$

with all μ_k and δ_k positive. It follows that

$$x = x - \delta_i \sum_{k=1}^{\ell} u^{(k)} = \sum_{k \neq i,j} (\delta_k - \delta_i) u^{(k)} - \delta_i u^{(j)}.$$

Since $u^{(1)}, \dots, u^{(i-1)}, u^{(i+1)}, \dots, u^{(\ell)}$ form a basis of $\{x : \sum_i x_i = 0\}$, it follows that $\mu_j = -\delta_i$, which contradicts the fact that μ_j and δ_i are positive.

To see $\bigcup_{i=1}^{\ell} L_i \subseteq P_0$, let $x = \sum_{j=1}^{\ell} \mu_j u^{(j)}$ with $\mu_j \in [0, 1]$. For $p \neq q$,

$$\begin{aligned} x_p - x_q &= \sum_{j=1}^{\ell} \mu_j \big(u_p^{(j)} - u_q^{(j)}\big) \\ &= \mu_p \left(\frac{1}{M} \times \frac{\ell - 1}{\ell} - \frac{1}{M} \times \frac{-1}{\ell} \right) + \mu_q \left(\frac{1}{M} \times \frac{-1}{\ell} - \frac{1}{M} \times \frac{\ell - 1}{\ell} \right) \\ &= \frac{1}{M} \left(\mu_p - \mu_q \right) \leqslant \frac{1}{M}. \end{aligned} \tag{8}$$

Hence, $x \in P_0$.

Conversely, let $x \in P_0$. We need to show $x \in L_i$, for some i. Since every set of $\ell - 1$ vectors among $u^{(1)}, \dots, u^{(\ell)}$ forms a basis of $\{x \in \mathbb{R}^\ell : \sum_{i=1}^{\ell} x_i = 0\}$, we can write $x = \sum_{j=1}^{\ell-1} \mu_j u^{(j)}$. Since $\sum_{j=1}^{\ell} u^{(j)} = 0_\ell$, we can write $x = \sum_{j=1}^{\ell} \mu_j u^{(j)}$ with $\mu_j \geqslant 0$ for all j and $\mu_i = 0$ for some i. Using (8) we obtain, for $j \neq q$,

$$1/M \geqslant (x_j - x_q)/M \overset{(8)}{=} (\mu_j - \mu_q)/M \geqslant \mu_j / M.$$

Thus $\mu_j \leqslant 1$ for all j which implies $x \in L_i$. □

Lemma 3.3. *Every $\lambda \in \{0, 1\}^\ell \setminus \{0_\ell, 1_\ell\}$ induces a vertex $u^{(\lambda)}$ of the standard polytope P_0 through*

$$u_i^{(\lambda)} = \begin{cases} \frac{1}{M} \times \frac{\ell - e(\lambda)}{\ell} & \text{if } \lambda_i = 1, \\ \frac{1}{M} \times \frac{-e(\lambda)}{\ell} & \text{if } \lambda_i = 0, \end{cases} \qquad i = 1, \dots, \ell,$$

with $e(\lambda) := |\{1 \leqslant i \leqslant \ell \mid \lambda_i = 1\}|$. There are no other vertices.

Proof. Obviously $u^{(\lambda)} = \sum_{i=1}^{\ell} \lambda_i u^{(i)}$, which yields that

$$u^{(\lambda)} \in \bigcap_{i : \lambda_i = 0} L_i \subset P_0.$$

The $u^{(\lambda)}$ with $e(\lambda) = 1$ are in fact the $u^{(j)}$ in the definition of the parallelotopes L_i. Due to symmetry with respect to permutations it suffices to concentrate on $u^{(\lambda)}$ with the first $e(\lambda)$ components equal to 1. Such a $u^{(\lambda)}$ solves

$$A_\ell u^{(\lambda)} = \begin{array}{c} \begin{array}{ccccccc} 1 & & \cdots & e(\lambda) & e(\lambda)+1 & \cdots & \ell \end{array} \\ \begin{pmatrix} 1 & & & & -1 & & \\ \vdots & & & & & \ddots & \\ 1 & & & & & & -1 \\ & 1 & & & -1 & & \\ & & \ddots & & \vdots & & \\ & & & 1 & -1 & & \\ 1 & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix} \end{array} u^{(\lambda)} = \frac{1}{M}\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix}.$$

Since A_ℓ is a non-singular matrix, $u^{(\lambda)}$ is a vertex. By Lemma 3.2, P_0 is the convex hull of all $2^\ell - 2$ vertices $u^{(\lambda)}$. Therefore, no other vertices exist. □

Proof of Theorem 3.1. Let x be a vertex of $P_0 \cap \bigcap_{i \in N(m)} \{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}$. Define $K = \{i \in N(m) \mid x_i = 0\}$ and let $k = |K|$. If $K = \emptyset$, i.e. $x_i > 0$ for all $i \in N(m)$, then x must already be a vertex of P_0 and, therefore, one of the $v^{(\lambda)}$ with $z(\lambda) = 0$. Otherwise, the vector $(x_i \mid i \notin K)$ consisting of the components of x with index in K must be a vertex of the $(\ell - k)$-dimensional standard polytope $P_0 \subset \mathbb{R}^{\ell-k}$ and thus x equals one of the $v^{(\lambda)}$ with $z(\lambda) = k$.

Conversely, every $v^{(\lambda)}$ is a vertex since it fulfills

$$\begin{pmatrix} A_{\ell - z(\lambda)} & \\ & I_{z(\lambda)} \end{pmatrix} v^{(\lambda)} = \frac{1}{M}\begin{pmatrix} 1_{\ell - z(\lambda) - 1} \\ 0_{z(\lambda)+1} \end{pmatrix}.$$

Finally, we have $v^{(\lambda)} \neq v^{(\bar{\lambda})}$ if λ and $\bar{\lambda}$ are two distinct vectors in $\{0,1\}^\ell \setminus \{0_\ell, 1_\ell\}$ such that there is $i, j \notin N(m)$ with $\lambda_i = \bar{\lambda}_j = 0$. Thus the number of vertices equals $|\{\lambda \in \{0,1\}^\ell \mid \lambda \neq 0_\ell, \lambda \neq 1_\ell, \exists i \notin N(m) : \lambda_i = 0\}| = 2^\ell - 2^{n(m)} - 1$. □

3.2. *Volumes*

A decomposition similar to Lemma 3.2 permits to compute in Theorem 3.4 the volume of an arbitrary restricted standard polytope, which equals the volume of the associated rounding polytope.

Theorem 3.4. *The volume of $P(m)$ depends only on $n := n(m)$ and is given by*

$$\mathrm{vol}_{\ell-1}(P(m)) = \frac{\sqrt{\ell}}{\binom{\ell}{n} M^{\ell-1}} \sum_{\substack{t \in \{0,1\}^{\ell-1}, \\ \sum_{i=1}^{\ell-1} t_i = n}} \prod_{j=1}^{\ell-2} \left(\frac{\ell - n + \sum_{k=1}^{j} t_k - j}{\ell - n + \sum_{k=1}^{j} t_k} \right)^{t_{j+1}}. \tag{9}$$

Corollary 3.5. *If $n(m) \in \{0, \ell - 1\}$, then the volume of $P(m)$ is given by*

$$\mathrm{vol}_{\ell-1}(P(m)) = \frac{\sqrt{\ell}}{M^{\ell-1}} \times \begin{cases} 1 & \textit{if } n(m) = 0, \\ 1/\ell! & \textit{if } n(m) = \ell - 1. \end{cases}$$

To prove Theorem 3.4 we establish in Lemma 3.6 a volume formula based on determinants. Simplifying this formula subsequently yields the theorem.

Lemma 3.6. *Let* $1 \notin N := N(m)$ (*otherwise permute the indices without changing volumes*). *Then*

$$\mathrm{vol}_{\ell-1}\left(P_0 \cap \bigcap_{i \in N}\{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}\right) = (\ell - n) \times \mathrm{vol}_{\ell-1}(U_1), \tag{10}$$

where U_1 *is the convex hull of* $\{v^{(\lambda)} \mid \lambda \in \{0,1\}^\ell,\ \lambda_1 = 0\}$. *If* π_1 *denotes the projection onto the components with index different from* 1 *then the volume of* U_1 *is*

$$\mathrm{vol}_{\ell-1}(U_1) = \frac{\sqrt{\ell}}{\binom{l-1}{n}} \sum_{\substack{t \in \{0,1\}^{\ell-1}, \\ \sum_{i=1}^{\ell-1} t_i = n}} \left|\det\left(\pi_1(v^{(\lambda^1(t))}), \ldots, \pi_1(v^{(\lambda^{\ell-1}(t))})\right)\right| \tag{11}$$

with

$$\lambda^j(t) = \sum_{i=1}^{j-\sum_{k=1}^{j} t_k} \varepsilon_{i+1} + \sum_{i=1}^{\sum_{k=1}^{j} t_k} \varepsilon_{\ell-n+i} \tag{12}$$

and ε_i *denoting the vector of the canonical basis in* $\mathbb{R}^\ell$ *having component* i *equal to 1 and all other components zero.*

Proof. Let U_i be the convex hull of $\{v^{(\lambda)} \mid \lambda \in \{0,1\}^\ell,\ \lambda_i = 0\}$. Since we can express every $v^{(\lambda)}$ with $\lambda_i = 0$ as a linear combination $\sum_{j \neq i}^{\ell} \mu_j u^{(j)}$ by setting $\mu_j = 1$ if $\lambda_j = 1$, $\mu_j = 0$ if $\lambda_j = 0$ and $j \notin N$, and $\mu_j = e(\lambda)/(\ell - z(\lambda))$ if $\lambda_j = 0$ and $j \in N$, we know that $U_i \subseteq L_i$. Therefore, the interior of $U_i \cap U_j$ is empty if $i \neq j$ and for all $i \in N$,

$$\mathrm{vol}_{\ell-1}(U_i) \leqslant \mathrm{vol}_{\ell-1}\left(L_i \cap \bigcap_{i \in N}\{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}\right) = 0.$$

By the definition of U_i as convex hull of vertices,

$$P_0 \cap \bigcap_{i \in N}\{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\} = \bigcup_{i \notin N} U_i.$$

Since permuting components i and j maps U_i in U_j and leaves the volume invariant,

$$\mathrm{vol}_{\ell-1}\left(P_0 \cap \bigcap_{i \in N}\{x \in \mathbb{R}^\ell \mid x_i \geqslant 0\}\right) = (\ell - n) \times \mathrm{vol}_{\ell-1}(U_1).$$

In order to calculate the volume of U_1, we decompose it into simplices. Let

$$\lambda^i = (0, 1_i, 0_{\ell-1-i})$$

and denote by S_1 the group of permutations of $\{1, \ldots, \ell\}$ leaving 1 fix. Then U_1 is the union of the simplices Δ_σ, $\sigma \in S_1$, defined as the convex hull of $u^{(\sigma(\lambda^i))}$, $i =$

$0, \dots, \ell-1$. Note that $\mathrm{int}(\Delta_\sigma) \cap \mathrm{int}(\Delta_\tau) = \emptyset$ if $\sigma \neq \tau$. The volume of a simplex Δ_σ is $\sqrt{\ell}$ times the full-dimensional volume of the projected simplex $\pi_1(\Delta_\sigma)$. The full-dimensional volume of $\pi_1(\Delta_\sigma)$ can be calculated by the determinant formula.

Let $\sigma, \tau \in S_1$, and define the equivalence relation

$$\sigma \sim \tau :\Longleftrightarrow [\sigma(i) \in N \iff \tau(i) \in N \quad \forall i]. \tag{13}$$

Then $\sigma \sim \tau$ implies that Δ_σ and Δ_τ can be mapped into each other by a permutation and thus have the same volume. Since each equivalence class consists of $(\ell - n + 1)!n!$ permutations, we arrive at the formula for the volume of U_1 stated in the theorem by summing over the representatives of each equivalence class. This is done by indexing the sum by vectors $t \in \{0,1\}^{\ell-1}$ where $t_i = 1$ means that all permutations σ in the corresponding equivalence class fulfill $\sigma(i+1) \in N$ and $t_i = 0$ signifies $\sigma(i+1) \notin N$. $\square$

Proof of Theorem 3.4. By (12), the vectors $\lambda^j(t) \in \{0,1\}^\ell$ have the form

$$\lambda^j(t) = (0, \underbrace{1, \dots, 1}_{j-\sum_{k=1}^{j} t_k}, 0, \dots, 0, \underbrace{1, \dots, 1}_{\sum_{k=1}^{j} t_k}, 0, \dots, 0) \tag{14}$$

with exactly j components equal to one. Let Λ be the square matrix with columns equal to the last $\ell - 1$ components of the vectors $\lambda^j(t)$, $j = 1, \dots, \ell-1$. Since $\lambda_i^j(t) = 1$ implies $\lambda_i^{j+1}(t) = 1$, we can transform Λ in an upper triangular matrix by permuting its rows. This transformation leaves the absolute value of the determinant of Λ unchanged. The same permutation shall be applied to

$$v(\Lambda) = \left(\pi_1(v^{\lambda^1(t)}), \dots, \pi_1(v^{\lambda^{\ell-1}(t)})\right).$$

By (7) and since $e(\lambda^j) = j$ and $z(\lambda^j) = n - \sum_{k=1}^{j-1} t_k$, it follows that after an appropriate permutation of rows

$$\det(v(\Lambda)) = \det \begin{pmatrix} \frac{1}{M} \times \frac{\ell-n+\sum_{k=1}^{1} t_k - 1}{\ell-n+\sum_{k=1}^{1} t_k} & \cdots & \frac{1}{M} \times \frac{\ell-n+\sum_{k=1}^{\ell-1} t_k-(\ell-1)}{\ell-n+\sum_{k=1}^{\ell-1} t_k} \\ & \ddots & \vdots \\ \bigstar & & \frac{1}{M} \times \frac{\ell-n+\sum_{k=1}^{\ell-1} t_k-(\ell-1)}{\ell-n+\sum_{k=1}^{\ell-1} t_k} \end{pmatrix}. \tag{15}$$

Here and in the remainder of the evaluation of $\det(v(\Lambda))$ we can ignore possible sign changes due to the absolute value in (11). The lower triangular part given as $\bigstar$ in (15) corresponds to zeros in the vectors $\lambda^j(t) (j = 1, \dots, \ell-1)$, and in the following only the first sub-diagonal will be of interest. By (7), (12) and (14), the sub-diagonal entry in the jth column of the permuted matrix $v(\Lambda)$ is equal to 0 if $t_{j+1} = 1$ and

$$-\frac{1}{M} \times \frac{\ell}{\ell - n + \sum_{k=1}^{j-1} t_k} \quad \text{if } t_{j+1} = 0.$$

To simplify (15), we subtract the first row of the matrix on the right hand side from all other rows. This gives

$$\det(v(\Lambda)) = \det \begin{pmatrix} \frac{1}{M} \times \frac{\ell-n+\sum_{k=1}^{1} t_k - 1}{\ell-n+\sum_{k=1}^{1} t_k} & \cdots & \cdots & \frac{1}{M} \times \frac{\ell-n+\sum_{k=1}^{\ell-1} t_k-(\ell-1)}{\ell-n+\sum_{k=1}^{\ell-1} t_k} \\ & 0 & \cdots & 0 \\ & & \ddots & \vdots \\ \bigstar & & & 0 \end{pmatrix}$$

where the sub-diagonal entry a_j in column j equals

$$a_j = \begin{cases} -\frac{1}{M} \times \frac{\ell-n+\sum_{k=1}^{j} t_k - j}{\ell-n+\sum_{k=1}^{j} t_k} & \text{if } t_{j+1} = 1, \\ -\frac{1}{M} & \text{if } t_{j+1} = 0. \end{cases}$$

Since by definition $\sum_{k=1}^{\ell-1} t_k = n$ it follows that

$$\det(v(\Lambda)) = \frac{\prod_{j=1}^{\ell-2} a_j}{M\ell},$$

which implies the result stated in Theorem 3.4. □

3.3. Examples

In order to illustrate the previous results we consider the rounding polytope for $m = (2, 2, 1)$, which is highlighted in Fig. 1(a). By Theorem 3.1 (with $N(m) = \emptyset$), the polytope's $2^3 - 2^0 - 1 = 6$ vertices are determined by the $v^{(\lambda)}$ given in Table 1. Adding $m/M = (2/5, 2/5, 1/5)$ to the $v^{(\lambda)}$ yields the vertices of $P(m)$, which we state in the order of appearance on a clockwise tour on the edges of $P(m)$:

Table 1
The vertices $v^{(\lambda)}$ of the standard polytope P_0 determined by Theorem 3.1 with $N(m) = \emptyset$

λ	$e(\lambda)$	$v^{(\lambda)}$
(0, 0, 1)	1	$\frac{1}{5}\left(-\frac{1}{3}, -\frac{1}{3}, \frac{2}{3}\right)$
(0, 1, 0)	1	$\frac{1}{5}\left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right)$
(0, 1, 1)	3	$\frac{1}{5}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$
(1, 0, 0)	1	$\frac{1}{5}\left(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}\right)$
(1, 0, 1)	2	$\frac{1}{5}\left(\frac{1}{3}, -\frac{2}{3}, \frac{1}{3}\right)$
(1, 1, 0)	2	$\frac{1}{5}\left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)$

$$\frac{1}{15}(5,5,5) \to \frac{1}{15}(4,7,4) \to \frac{1}{15}(5,8,2) \to \frac{1}{15}(7,7,1)$$
$$\to \frac{1}{15}(8,5,2) \to \frac{1}{15}(7,4,4).$$

It follows from Corollary 3.5 that $\mathrm{vol}_{\ell-1}(P(2,2,1)) = \sqrt{3}/5^2 \approx 0.069$.

4. Rounding polytopes for divisor methods

Let R be the divisor method with sign-post sequence s. Set $s(-1) := 0$. In Sections 4.1 and 4.2 we determine the vertices and the volume, respectively, of the associated rounding polytopes $P(m) := P_R(m)$. In Section 4.3 we simplify the volume formula for q-stationary divisor methods. Our results are illustrated in Section 4.4. Note that the cases when $m_i = 0$ for some i need no separate treatment. However, the set $\bar{N}(m) := \{i \mid s(m_i - 1) = 0\}$ will take over the role of $N(m)$, and we set $\bar{n}(m) := |\bar{N}(m)|$. The assumption $M > \ell$ implies $\bar{n}(m) \leqslant \ell - 1$. Note also that $N(m) = \bar{N}(m)$ if $s(0) > 0$.

4.1. Vertices

Theorem 4.1 gives the vertices of rounding polytopes for divisor methods.

Theorem 4.1. *The polytope $P(m)$ has $2^\ell - 2^{\bar{n}(m)} - 1$ vertices $v^{(\lambda)}$, which are induced by $\lambda \in \{0,1\}^\ell \setminus \{0_\ell, 1_\ell\}$ with $\lambda_j = 0$ for some index $j \notin \bar{N}(m)$. The components of $v^{(\lambda)}$ are*

$$v_i^{(\lambda)} = \begin{cases} s(m_i)/c(\lambda) & \text{if } \lambda_i = 1, \\ s(m_i - 1)/c(\lambda) & \text{if } \lambda_i = 0, \end{cases} \qquad i = 1, \ldots, \ell, \tag{16}$$

where the normalization is $c(\lambda) = \sum_{i:\lambda_i=1} s(m_i) + \sum_{i:\lambda_i=0} s(m_i - 1)$.

If $\bar{n}(m) = \ell - 1$, then $v^{(0_\ell)} := 0_\ell$ is also a vertex and $P(m)$ has $2^\ell - 2^{\bar{n}(m)} = 2^{\ell-1}$ vertices.

There are no other vertices than the indicated $v^{(\lambda)}$.

Remark 4.2. If $s(0) = 0$ and there exists $m_i = 0$ then $s(m_i) = s(m_i - 1) = 0$. This implies that $P(m)$ is degenerate in the sense that $w_i = 0$ for all $w \in P(m)$. Thus, $\dim(P(m)) \leqslant \ell - 2$ and $P(m) \subsetneq P(\tilde{m})$ for some $\tilde{m}$ with $\tilde{m}_j \geqslant 1$ for all j.

Proof of Theorem 4.1. Without loss of generality assume that m has ordered components, i.e. $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_\ell$ (otherwise permute the components of m appropriately). Since $M > \ell$ it holds that $m_1 \geqslant 2$ and $s(m_1 - 1) > 0$.

Obviously fulfilling the inequalities (2), every $v^{(\lambda)}$ lies in $P(m)$. To see that $v^{(\lambda)}$ is indeed a vertex, we concentrate first on $\bar{n}(m) = 0$ and without loss of generality on $v^{(\lambda)}$ with $\lambda = (0_k, 1_{\ell-k})$. Now it is easy to see that $v^{(\lambda)}$ solves the system $Av^{(\lambda)} = 0$ where A is the matrix

$$\begin{array}{c}
\begin{array}{ccccccc}
1 & \cdots & \cdots & k & k+1 & \cdots & \ell
\end{array} \\
\begin{pmatrix}
-s(m_{k+1}) & & & & s(m_1-1) & & \\
\vdots & & & & & \ddots & \\
-s(m_\ell) & & & & & & s(m_1-1) \\
 & -s(m_{k+1}) & & & s(m_2-1) & & \\
 & & \ddots & & \vdots & & \\
 & & & -s(m_{k+1}) & s(m_k-1) & & \\
1 & \cdots & \cdots & \cdots & \cdots & \cdots & 1
\end{pmatrix}
\end{array}$$

for which only non-zero entries are shown. Since $s(m_1 - 1) > 0$ and all $s(m_j) > 0$, A is of full rank ℓ. Hence, $v^{(\lambda)}$ is a vertex. If $\bar{n}(m) > 0$ then every component $v_i^{(\lambda)}$ with value zero fulfills the constraint $w_i \geqslant 0$ with equality, and we can argue in analogy to the case $\bar{n}(m) = 0$ by replacing the dimension ℓ by the number of non-zero components of $v^{(\lambda)}$.

No other vertices can exist since the convex hull of all $v^{(\lambda)}$ of form (16) is the whole polytope $P(m)$. This will be shown by establishing that

$$P(m) = \bigcup_{i \notin \bar{N}(m)} Q_i, \tag{17}$$

where Q_i is the convex hull of all $v^{(\lambda)}$ with $\lambda_i = 0$ and $\operatorname{int}(Q_i) \cap \operatorname{int}(Q_j) = \emptyset$ if $i \neq j$.

By definition, all Q_i are subsets of $P(m)$, which implies $\supseteq$ in (17). To see $\subseteq$ in (17), we first show that

$$\tilde{Q}_i = \{w \in P(m) \mid w_i s(m_k - 1) \leqslant w_k s(m_i - 1) \leqslant w_i s(m_k)\ \forall k \neq i\} \tag{18}$$

and Q_i coincide. Obviously, every $v^{(\lambda)}$ with $\lambda_i = 0$ is an element of the polytope $\tilde{Q}_i$, thus $Q_i \subseteq \tilde{Q}_i$. Conversely, every vertex w of $\tilde{Q}_i$ has its component w_k, $k \neq i$, determined as $s(m_k)w_i/s(m_i - 1)$ or $s(m_k - 1)w_i/s(m_i - 1)$. The condition $\sum_{k=1}^{\ell} w_k = 1$ implies that $w = v^{(\lambda)}$ for a λ with $\lambda_i = 0$. Hence, $\tilde{Q}_i \subseteq Q_i$.

Next, let w be a point in $P(m)$. Then we can choose an index $i \notin \bar{N}(m)$ such that $s(m_i - 1) > 0$ and $w_i/s(m_i - 1) \leqslant w_j/s(m_j - 1)$ for all $j \notin \bar{N}(m)$. Since w fulfills the inequalities (2) it follows from (18) that w is an element of Q_i.

Finally, according to (18), a point $w \in \operatorname{int}(Q_i) \cap \operatorname{int}(Q_j)$ fulfills

$$w_i s(m_j - 1) < w_j s(m_i - 1) < w_i s(m_j - 1).$$

Hence, $\operatorname{int}(Q_i) \cap \operatorname{int}(Q_j) = \emptyset$ if $i \neq j$. $\square$

4.2. Volumes

The knowledge about the vertices now allows us to decompose the projected cuboids Q_i from (17) into simplices whose volumes can be computed by the determinant formula. This ultimately yields the volume of $P(m)$ given in Theorem 4.3.

Theorem 4.3. *If $s(m_i) > 0$ for all i then*

$$\mathrm{vol}_{\ell-1}(P(m)) = \frac{\sqrt{\ell}}{c_0(\ell-1)!} \sum_{i \notin \bar{N}(m)} \left(s(m_i - 1) \prod_{j \neq i} d_j \right) \sum_{\sigma \in S_i} \frac{1}{\prod_{j \neq i} c_i^{\sigma}(j)}, \tag{19}$$

where $d_j = s(m_j) - s(m_j - 1)$, S_i is the group of permutations of $\{1, \ldots, \ell\}$ leaving i fix, $c_0 = \sum_{j=1}^{\ell} s(m_j - 1)$, and

$$c_i^{\sigma}(j) = \sum_{k=1, k \neq i}^{j} s(m_{\sigma(k)}) + \sum_{k=j+1,\ k \neq i}^{\ell} s(m_{\sigma(k)} - 1) + s(m_i - 1).$$

Remark 4.4. It follows directly from Remark 4.2 that if $s(m_i) = 0$ for some i, i.e. if $s(0) = 0$ and $m_i = 0$, then $\mathrm{vol}_{\ell-1}(P(m)) = 0$.

Proof of Theorem 4.3. Since permuting the components of a rounding result m does not change the volume of $P(m)$ we can assume without loss of generality that m has ordered components, i.e. $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_\ell$. This assumption implies $\bar{N}(m) = \{\ell - \bar{n}(m) + 1, \ldots, \ell\}$.

The proof of Theorem 4.1 establishes

$$\mathrm{vol}_{\ell-1}(P(m)) = \sum_{i \notin \bar{N}(m)} \mathrm{vol}_{\ell-1}(Q_i) = \sum_{i=1}^{\ell - \bar{n}(m)} \mathrm{vol}_{\ell-1}(Q_i). \tag{20}$$

Since all the Q_i can be treated analogously, we will only demonstrate the calculation of the volume of Q_1. The result for Q_i is obtained by interchanging indices. If we adopt the notation of the proof of Theorem 3.6 the arguments used there yield

$$\begin{aligned} &\mathrm{vol}_{\ell-1}(Q_1) \\ &\quad = \frac{\sqrt{\ell}}{(\ell-1)!} \sum_{\sigma \in S_1} \left| \det \left[\pi_1(v^{(\sigma(\lambda^1))} - v^{(0_\ell)}), \ldots, \pi_1(v^{(\sigma(\lambda^{\ell-1}))} - v^{(0_\ell)}) \right] \right|, \end{aligned} \tag{21}$$

where $v_k^{(0_\ell)} = s(m_k - 1)/c(0_\ell) = s(m_k - 1)/c_0$ for all $k = 1, \ldots, \ell$. In order to evaluate the determinant in (21), we need to study the vertex $v^{(\sigma(\lambda^j))}$. Its first component is $s(m_1 - 1)/c(\sigma(\lambda^j))$, and its $\sigma(k)$th component equals $s(m_{\sigma(k)})/c(\sigma(\lambda^j))$ if $k = 2, \ldots, j+1$, and $s(m_{\sigma(k)} - 1)/c(\sigma(\lambda^j))$ if $k = j+2, \ldots, \ell$. Obviously

$c_1^\sigma(j+1) = c(\sigma(\lambda^j))$, and setting $x^{(\sigma(\lambda^j))} := v^{(\sigma(\lambda^j))} \times c(\sigma(\lambda^j))$ we find that the determinant in (21) equals $D(\sigma)/c_0^{\ell-1}\prod_{j=2}^{\ell} c_1^\sigma(j)$ with

$$D(\sigma) = \det\left[\pi_1(c_0 x^{(\sigma(\lambda^1))} - c_1^\sigma(2)x^{(0_\ell)}), \dots, \pi_1(c_0 x^{(\sigma(\lambda^\ell))} - c_1^\sigma(\ell)x^{(0_\ell)})\right]. \tag{22}$$

Now if $k \in \{\sigma(2), \dots, \sigma(j+1)\}$ then

$$\begin{aligned}\left[\pi_1(c_0 x^{(\sigma(\lambda^j))} - c_1^\sigma(j+1)x^{(0_\ell)})\right]_k &= c_0 s(m_k) - c_1^\sigma(j+1)s(m_k-1)\\ &= c_0 d_k - s(m_k-1)\sum_{p=2}^{j+1} d_{\sigma(p)},\end{aligned}$$

since $c_1^\sigma(j+1) = c_0 + \sum_{p=2}^{j+1} d_{\sigma(p)}$. If $k \in \{\sigma(j+2), \dots, \sigma(\ell)\}$ then

$$\begin{aligned}\left[\pi_1(c_0 x^{(\sigma(\lambda^j))} - c_1^\sigma(j+1)x^{(0_\ell)})\right]_k &= c_0 s(m_k-1) - c_1^\sigma(j+1)s(m_k-1)\\ &= -s(m_k-1)\sum_{p=2}^{j+1} d_{\sigma(p)}.\end{aligned}$$

In the following evaluation of $D(\sigma)$ we can ignore possible sign changes because of the absolute value in (21). Switching row $\sigma(j)$ in row j yields that

$$D(\sigma) = \det\begin{pmatrix} c_0 d_{\sigma(2)} - s(m_{\sigma(2)}-1)d_{\sigma(2)} & \cdots & c_0 d_{\sigma(2)} - s(m_{\sigma(2)}-1)\sum_{p=2}^{\ell} d_{\sigma(p)}\\ -s(m_{\sigma(3)}-1)d_{\sigma(2)} & \cdots & c_0 d_{\sigma(3)} - s(m_{\sigma(3)}-1)\sum_{p=2}^{\ell} d_{\sigma(p)}\\ \vdots & \ddots & \vdots\\ -s(m_{\sigma(\ell)}-1)d_{\sigma(2)} & \cdots & c_0 d_{\sigma(\ell)} - s(m_{\sigma(\ell)}-1)\sum_{p=2}^{\ell} d_{\sigma(p)} \end{pmatrix}.$$

By factoring out $d_{\sigma(2)}$ and adding $-\sum_{p=2}^{k+1} d_{\sigma(p)}$ times column 1 to column k, $k = 1, \dots, \ell-1$,

$$D(\sigma) = d_{\sigma(2)} \times \det\begin{pmatrix} c_0 - s(m_{\sigma(2)}-1) & -c_0 d_{\sigma(3)} & \cdots & -c_0\sum_{p=3}^{\ell} d_{\sigma(p)}\\ -s(m_{\sigma(3)}-1) & c_0 d_{\sigma(3)} & \cdots & c_0 d_{\sigma(3)}\\ -s(m_{\sigma(4)}-1) & 0 & & c_0 d_{\sigma(4)}\\ \vdots & \vdots & \ddots & \vdots\\ -s(m_{\sigma(\ell)}-1) & 0 & \cdots & c_0 d_{\sigma(\ell)} \end{pmatrix}.$$

Now adding each row k, $k \geqslant 2$ to row 1, we obtain that

$$
\begin{aligned}
D(\sigma) &= d_{\sigma(2)} \times \det \begin{pmatrix} s(m_1-1) & 0 & \cdots & 0 \\ -s(m_{\sigma(3)}-1) & c_0 d_{\sigma(3)} & \cdots & c_0 d_{\sigma(3)} \\ -s(m_{\sigma(4)}-1) & 0 & & c_0 d_{\sigma(4)} \\ \vdots & \vdots & \ddots & \vdots \\ -s(m_{\sigma(\ell)}-1) & 0 & \cdots & c_0 d_{\sigma(\ell)} \end{pmatrix} \\
&= s(m_1-1) c_0^{\ell-2} \prod_{j=2}^{\ell} d_{\sigma(j)}.
\end{aligned}
$$

It follows that the modulus of the determinant in (21) is equal to

$$
\frac{s(m_1-1)\prod_{j=2}^{\ell} d_{\sigma(j)}}{c_0 \prod_{j=2}^{\ell} c_1^{\sigma}(j)} = \frac{s(m_1-1)\prod_{j=2}^{\ell} d_j}{c_0 \prod_{j=2}^{\ell} c_1^{\sigma}(j)}. \tag{23}
$$

When calculating $\mathrm{vol}_{\ell-1}(Q_i)$ instead of $\mathrm{vol}_{\ell-1}(Q_1)$, the result in (23) becomes

$$
\frac{s(m_i-1)\prod_{j\neq i}^{\ell} d_j}{c_0 \prod_{j\neq i}^{\ell} c_i^{\sigma}(j)}.
$$

Summing the pieces as in (20) yields formula (19) claimed in the theorem. □

4.3. Volumes for stationary divisor methods

Let R be the q-stationary divisor method, i.e. $s(k) = k + q$. Then the differences $d_k = s(m_k) - s(m_k - 1)$ are equal to 1 if $m_k \geqslant 1$ and q if $m_k = 0$. This permits simplification of formula (19) to the result in Theorem 4.5, which shows that the volume of $P(m) := P_R(m)$ depends only on $n(m) = |\{i \mid m_i = 0\}|$.

Theorem 4.5. *The volume of $P(m)$ depends only on $n := n(m)$ and is given by*

$$
\begin{aligned}
\mathrm{vol}_{\ell-1}\,(P(m)) = {} & \frac{q^n \sqrt{\ell}}{\binom{\ell-1}{n}(M+\ell q-1)} \\
& \times \sum_{\substack{t\in\{0,1\}^{\ell-1},\\ \sum_{i=1}^{\ell-1} t_i = n}} \frac{1}{\prod_{j=1}^{\ell-2}(M + j + (\ell - \sum_{k=j+1}^{\ell-1} t_k)(q-1))},
\end{aligned} \tag{24}
$$

where we set $0^0 := 1$.

Remark 4.6. If $q = 0$ and $n > 0$, then $q^n = 0$ and $\mathrm{vol}_{\ell-1}(P(m)) = 0$. If $q = 1$, then all rounding polytopes have the same volume.

Corollary 4.7. *If $n(m) \in \{0, \ell - 1\}$, then the volume of $P(m)$ for $q \in [0, 1]$ is given by*

$$\mathrm{vol}_{\ell-1}(P(m)) = \sqrt{\ell} \times \begin{cases} \frac{1}{\prod_{j=1}^{\ell-1}(M+\ell q-j)} & \text{if } n(m) = 0,\\ \frac{q^{\ell-1}}{\prod_{j=2}^{\ell}(M+jq-1)} & \text{if } n(m) = \ell - 1. \end{cases}$$

The case $n(m) = 0$ in Corollary 4.7 is treated in [13, p. 204, Theorem 6.2.10].

Proof of Theorem 4.5. The case $q = 0$ and $n > 0$ is an immediate consequence of Remark 4.4. For $q > 0$ as well as for $q = 0$ and $n = 0$, it holds that $s(m_i) > 0$ for all i. We assume that m is ordered as $m_1 \geqslant \cdots \geqslant m_\ell$, which implies that $s(m_1 - 1) > 0$, and we can apply formula (19) from Theorem 4.3.

Now $s(m_i - 1) = m_i + q - 1$ for all $i \leqslant \ell - n$. Moreover, $d_j = 1$ if $j \leqslant \ell - n$, and $d_j = q$ if $j \geqslant \ell - n + 1$. Thus $c_0 = M + (\ell - n)(q - 1)$ and

$$\begin{aligned} c_i^\sigma(j) &= \sum_{k=1,\ k\neq i}^{j} s(m_{\sigma(k)}) + \sum_{k=j+1,\ k\neq i}^{\ell} s(m_{\sigma(k)} - 1) + s(m_i - 1)\\ &= \sum_{k=1,\ k\neq i}^{j} (m_{\sigma(k)} + q) + \sum_{\substack{k=1\\ k\in\sigma(\{j+1,\dots,\ell\}\setminus\{i\})}}^{\ell-n} (m_k + q - 1) + (m_i + q - 1)\\ &= \begin{cases} M + jq + (1 + p_i^\sigma(j))(q - 1) & \text{if } j \leqslant i - 1,\\ M + (j - 1)q + (1 + p_i^\sigma(j))(q - 1) & \text{if } j \geqslant i + 1, \end{cases} \end{aligned} \tag{25}$$

where $p_i^\sigma(j) = |\sigma(\{j + 1, \dots, \ell\} \setminus \{i\}) \cap \{1, \dots, \ell - n\}|$. It follows from (19) that

$$\begin{aligned} \mathrm{vol}_{\ell-1}(P(m)) &= \frac{q^n\sqrt{\ell}}{(\ell - 1)!} \times \frac{\sum_{i=1}^{\ell-n}(m_i + q - 1)}{(M + (\ell - n)(q - 1))} \sum_{\sigma\in S_i} \frac{1}{\prod_{j\neq i} c_i^\sigma(j)}\\ &= \frac{q^n\sqrt{\ell}}{(\ell - 1)!} \sum_{\sigma\in S_i} \frac{1}{\prod_{j\neq i} c_i^\sigma(j)}. \end{aligned} \tag{26}$$

Let $\Sigma_i = \sum_{\sigma\in S_i} 1/\prod_{j\neq i} c_i^\sigma(j)$. Then Σ_i is independent of the index i, which we prove by showing that $\Sigma_i = \Sigma_1$ for all $i \leqslant \ell - n$. To do so, we introduce the bijection $f : S_i \to S_1$ for $1 < i \leqslant \ell - n$. Since $f(\sigma) \in S_1$ it must hold that $f(\sigma)(1) = 1$. The remaining components of $f(\sigma)$ are defined as follows. If $\sigma^{-1}(1) \leqslant i$ then $f(\sigma)(j) = \sigma(j - 1)$ for all $j \in \{2 \dots, i\} \setminus \{\sigma^{-1}(1) + 1\}$, $f(\sigma)(\sigma^{-1}(1) + 1) = i$, and $f(\sigma)(j) = \sigma(j)$ for all $j \geqslant i + 1$. Otherwise, if $\sigma^{-1}(1) \geqslant i + 1$ then $f(\sigma)(j) = \sigma(j - 1)$ for all $2 \leqslant j \leqslant i$, $f(\sigma)(j) = \sigma(j)$ for all $j \in \{i + 1, \dots, \ell\} \setminus \{\sigma^{-1}(1)\}$, and $f(\sigma)(\sigma^{-1}) = i$. Then it is easy to see that $p_1^{f(\sigma)}(j) = p_i^\sigma(j - 1)$ if $2 \leqslant j \leqslant i$, and $p_1^{f(\sigma)}(j) = p_i^\sigma(j)$ if $j \geqslant i+1$. It follows that $c_i^\sigma(j) = c_1^{f(\sigma)}(j+1)$ if $1 \leqslant j \leqslant i-1$, and $c_i^\sigma(j) = c_1^{f(\sigma)}(j)$ if $i+1 \leqslant j \leqslant \ell$, which yields

$$\frac{1}{\prod_{j \neq i} c_i^\sigma(j)} = \frac{1}{\prod_{j=2}^{\ell} c_1^{f(\sigma)}(j)}$$

and $\Sigma_i = \Sigma_1$. Hence (26) simplifies to

$$\mathrm{vol}_{\ell-1}(P(m)) = \frac{q^n \sqrt{\ell}}{(\ell-1)!} \times \sum_{\sigma \in S_1} \frac{1}{\prod_{j=2}^{\ell} c_1^\sigma(j)}. \tag{27}$$

It is easy to see that $c_1^\sigma(j) = c_1^\tau(j)$ if $\sigma \sim \tau$ in the sense of (13). As in the proof of Lemma 3.6 we index each equivalence class by a vector $t \in \{0,1\}^{\ell-1}$, $\sum_{k=1}^{\ell-1} t_k = n$, such that each permutation $\sigma \in S_1$ in an equivalence class associated with t satisfies $\sigma(k+1) \leqslant \ell - n$ if $t_k = 0$, and $\sigma(k+1) \geqslant \ell - n + 1$ if $t_k = 1$. For a permutation $\sigma \in S_1$ associated to t,

$$p_1^\sigma(j) = \sum_{k=j}^{\ell-1}(1-t_k) = (\ell - j) - \sum_{k=j}^{\ell-1} t_k.$$

Plugging this into (25) and the result for $c_1^\sigma(j)$ into (27) gives

$$\begin{aligned}&\mathrm{vol}_{\ell-1}(P(m))\\ &= \frac{q^n\sqrt{\ell}}{\binom{\ell-1}{n}} \times \sum_{\substack{t \in \{0,1\}^{\ell-1},\\ \sum_{i=1}^{\ell-1} t_i = n}} \frac{1}{\prod_{j=2}^{\ell}(M + (j-1)q + (1+\ell-j-\sum_{k=j}^{\ell-1} t_k)(q-1))},\end{aligned}$$

which implies the claimed formula (24). □

4.4. Examples

As in Section 3.3, we illustrate our results by the rounding polytope for $m = (2, 2, 1)$, which is highlighted for $q = 0, 0.5$, and 1 in Fig. 1(b), (c), and (d), respectively. For $q = 0.5$ and $q = 1$, $\bar{n}(m) = 0$ and $P(m)$ has $2^3 - 2^0 - 1 = 6$ vertices (cf. Theorem 4.1). For $q = 0$, $\bar{n}(m) = 1$ and $P(m)$ has $2^3 - 2^1 - 1 = 5$ vertices ($v^{(1,1,0)}$ is not a vertex since $\lambda_j \neq 0$ for all $j \notin \bar{N}(m) = \{1, 2\}$). Table 2 gives the coordinates of these vertices.

In the cases $q = 0.5$ and $q = 1$ the vertices of $P(m)$ can be arranged on a clockwise tour on the edges of $P(m)$ according to the sequence

$$(0,0,1) \to (0,1,1) \to (0,1,0) \to (1,1,0) \to (1,0,1) \to (1,0,0),$$

of vertex-inducing vectors λ. In the case $q = 0$, $\lambda = (1,1,0)$ would be skipped.

Table 2
The vertices $v^{(\lambda)}$ of $P(m)$, $m = (2, 2, 1)$, determined by Theorem 4.1

	λ	$v^{(\lambda)}c(\lambda)$	$c(\lambda)$	$v^{(\lambda)}$
$q = 0$	(0, 0, 1)	(1, 1, 1)	3	$\frac{1}{3}(1, 1, 1)$
	(0, 1, 0)	(1, 2, 0)	3	$\frac{1}{3}(1, 2, 0)$
	(0, 1, 1)	(1, 2, 1)	4	$\frac{1}{4}(1, 2, 1)$
	(1, 0, 0)	(2, 1, 0)	3	$\frac{1}{3}(2, 1, 0)$
	(1, 0, 1)	(2, 1, 1)	4	$\frac{1}{4}(2, 1, 1)$
$q = 0.5$	(0, 0, 1)	(1.5, 1.5, 1.5)	4.5	$\frac{1}{9}(3, 3, 3)$
	(0, 1, 0)	(1.5, 2.5, 0.5)	4.5	$\frac{1}{9}(3, 5, 1)$
	(0, 1, 1)	(1.5, 2.5, 1.5)	5.5	$\frac{1}{11}(3, 5, 3)$
	(1, 0, 0)	(2.5, 1.5, 0.5)	4.5	$\frac{1}{9}(5, 3, 1)$
	(1, 0, 1)	(2.5, 1.5, 1.5)	5.5	$\frac{1}{11}(5, 3, 3)$
	(1, 1, 0)	(2.5, 2.5, 0.5)	5.5	$\frac{1}{11}(5, 5, 1)$
$q = 1$	(0, 0, 1)	(2, 2, 2)	6	$\frac{1}{6}(2, 2, 2)$
	(0, 1, 0)	(2, 3, 1)	6	$\frac{1}{6}(2, 3, 1)$
	(0, 1, 1)	(2, 3, 2)	7	$\frac{1}{7}(2, 3, 2)$
	(1, 0, 0)	(3, 2, 1)	6	$\frac{1}{6}(3, 2, 1)$
	(1, 0, 1)	(3, 2, 2)	7	$\frac{1}{7}(3, 2, 2)$
	(1, 1, 0)	(3, 3, 1)	7	$\frac{1}{7}(3, 3, 1)$

Finally, by Corollary 3.5,

$$\mathrm{vol}_{\ell-1}(P(m)) = \begin{cases} \frac{\sqrt{3}}{4\times 3} \approx 0.144 & (q = 0), \\ \frac{\sqrt{3}}{5.5\times 4.5} \approx 0.070 & (q = 0.5), \\ \frac{\sqrt{3}}{7\times 6} \approx 0.041 & (q = 1). \end{cases}$$

5. An application in political science

In the electoral apportionment problem, the weight vectors w are vote fractions of parties and the rounding result m is the vector of the seats in a parliament allocated to the different parties. The number of seats for the ith party is an approximation to the idealized share of seats the party should obtain, i.e. $m_i \approx w_i M$. Since the number of votes is usually much larger than the number of seats M, there will be an inevitable gap $m_i - w_i M$, which expresses whether party i obtains more or less seat fractions than its ideal share projects.

Schuster et al. [18] investigate whether a rounding method leads to a systematic advantage for large (or for small) parties. To formalize this question, they condition uniformly distributed weight vectors w to be ordered as $w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell$ and define the seat-bias of the ith largest party as

$$B_i(M) := \mathrm{E}[m_i - w_i M \mid w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell].$$

A bias $B_i(M) = 0.3$, for example, means that in 10 elections for a parliament of M seats the ith largest party gains on average three seats. Here we show how Schuster et al. used the results we developed in Sections 3 and 4 to find their formulas for $B_i(M)$, which are stated without proof in [18].

The distribution of w conditional on $\{w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell\}$ is uniform on the ordered probability simplex $S_\geqslant := \{w \in S \mid w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell\}$, which, due to symmetry, has volume $\mathrm{vol}_{\ell-1}(S_\geqslant) = \mathrm{vol}_{\ell-1}(S)/\ell!$. First we compute the expected ideal share of seats of the ith largest party.

Lemma 5.1. *The expected ideal share of seats of the* ith *largest party equals*

$$I_i(M) := \mathrm{E}[w_i M \mid w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell] = \frac{M}{\ell}\sum_{j=0}^{\ell-i}\frac{1}{\ell - j} = \frac{M}{\ell}\sum_{j=i}^{\ell}\frac{1}{j}. \tag{28}$$

Proof. The vector $I(M) = (I_1(M), \ldots, I_\ell(M))$ equals M times $\mathrm{E}[w \mid w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell]$, and the latter expectation is the center of mass of the simplex $S_\geqslant$. The center of mass of a simplex is the (arithmetic) mean of its vertices. The vertices of $S_\geqslant$ are the vectors $w^{(j)}$, $j = 0, \ldots, \ell - 1$, with ith component $w_i^{(j)} = 1/(\ell - j)$ if $i \leqslant \ell - j$ and $w_i^{(j)} = 0$ else (cf. [8]). Multiplying the mean of these vertices by M yields (28). $\square$

Next we compute $\mathrm{E}[m_i|w_1 \geqslant w_2 \geqslant \cdots \geqslant w_\ell]$ by summing over all possible rounding results m. Since the studied methods round ordered weight vectors to ordered rounding results, the sum is over all m in $G_\geqslant := \{m \in G \mid m_1 \geqslant m_2 \geqslant \cdots \geqslant m_\ell\}$. The terms m_i in the sum are weighted by the probability that m is the rounding result. Since the weights are uniformly distributed this probability is equal to $\mathrm{vol}_{\ell-1}(P(m) \cap S_\geqslant)/\mathrm{vol}_{\ell-1}(S_\geqslant)$.

Let $b(m) := \prod_{i=1}^{\ell-1} |\{j \mid m_{i+j} = m_i,\ 0 \leqslant j \leqslant \ell - i\}|$ count the permutations that leave m invariant. Then $\mathrm{vol}_{\ell-1}(P(m) \cap S_\geqslant) = \mathrm{vol}_{\ell-1}(P(m))/b(m)$, and

$$B_i(M) = \frac{\ell!}{\mathrm{vol}_{\ell-1}(S)} \times \sum_{m \in G_\geqslant} m_i \frac{\mathrm{vol}_{\ell-1}(P(m))}{b(m)} - I_i(M).$$

With the results for $\mathrm{vol}_{\ell-1}(P(m))$ developed in Sections 3 and 4, it is a lengthy but straightforward calculation to find the formulas for $B_i(M)$ given in [18].

Acknowledgements

We thank Friedrich Pukelsheim for introducing us to the topic of rounding methods, and the referee for very helpful suggestions.

References

[1] B.D. Athanasopoulos, Probabilistic approach to the rounding problem with applications to fair representation, in: Approximation, Probability, and Related Fields (Santa Barbara, CA, 1993), Plenum, New York, 1994, pp. 75–99.
[2] B.D. Athanasopoulos, Probabilistic approaches to the rounding problem, Serdica 20 (1994) 306–327.
[3] M.L. Balinski, S.T. Rachev, Rounding proportions: rules of rounding, Numer. Funct. Anal. Optim. 14 (1993) 475–501.
[4] M.L. Balinski, S.T. Rachev, Rounding proportions: methods of rounding, Math. Sci. 22 (1997) 1–26.
[5] M.L. Balinski, H.P. Young, Fair Representation—Meeting the Ideal of One Man, One Vote, second ed., Brookings Institution Press, Washington, DC, 2001.
[6] H. Bergold, Wie gerecht sind unsere Verhältniswahlen? Didaktik Math. 7 (1979) 266–278.
[7] H. Bergold, Mehr Chancen für Außenseiter, Didaktik Math. 18 (1990) 38–51.
[8] W. Bühler, Characterization of the extreme points of a class of special polyhedra, Z. Oper. Res. 19 (1975) 131–137.
[9] G. Dorfleitner, T. Klein, Rounding with multiplier methods: an efficient algorithm and applications in statistics, Statist. Papers 40 (1999) 143–157.
[10] M. Drton, U. Schwingenschlögl, On methods for rounding probabilities and other fractions, Technical Report 450, Institut für Mathematik, Universitat Augsburg, 2003.
[11] C.H. Edwards Jr., Advanced Calculus of Several Variables, Dover Publications, New York, 1994 (Corrected reprint of the 1973 original).
[12] M. Happacher, The discrepancy distribution of stationary multiplier rules for rounding probabilities, Metrika 53 (2001) 171–181.
[13] K. Kopfermann, Mathematische Aspekte der Wahlverfahren—Mandatsverteilung bei Abstimmungen, Wissenschaftsverlag, Mannheim, 1991.
[14] A.W. Marshall, I. Olkin, F. Pukelsheim, A majorization comparison of apportionment methods in proportional representation, Social Choice and Welfare 19 (2002) 885–900.
[15] G. Pólya, Sur la représentation proportionelle en matière électorale, L'Enseignement Math. 20 (1918) 355–379.
[16] F. Pukelsheim, Efficient rounding of sampling allocations, Statist. Probab. Lett. 35 (1997) 141–143.
[17] F. Pukelsheim, S. Rieder, Efficient rounding of approximate designs, Biometrika 79 (1992) 763–770.
[18] K. Schuster, F. Pukelsheim, M. Drton, N.R. Draper, Seat-biases of apportionment methods for proportional representation, Electoral Studies 22 (2003) 651–676.
[19] R. Webster, Convexity, Oxford University Press, Oxford, 1994.
[20] D. Wode, Über die Wahrscheinlichkeit absoluter Mehrheiten, Didaktik Math. 9 (1981) 104–110.