An adaptive algorithm for vector partitioning

Komei Fukuda*†
IFOR, ETH Zurich
Ch-8092, Switzerland
fukuda@ifor.math.ethz.ch

Shmuel Onn ‡
Technion - IIT
Haifa, Israel
onn@ie.technion.ac.il

Vera Rosta
Webster University
Geneva, Switzerland
rosta@math.mcgill.ca

November 5, 2001

Abstract

The *vector partition problem* concerns the partitioning of a set A of n vectors in d-space into p parts so as to maximize an objective function c which is convex on the sum of vectors in each part. Here all parameters d, p, n are considered variables. In this paper, we study the adjacency of vertices in the associated *partition polytopes*. Using our adjacency characterization for these polytopes, we are able to develop an adaptive algorithm for the vector partition problem that runs in time $O(q(L) \cdot v)$ and in space $O(L)$, where q is a polynomial function, L is the input size and v is the number of vertices of the associated partition polytope. It is based on an output-sensitive algorithm for enumerating all vertices of the partition polytope. Our adjacency characterization also implies a polynomial upper bound on the combinatorial diameter of partition polytopes. We also establish a partition polytope analogue of the lower bound theorem, indicating that the output-sensitive enumeration algorithm can be far superior to previously known algorithms that run in time polynomial in the size of the worst-case output.

Key words: partition polytope, vertex enumeration, output-sensitive, polytope diameter, combinatorial optimization

1 Partition problems and partition polytopes

The *vector partition problem* concerns the partitioning of a multiset A of n vectors in d-space into p parts so as to maximize an objective function which is convex on the sum of vectors in each part. More precisely, with each *p-partition* of A, namely, an ordered tuple $\pi = (\pi_1, \ldots, \pi_p)$ of p pairwise disjoint (possibly empty) multisets whose union is A, we associate the $d \times p$ matrix

$$A^\pi \quad := \quad \left[\sum_{a\in\pi_1} a, \ldots, \sum_{a\in\pi_p} a\right] \quad \in \quad \mathbb{R}^{d\times p} \quad ;$$

*Corresponding author.

†Research supported in part by a grant from the Swiss National Foundation.

‡Research supported in part by a grant from the Israel Science Foundation, by a VPR grant at the Technion, and by the Fund for the Promotion of Research at the Technion.

given now a convex functional $c : \mathbb{R}^{d\times p} \longrightarrow \mathbb{R}$, the problem is to find a p-partition π maximizing the objective value $c(A^\pi)$ over all p-partitions of A. Such partition problems arise in a variety of areas ranging from economics to symbolic computation - see [9, 10, 13] and references therein.

Formally we are concerned with the following combinatorial optimization problem (with the reals replaced by the rationals when the Turing computation model is considered).

Vector Partition Problem: Given positive integers p, d, n, a multiset A of n points in $\mathbb{R}^d$, and a convex functional $c : \mathbb{R}^{d\times p} \longrightarrow \mathbb{R}$, find a p-partition π^* attaining maximum objective value,

$$c(A^{\pi^*}) \quad = \quad \max \{ c(A^\pi) \ : \ \pi \text{ is a } p\text{-partition of } A \} \quad .$$

The problem can be reduced to maximizing the same objective over the *p-partition polytope* $\mathcal{P}_A^p$ of A defined to be the convex hull in $\mathbb{R}^{d\times p}$ of all p^n matrices A^π associated with p-partitions,

$$\mathcal{P}_A^p \quad := \quad \text{conv} \{ A^\pi \ : \ \pi \text{ is a } p\text{-partition of } A \} \quad \subset \quad \mathbb{R}^{d\times p} \quad .$$

There will always be an optimal solution which is a vertex of $\mathcal{P}_A^p$, so the partition problem can be solved by picking the best vertex. When the functional c is presented by an evaluation oracle, it is necessary in worst case to query the oracle on every vertex. Thus, the complexity of the partition problem is intimately related to the vertex complexity of the corresponding partition polytope.

Much of the previous work has concentrated on worst case analysis: let $v_{p,d}(n)$ denote the maximal number of vertices of $\mathcal{P}_A^p$ for any set A of n points in d-space. Since a partition π is *separable* whenever A^π is a vertex [7], an upper bound on $v_{p,d}(n)$ follows from the results of [1] where the maximal number of separable partitions was determined to be $\Theta(n^{d\binom{p}{2}})$ for every fixed $p \geq 2$ and $d \geq 3$. This bound was exploited in [10] to provide an efficient procedure for solving the more general class of *shaped partition problems*, where partitions are restricted to be those π whose *shape* $(|\pi_1|, \ldots, |\pi_p|)$ lies in a prescribed (but arbitrary) set of shapes of n.

More recently, using methods suitable for partitions without shape restrictions, it was shown in [3] and [12] that, for every fixed d and $p \geq 2$, the number $v_{p,d}(n)$ obeys the sharper bounds

$$\Omega(n^{\lfloor \frac{d-1}{2} \rfloor p}) \quad \leq \quad v_{p,d}(n) \quad \leq \quad O(n^{d(p-1)-1}) \quad .$$

The main purpose of the present article is to provide an efficient *adaptive* algorithm - that is - one whose running time depends on the complexity of the actual problem instance at hand. Thus, the running time of the algorithm will be proportional to the vertex complexity of the specific partition polytope corresponding to the given instance of the vector partition problem. We prove:

Theorem 1: For any variable d, p, n, the partition problem with an oracle presented c can be solved in time $O(q(L)\cdot v)$ and in space $O(L)$, where q is a polynomial function, L is the input size and v is the number of vertices of the associated partition polytope.

In order to prove Theorem 1, we first obtain a characterization of vertex adjacency in partition polytopes (Lemmas 2.2 – 2.5). Our characterization allows us to generate all neighbor vertices of a given vertex in polynomial time, see Theorem 2.6. Another byproduct of our adjacency characterization is a polynomial upper bound on the combinatorial diameter of partition polytopes:

Theorem 2: The diameter of the p-partition polytope $\mathcal{P}_A^p$ of any n-set A in $\mathbb{R}^d$ is at most $n\binom{p}{2}$.

As we show, considering on the way a partition-analogue of the so-called Lower Bound Theorem, partition polytopes may have much fewer vertices than the worst case $v_{p,d}(n)$, in which case the algorithm provided in the present article drastically outperforms available ones. We say that A is a *free set* if it is a finite set of nonzero vectors, no two of which are positive multiples of one another. If A is not free than $\mathcal{P}_A^p$ may have as few as p^d extreme points independent of $n = |A|$ (see Section 5); on the other hand, any set can be replaced by a free set with no more points without affecting the corresponding partition problem and partition polytope (See Lemma 2.1 in the next section). Consequently it is natural to consider the lower bound theorem for free sets and maximal dimensional partition polytopes. Let $l_{p,d}(n)$ denote the minimal number of vertices of a maximal dimensional p-partition polytope $\mathcal{P}_A^p$ of any free set A of n vectors in d-space. We show:

Theorem 3: For every d, p and $n \geq d$ we have $l_{p,d}(n) \leq p^{d-1}(n-d+2)^{p-1}$.

Theorem 3 implies that for every fixed d and p we have $l_{p,d}(n) \leq O(n^{p-1})$ while $v_{p,d}(n) \geq \Omega(n^{\lfloor \frac{d-1}{2} \rfloor p})$, indicating that the output-sensitive enumeration algorithm can be far superior to previously known algorithms that run in time polynomial in the size of the worst-case output.

In Section 2, we give crucial properties and a characterization of edges of partition polytopes. Some of the properties will be used to prove Theorem 2 in Section 3. In Section 4, by exploiting our characterization of edges, we prove Theorem 1 which provides an efficient vertex enumeration algorithm for partition polytopes. Finally, we give a proof of Theorem 3 in Section 5.

2 Constructing the neighborhood of a vertex

Throughout, e_i stands for the ith standard unit vector in Euclidean real space. The tensor product of an ordered pair u, v of vectors is the matrix $u \otimes v$ whose (i,j)th entry is $u_i v_j$. The inner product of two matrices U, V of the same dimensions is $\langle U, V \rangle := \sum_{i,j} U_{i,j} \cdot V_{i,j}$. We make the convention that a sum over an empty set of vectors (matrices) is the zero vector (matrix) of dimension which is clear from the context. An edge $E = [u, v]$ of a polytope P is *in direction* e if $v - u$ is a scalar multiple of the vector e, that is, $E = P \cap (u + \mathrm{lin}(e))$. The *direction* of an edge $E = [u, v]$ is defined as $\mathrm{lin}(u - v)$.

A p-partition $\pi = (\pi_1, \ldots, \pi_p)$ of a multiset A of vectors in $\mathbb{R}^d$ will be also interpreted as the function from A to $\{1, \ldots, p\}$ with $\pi(a)$ being the index for which $a \in \pi_{\pi(a)}$. With this notation, the matrix associated with π is

$$A^\pi \quad = \quad \left[\sum_{a \in \pi_1} a, \ldots, \sum_{a \in \pi_p} a \right] \quad = \quad \sum_{a \in A} a \otimes e_{\pi(a)} \quad \in \quad \mathbb{R}^{d \times p} \quad .$$

We start by showing results that enable us to impose certain restrictions on the multiset A without losing generality.

Lemma 2.1 *Let π be a p-partition of a multiset A of vectors in d-space with A^π a vertex of $\mathcal{P}_A^p$. Suppose $a, b \in A$ satisfy $b = \alpha \cdot a$. If $\alpha > 0$ then $\pi(b) = \pi(a)$ whereas if $\alpha < 0$ then $\pi(b) \neq \pi(a)$.*

Proof. We prove the claim for $\alpha > 0$ and leave the analogous proof for $\alpha < 0$ to the reader. Let $\langle C, \cdot \rangle$ be a linear functional uniquely maximized over $\mathcal{P}_A^p$ at A. Suppose indirectly $\pi(b) \neq \pi(a)$. Let

$\bar{\pi}$ be obtained from π by the single modification $\bar{\pi}(a) := \pi(b)$ and let $\hat{\pi}$ be obtained from π by the single modification $\hat{\pi}(b) := \pi(a)$. Then

$$\langle C, a \otimes (e_{\pi(a)} - e_{\pi(b)}) \rangle \quad = \quad \langle C, A^{\pi} - A^{\bar{\pi}} \rangle \quad = \quad \langle C, A^{\pi} \rangle - \langle C, A^{\bar{\pi}} \rangle \quad > \quad 0 \quad ,$$

and

$$\langle C, b \otimes (e_{\pi(a)} - e_{\pi(b)}) \rangle \quad = \quad \langle C, A^{\hat{\pi}} - A^{\pi} \rangle \quad = \quad \langle C, A^{\hat{\pi}} \rangle - \langle C, A^{\pi} \rangle \quad < \quad 0 \quad .$$

But $\langle C, b \otimes (e_{\pi(a)} - e_{\pi(b)}) \rangle = \alpha \cdot \langle C, a \otimes (e_{\pi(a)} - e_{\pi(b)}) \rangle$ which is a contradiction. $\square$

The lemma implies that any vector partition problem has an optimal partition containing all points of A which are positive multiples of one another in the same part. Thus, all such points can be replaced by the single point which is their sum, giving the same partition polytope. Thus, without loss of generality we may assume that no point of A is a positive multiple of another. In particular, A contains no multiple points hence is a set rather than a multiset. We may also assume A contains no zero vectors, as these can be placed in an arbitrary part of the partition without affecting the objective function. So, from now on, we shall always assume that A is a *free set*, defined to be a finite set of nonzero points, no two of which are positive multiples of one another.

We proceed to determine a characterization of vertex neighborhoods of partition polytopes.

Lemma 2.2 *Let A be a free set in $\mathbb{R}^d$ and let π be a p-partition of A with A^{π} a vertex of $\mathcal{P}_A^p$. Then any edge of $\mathcal{P}_A^p$ containing A^{π} is in direction $a \otimes (e_t - e_{\pi(a)})$ for some $a \in A$ and $t \neq \pi(a)$.*

Proof. Consider any edge $E = [A^{\pi}, A^{\tau}]$ of $\mathcal{P}_A^p$ with τ another p-partition of A. Pick any $a \in A$ with $t := \tau(a) \neq \pi(a)$ and let $\langle C, \cdot \rangle$ be any linear functional uniquely maximized over $\mathcal{P}_A^p$ at E. Let $\bar{\pi}$ be obtained from π by the single modification $\bar{\pi}(a) := \tau(a)$ and let $\bar{\tau}$ be obtained from τ by the single modification $\bar{\tau}(a) := \pi(a)$. Then

$$\langle C, a \otimes (e_{\pi(a)} - e_{\tau(a)}) \rangle \quad = \quad \langle C, A^{\pi} - A^{\bar{\pi}} \rangle \quad = \quad \langle C, A^{\pi} \rangle - \langle C, A^{\bar{\pi}} \rangle \quad \geq \quad 0 \quad ,$$

and

$$\langle C, a \otimes (e_{\tau(a)} - e_{\pi(a)}) \rangle \quad = \quad \langle C, A^{\tau} - A^{\bar{\tau}} \rangle \quad = \quad \langle C, A^{\tau} \rangle - \langle C, A^{\bar{\tau}} \rangle \quad \geq \quad 0 \quad .$$

Thus, $\langle C, a \otimes (e_{\pi(a)} - e_{\tau(a)}) \rangle = 0$ hence $\langle C, A^{\bar{\pi}} \rangle = \langle C, A^{\pi} \rangle$. Since C is uniquely maximized over $\mathcal{P}_A^p$ at E, it follows that $A^{\bar{\pi}} \in E$ hence E is indeed in direction $A^{\bar{\pi}} - A^{\pi} = a \otimes (e_t - e_{\pi(a)})$. $\square$

One direct outcome of Lemma 2.2 is a polynomial bound on the number of edge directions of $\mathcal{P}_A^p$.

Theorem 2.3 *The number of edge directions in $\mathcal{P}_A^p$ is at most $n\binom{p}{2}$.*

Let A be a free set in $\mathbb{R}^d$. For each p-partition π of A, vector $a \in A$, and index $t \neq \pi(a)$, let $LP(\pi, a, t)$ be the following system of linear inequalities in the $d \times p$ matrix of variables C:

$$LP(\pi, a, t) : \left\{ \begin{array}{ll} \langle C, b \otimes (e_{\pi(b)} - e_k) \rangle \geq 1 & \forall\, b \in A\ \forall\, k \neq \pi(b) \text{ s.t. } b \notin \mathrm{lin}(a) \text{ or } \{k, \pi(b)\} \neq \{t, \pi(a)\} \\ \langle C, a \otimes (e_{\pi(a)} - e_t) \rangle = 0 & \end{array} \right.$$

Lemma 2.4 *Let A be a free set in $\mathbb{R}^d$, let π be a p-partition of A, and consider any $a \in A$ and $t \neq \pi(a)$. Then A^{π} lies on an edge of $\mathcal{P}_A^p$ in direction $a \otimes (e_t - e_{\pi(a)})$ if and only if the system of linear inequalities $LP(\pi, a, t)$ has a solution $C \in \mathbb{R}^{d \times p}$ (the linear program $LP(\pi, a, t)$ is feasible).*

Proof. Suppose first A^π lies on an edge E in direction $a \otimes (e_t - e_{\pi(a)})$, and let $\langle C, \cdot\rangle$ be a linear functional uniquely maximized over $\mathcal{P}^p_A$ at E. Since its value on E is constant we do have $\langle C, a \otimes (e_{\pi(a)} - e_t)\rangle = 0$. Now, consider any $b \in A$ and $k \neq \pi(b)$ such that $b \notin \mathrm{lin}(a)$ or $\{k, \pi(b)\} \neq \{t, \pi(a)\}$ and let $\bar\pi$ be obtained from π by the single modification $\bar\pi(b) := k$. Then

$$A^{\bar\pi} - A^\pi \quad = \quad b \otimes (e_k - e_{\pi(b)}) \quad \notin \quad \mathrm{lin}(a \otimes (e_t - e_{\pi(a)}))$$

hence $A^{\bar\pi} \notin E$, which implies

$$\langle C, b \otimes (e_{\pi(b)} - e_k)\rangle \quad = \quad \langle C, A^\pi - A^{\bar\pi}\rangle \quad = \quad \langle C, A^\pi\rangle - \langle C, A^{\bar\pi}\rangle \quad > \quad 0 \quad .$$

Therefore, a suitable positive multiple of C satisfies the linear inequality system $L(\pi, a, t)$.

Conversely, suppose C is a $d \times p$ matrix satisfying $L(\pi, a, t)$. Then the linear functional $\langle C, \cdot\rangle$ vanishes on the line $L := \mathrm{lin}(a \otimes (e_t - e_{\pi(a)}))$ and is constant over the segment $E := \mathcal{P}^p_A \cap (A^\pi + L)$. We show that $\langle C, \cdot\rangle$ is uniquely maximized over $\mathcal{P}^p_A$ at E which is therefore an edge of $\mathcal{P}^p_A$.

Consider any p-partition τ with $A^\tau \notin E$. Let $S \;:=\; \{b \in A \;:\; b \otimes (e_{\tau(b)} - e_{\pi(b)}) \in L\;\}$ and note that $b \in A \setminus S$ if and only if $\tau(b) \neq \pi(b)$ and either $b \notin \mathrm{lin}(a)$ or $\{\tau(b), \pi(b)\} \neq \{t, \pi(a)\}$. Thus $\langle C, b \otimes (e_{\pi(b)} - e_{\tau(b)})\rangle$ is 0 if $b \in S$ and is greater or equal 1 if $b \in A \setminus S$. Now $\sum_{b\in S} b \otimes (e_{\tau(b)} - e_{\pi(b)})$ is in L whereas $\sum_{b\in A} b \otimes (e_{\tau(b)} - e_{\pi(b)}) = A^\tau - A^\pi$ is not, so S is strictly contained in A. Therefore,

$$\langle C, A^\pi - A^\tau\rangle \;=\; \langle C, \sum_{b\in S} b \otimes (e_{\pi(b)} - e_{\tau(b)})\rangle \;+\; \langle C, \sum_{b\in A\setminus S} b \otimes (e_{\pi(b)} - e_{\tau(b)})\rangle \;\geq\; |A \setminus S| \cdot 1 \;>\; 0 \;.$$

Thus, $\langle C, A^\pi\rangle > \langle C, A^\tau\rangle$ for every $A^\tau \notin E$, completing the proof. $\square$

Let $a \in A$ be a point in a free set. While A contains no positive scalar multiple of a, it may contain one negative scalar multiple of a, which will be denoted by $\bar a$. We say that a is *single* in A if $\mathrm{lin}(a) \cap A = \{a\}$ whereas a is *double* in A if $\mathrm{lin}(a) \cap A = \{a, \bar a\}$. For each p-partition π of A, vector $a \in A$, and index $t \neq \pi(a)$, let $\tau(\pi, a, t)$ be the p-partition of A defined as follows:

$$\tau(\pi, a, t) := \tau : \begin{cases} \tau(a) := t \\ \text{if } a \text{ is double in } A \text{ then: if } \pi(\bar a) = t \text{ then } \tau(\bar a) := \pi(a) \text{ else } \tau(\bar a) := \pi(\bar a) \\ \forall\, b \notin \mathrm{lin}(a) \quad \tau(b) := \pi(b) \end{cases}$$

Lemma 2.5 *Let A be a free set in $\mathbb{R}^d$, let π be a p-partition of A, and consider any $a \in A$ and $t \neq \pi(a)$. If $\mathcal{P}^p_A$ has an edge $E = [A^\pi, A^\tau]$ in direction $a \otimes (e_t - e_{\pi(a)})$ then $\tau = \tau(\pi, a, t)$.*

Proof. Suppose $\mathcal{P}^p_A$ has such an edge E. By Lemma 2.4 there is a $d \times p$ matrix C satisfying $LP(\pi, a, t)$, and the linear functional $\langle C, \cdot\rangle$ is uniquely maximized over $\mathcal{P}^p_A$ at E.

Consider any $b \in A$ with $\tau(b) \neq \pi(b)$. If $b \notin \mathrm{lin}(a)$ or $\{\tau(b), \pi(b)\} \neq \{t, \pi(a)\}$ then, letting $\bar\tau$ be obtained from τ by the modification $\bar\tau(b) := \pi(b)$, we get $\langle C, A^{\bar\tau} - A^\tau\rangle = \langle C, b \otimes (e_{\pi(b)} - e_{\tau(b)})\rangle \geq 1$ contradicting the fact that $\langle C, \cdot\rangle$ is maximized at A^τ. It follows that whenever $\tau(b) \neq \pi(b)$, we must have $b \in \mathrm{lin}(a)$ and $\{\tau(b), \pi(b)\} = \{t, \pi(a)\}$. Thus, τ and π can differ only on $b \in \{a, \bar a\}$ and must differ on at least one such b.

If $\tau(a) \neq \pi(a)$ then the condition $\{\tau(a), \pi(a)\} = \{t, \pi(a)\}$ implies $\tau(a) = t$. If a is double and $\tau(\bar a) \neq \pi(\bar a)$ then, since $\pi(\bar a) \neq \pi(a)$ by Lemma 2.1, the condition $\{\tau(\bar a), \pi(\bar a)\} = \{t, \pi(a)\}$ implies

$\pi(\bar{a}) = t$ and $\tau(\bar{a}) = \pi(a)$; so if $\pi(\bar{a}) \neq t$ then necessarily $\tau(\bar{a}) = \pi(\bar{a})$. Thus, if either a is single or a is double but $\pi(\bar{a}) \neq t$, then τ and π differ only by $\tau(a) = t$ hence $\tau = \tau(\pi, a, t)$ as claimed.

Suppose now that a is double and $\pi(\bar{a}) = t$. Let $\alpha > 0$ be the scalar such that $\bar{a} = -\alpha \cdot a$. Let π^a be obtained from π by the single modification $\pi^a(a) := t$, let $\pi^{\bar{a}}$ be obtained from π by the single modification $\pi^{\bar{a}}(\bar{a}) := \pi(a)$, and let $\pi^{a\bar{a}} = \tau(\pi, a, t)$ be obtained from π by both modifications. Now τ and π can differ only on $b \in \{a, \bar{a}\}$ hence τ is one of $\pi^a, \pi^{\bar{a}}, \pi^{a\bar{a}}$, and $A^{\pi^a}, A^{\pi^{\bar{a}}}, A^{\pi^{a\bar{a}}} \in E$, hence $E = [A^\pi, A^\tau] = \text{conv}\{A^\pi, A^{\pi^a}, A^{\pi^{\bar{a}}}, A^{\pi^{a\bar{a}}}\}$. But now $A^{\pi^a} - A^\pi = a \otimes (e_t - e_{\pi(a)})$, $A^{\pi^{\bar{a}}} - A^\pi = \alpha \cdot a \otimes (e_t - e_{\pi(a)})$, and $A^{\pi^{a\bar{a}}} - A^\pi = (1+\alpha) \cdot a \otimes (e_t - e_{\pi(a)})$, hence $[A^\pi, A^\tau] = E = \text{conv}\{A^\pi, A^{\pi^a}, A^{\pi^{\bar{a}}}, A^{\pi^{a\bar{a}}}\} = [A^\pi, A^{\pi^{a\bar{a}}}]$ implying $\tau = \pi^{a\bar{a}} = \tau(\pi, a, t)$ as claimed. $\square$

Lemmas 2.2 – 2.5 yield at once the following characterization of neighborhoods of vertices of partition polytopes, which leads to an efficient procedure for generating the neighborhood of any given vertex of $\mathcal{P}_A^p$, providing the basic engine for the adaptive algorithm developed in next section.

Theorem 2.6 *Let A be a free set in $\mathbb{R}^d$ and let π be a p-partition of A with A^π a vertex of $\mathcal{P}_A^p$. Then the set of neighbors of A^π is*

$$N(A^\pi) \quad = \quad \{\, A^{\tau(\pi,a,t)} \ : \ a \in A,\ t \neq \pi(a),\ \textit{and } LP(\pi, a, t) \textit{ is feasible}\,\} \quad .$$

3 Parametric paths and the diameter of partition polytopes

In the previous section, we studied the neighbors of a given vertex of a p-partition polytope. Here we combine the results of Section 2 with results on parametric linear programming and establish Theorem 2, which generalizes a result of [2] on k-set polytopes that correspond to the special case of $p = 2$.

We start by reviewing the fundamental facts on parametric linear programming. Throughout this section we assume A and C are given matrices in $\mathbb{R}^{d\times n}$ and in $\mathbb{R}^{d\times p}$, respectively, and the columns of A form a free set.

The parametric linear programming method [8] starts with any vertex of $\mathcal{P}_A^p$, say V^0, and an objective function $\langle C^0, X\rangle$ which is maximized uniquely at V^0 over $\mathcal{P}_A^p$. For the moment, we assume that such V^0 and C^0 are given. We shall show later how one can find such vectors efficiently for the partition polytope. The key idea is to solve a sequence of the parametric LP problem

$$\begin{array}{lll} \text{LP}(\lambda): & \max & f^\lambda(X) := \langle \lambda C + (1-\lambda)C^0, X\rangle \\ & & \text{subject to } X \in \mathcal{P}_A^p, \end{array}$$

for increasing values of λ until it reaches one. More specifically, we compute an increasing sequence of parameter values $\lambda^0 < \lambda^1 < \cdots < \lambda^s < 1$ and vertices $V^0, V^1, \ldots, V^s$ of $\mathcal{P}_A^p$ such that for each $k = 0, 1, \ldots, s-1$,

(P1) the consecutive vertices V^k and V^{k+1} are adjacent in $\mathcal{P}_A^p$,

(P2) V^k is optimal for LP(λ) with all $\lambda^k \leq \lambda \leq \lambda^{k+1}$,

(P3) V^s is optimal for LP(1).

The existence of such a sequence is guaranteed by the following lemma.

Lemma 3.1 *Assume that the function $f^{\lambda}(X)$ is not constant on any 2-faces of the polytope $\mathcal{P}_A^p$, for any fixed $0 \le \lambda \le 1$. Let V^k be a vertex of $\mathcal{P}_A^p$ and let $0 \le \lambda^k < 1$ be a number such that the functional $f^{\lambda}(X)$ is maximized uniquely at V^k for any λ larger than λ^k but sufficiently close. Then, V^k maximizes $f^1(X)$ over $\mathcal{P}_A^p$ or there is a unique λ^{k+1} such that*

(a) $\lambda^k < \lambda^{k+1} < 1$,

(b) $f^{\lambda^{k+1}}(X)$ *is uniquely maximized over $\mathcal{P}_A^p$ at an edge $[V^k, V^{k+1}]$ for some neighbor V^{k+1} of V^k.*

Proof. Suppose that all assumptions are satisfied. If V^k maximizes $f^1(X)$ over $\mathcal{P}_A^p$, then we are done. So let us assume that this is not the case. This implies that there is a vertex V adjacent to V^k such that $f^1(V) > f^1(V^k)$. Let us define a partition (N^+, N^-) of the neighbor set $N(V^k)$ by

$$\begin{aligned} N^+ &:= \{V \in N(V^k) : f^1(V) > f^1(V^k)\} \quad , \\ N^- &:= \{V \in N(V^k) : f^1(V) \le f^1(V^k)\} \quad . \end{aligned}$$

For each $V \in N^+$, define $\lambda(V)$ as the unique number satisfying

$$f^{\lambda(V)}(V) = f^{\lambda(V)}(V^k) \quad . \tag{1}$$

To see why $\lambda(V)$ is well-defined, consider the linear function $F(\lambda) := f^{\lambda}(V) - f^{\lambda}(V^k)$. It is negative for all $\lambda > \lambda^k$ sufficiently close to λ^k since $f^{\lambda}(X)$ is maximized uniquely at V^k over $\mathcal{P}_A^p$ for all such λ, and it is positive for $\lambda = 1$ since $f^1(V) > f^1(V^k)$. Therefore there is a unique $\lambda^k < \lambda(V) < 1$ with $F(\lambda(V)) = 0$. In fact, by solving the equation (1) in terms of $\lambda(V)$, we have

$$\lambda(V) = \frac{\langle C^0, V - V^k \rangle}{\langle C^0 - C, V - V^k \rangle} \quad . \tag{2}$$

One can also verify that for $V \in N^+$,

$$f^{\lambda}(V) \le f^{\lambda}(V^k) \quad \forall \lambda \text{ satisfying } \lambda^k \le \lambda \le \lambda(V) \quad . \tag{3}$$

Furthermore, for $V \in N^-$,

$$f^{\lambda}(V) \le f^{\lambda}(V^k) \quad \forall \lambda \text{ satisfying } \lambda^k \le \lambda \quad . \tag{4}$$

Now, set $\lambda^{k+1} := \min\{\lambda(V) : V \in N^+\}$ and V^{k+1} to be a vertex in N^+ attaining this minimum. Using the relations (3), (4), we conclude

$$f^{\lambda^{k+1}}(V^{k+1}) = f^{\lambda^{k+1}}(V^k) \ge f^{\lambda^{k+1}}(V) \quad \forall V \in N(V^k) \quad .$$

This implies that the halfspace $\{X : f^{\lambda^{k+1}}(V^{k+1}) \ge f^{\lambda^{k+1}}(X)\}$ contains all edges incident to V^k and thus all points in $\mathcal{P}_A^p$. Consequently, with the equation (1), we know that the function $f^{\lambda^{k+1}}(X)$ is maximized at $[V^k, V^{k+1}]$ over $\mathcal{P}_A^p$. By the assumption that $f^{\lambda}(X)$ is not constant on any 2-faces of the polytope $\mathcal{P}_A^p$, the vertex V^{k+1}, a minimizer of $\lambda(V)$ in N^+, must be unique. This proves (b). Finally, since V^k is not a maximizer of $f^1(X)$, $\lambda^{k+1} < 1$, and (a) follows, completing the proof. $\square$

The assumption that the function $f^{\lambda}(X)$ is not constant on any 2-faces of the polytope $\mathcal{P}_A^p$ can be easily satisfied by a symbolic perturbation of the initial matrix C^0. Therefore, a parametric LP path exists between any two vertices of a convex polytope in general.

Lemma 3.2 *Let V^0, V be any two vertices of $\mathcal{P}_A^p$ and let $\langle C^0, X\rangle$ and $\langle C, X\rangle$ be linear functionals maximized uniquely at V^0 and V, respectively. Then the length s of any parametric LP path V^0, V^1, ..., $V^s = V$ is at most the number of edge directions in $\mathcal{P}_A^p$.*

Proof. Let V^0, V be any two vertices of $\mathcal{P}_A^p$, and let $\langle C^0, X\rangle$ and $\langle C, X\rangle$ be linear functionals maximized uniquely at V^0 and V, respectively. By Lemma 3.1, one can construct a parametric LP path V^0, V^1, ..., $V^s = V$ with parameters $\lambda^0 = 0 < \lambda^1 < \cdots < \lambda^s < 1$. The parameter λ^k is determined by the equation (2) and uniquely by any nonzero multiple of $V^{k+1} - V^k$. Thus, no two distinct edges of the same direction can appear in a parametric LP path, and the result follows. □

Combining Lemmas 2.2 and 3.2, we can now deduce Theorem 2.

Proof of Theorem 2. By Lemma 3.2 the diameter of $\mathcal{P}_A^p$ is at most the number of edge directions of $\mathcal{P}_A^p$ which, by Lemma 2.2 (see Theorem 2.3) is at most $n\binom{p}{2}$. □

One might expect that LP(1) can be solved efficiently by numerically tracing the parametric LP path from any initial vertex V^0. In fact it is possible to design such an algorithm so that it runs in strongly polynomial time. However, there is a much simpler way to solve LP(1).

Lemma 3.3 *Given a matrix $C \in \mathbb{R}^{d\times p}$, let π be any p-partition of A defined as follows: for each $a \in A$, let $\pi(a) \in \{1, \ldots, p\}$ be any index such that $\langle C, a\otimes e_{\pi(a)}\rangle = \max\{\langle C, a\otimes e_i\rangle : i \in \{1, \ldots, p\}\}$. Then A^π maximizes the linear functional $\langle C, X\rangle$ over $\mathcal{P}_A^p$. Moreover, if $\pi(a)$ is uniquely defined for every $a \in A$ then A^π is the unique vertex of $\mathcal{P}_A^p$ at which $\langle C, X\rangle$ is maximized.*

Proof. For every p-partition τ of A, the value of the linear functional $\langle C, X\rangle$ at A^τ is given by

$$\langle C, A^\tau\rangle \quad = \quad \sum_{a\in A}\langle C, a\otimes e_{\tau(a)}\rangle \quad .$$

Any p-partition π as defined in the lemma maximizes every term of this sum and hence the entire sum. Moreover, each summand $\langle C, a\otimes e_{\tau(a)}\rangle$ is uniquely maximized by π if there is a unique index $\pi(a)$ maximizing $\langle C, a\otimes e_i\rangle$, in which case the entire sum is uniquely maximized at A^π. □

It should be remarked that, in the more general context of *shaped* partitions, linear optimization is more involved and had been studied in [11].

The following corollary gives a simple way to find one vertex of $\mathcal{P}_A^p$, which will be used in the next section for the intialization of a vertex enumeration algorithm.

Corollary 3.4 *Let π^0 be the p-partition of A defined by*

$$\begin{aligned}
\pi_1^0 &:= \{a \in A : \textit{the first nonzero coordinate of } a \textit{ is positive}\}\\
\pi_2^0 &:= \{a \in A : \textit{the first nonzero coordinate of } a \textit{ is negative}\}\\
\pi_i^0 &:= \emptyset, \quad 3 \le i \le p\ ,
\end{aligned}$$

let $V^0 := A^{\pi^0}$, *and let* $C^0 = C^0(\epsilon)$ *be the* $d \times p$ *matrix defined by*

$$C^0_{ij} := \begin{cases} \epsilon^i & \text{if } j = 1, \\ -\epsilon^i & \text{if } j = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Then the linear functional $\langle C^0, X\rangle$ *is maximized uniquely at* V^0 *over* $\mathcal{P}^p_A$ *for sufficiently small* $\epsilon > 0$. *Consequently,* V^0 *is a vertex of* $\mathcal{P}^p_A$.

Proof. Consider any $a \in A$ (which is nonzero since A is free). If the first nonzero coordinate of a is positive then for all sufficiently small $\epsilon > 0$, the value $\langle C, a \otimes e_i\rangle$ is positive if $i = 1$, negative if $i = 2$, and zero for all other i. Thus, by Lemma 3.3, any π for which A^π maximizes $\langle C, X\rangle$ over $\mathcal{P}^p_A$ must have $\pi(a) = 1$. Similarly, we conclude that if the first nonzero coordinate of a is negative then any π for which A^π maximizes $\langle C, X\rangle$ over $\mathcal{P}^p_A$ must have $\pi(a) = 2$. By Lemma 3.3 it follows that $V = A^{\pi^0}$ is indeed the unique maximizer of $\langle C, X\rangle$ over $\mathcal{P}^p_A$. □

4 The adaptive algorithm: output sensitive vertex enumeration

Generation of all vertices of a given polytope is often a very demanding task, because the size of output is in general an exponential function of the size of input representation. To perform such a task, it is highly desirable to have time complexity polynomially bounded in the size of input and output, while memory requirement is kept as small as possible.

In this section, we present a polynomial algorithm which requires very little memory. To design such an algorithm, we exploit both the idea of reverse search [4, 5] and the efficient recognition of neighbors of a vertex in partition polytopes.

The main idea of reverse search is extremely simple. For a successful application of reverse search, we need a finite local search algorithm that can be initiated with any feasible object and find the optimal object in a finite sequence of operation to move from one feasible object to an "adjacent" feasible object. By reversing the finite algorithm from the optimal object in all possible ways, one can generate all feasible objects.

For the vertex enumeration of a convex polytope P, one can set a finite local search as the simplex algorithm with a finite pivot rule [6] to find an optimal vertex minimizing a given linear objective function. Below we give one particular implementation of reverse search which is appropriate for the case when P is a p-partition polytope $\mathcal{P}^p_A$. All the assumptions we shall make on P are naturally fulfilled for this special case, due to the results we proved in Section 2.

First let us assume that the graph $G(P) = (V(P), E(P))$ of a polytope P is given in the sense that the neighbor set $N(v)$ of each vertex $v \in V(P)$ is easily computable. We will make this point more precise later. For the moment, $N(v)$ is given by an oracle which lists the members in a linear order. Also, we assume that a linear function $g(x) = c^T x$ which is not constant on any edge of P is given, and the unique maximizer $s \in V(P)$ of $g(x)$ over P is also given. Define our local search function $f : V(P) \setminus \{s\} \to V(P)$ by $f(v) := \min\{v' \in N(v) : g(v') > g(v)\}$. The *trace* $T_f = (V(P), E_f)$ of the local search function f is the directed spanning tree of $G(P)$ rooted at s

whose directed edge set is defined by

$$E_f := \{(v, f(v)) : v \in V(P) \setminus \{s\}\}.$$

An enumeration algorithm follows the trace T_f from the optimal vertex s in the depth-first-search manner. Going down the tree means following an edge against its orientation, and going up is merely applying the function f. In order to make the procedure concrete, we now make the neighbor listing oracle precise. An adjacency oracle is a function $Adj(v, k)$ defined at each $v \in V(P)$ for $k = 1, \ldots, \delta$. Here δ is a given constant bounding the maximum degree of vertices in $G(P)$. There are three conditions that an adjacency oracle must satisfy:

(a) for each vertex v and each number k with $1 \leq k \leq \delta$ the oracle returns $Adj(v, k)$, a vertex adjacent to v or extraneous 0 (zero),

(b) if $Adj(v, k) = Adj(v, k') \neq 0$ for some $v \in V$, k and k', then $k = k'$,

(c) for each vertex v, $\{Adj(v, k) : Adj(v, k) \neq 0, 1 \leq k \leq \delta\} = N(v)$.

The meaning of these conditions should be clear. Essentially, an adjacency oracle must give the complete unduplicated list of neighbors of a given vertex v after evaluating the δ neighbor indices $k = 1, \ldots, \delta$.

The reverse search algorithm is uniquely determined by a quadruple (Adj,δ,s,f), see Figure 1. The complexity of the algorithm is given by the following theorem.

```
procedure ReverseSearch(Adj,δ,s,f);
    v := s; j := 0; (* j: neighbor counter *)
    repeat
        while j < δ do
            j := j + 1;
            next := Adj(v, j);
            if next ≠ 0 then
                if f(next) = v then  (* reverse traverse *)
                    v := next; j := 0
                endif
            endif
        endwhile;
        if v ≠ s then (* forward traverse *)
            u := v;    v := f(v);
            j := 0;    repeat j := j + 1 until Adj(v, j) = u (* restore j *)
        endif
    until v = s and j = δ.
```

Figure 1: The reverse search algorithm

Theorem 4.1 *[5] The time complexity of ReverseSearch is $O(\delta\ t(Adj)|V(P)| + t(f)|E(P)|)$, where $t(g)$ denotes the time to evaluate the function g.*

To apply this general result to the special case of partition polytopes, we must specify how we implement a quadruple (Adj, δ, s, f) for $P = \mathcal{P}_A^p$. Here we assume that a free point set $A \in \mathbb{Q}^{d \times n}$ is given. We can use Theorem 2.6 to define an adjacency oracle Adj by

$$Adj(A^\pi, (a,t)) := \begin{cases} A^{\tau(\pi,a,t)} & \text{if } t \neq \pi(a) \text{ and } LP(\pi,a,t) \text{ is feasible,} \\ 0 & \text{otherwise.} \end{cases}$$

The function Adj is defined for any $V \in V(\mathcal{P}_A^p)$ and any pair $(a,t) \in A \times \{1, \dots, p\}$. The neighbor index (a,t) is doubly indexed and the maximum degree bound δ can be taken to be $n \cdot p$.

To define a local search function f for $\mathcal{P}_A^p$, we need to fix a linear functional $g(X) = \langle C, X \rangle$ which is not constant on any edge of $\mathcal{P}_A^p$ and find the unique vertex S maximizing g over $\mathcal{P}_A^p$. For this we simply apply Corollary 3.4 and set $C := C^0$ and $S := V^0$ as defined in the lemma. A local search function $f(V)$ is easily fixed by using any rule to determine the "best neighbor" of a vertex v that has larger g value. Typically, one uses a lexicographic ordering of pairs(a,t) to break ties.

We are now finally in position to establish the main result of this article, Theorem 1.

Proof of Theorem 1. We only need to show that the particular implementation of reverse search we presented above does exactly what is claimed in the theorem. According to Theorem 4.1, the time complexity is $O(\delta\ t(Adj)|V(\mathcal{P}_A^p)| + t(f)|E(\mathcal{P}_A^p)|)$. Since $|E(\mathcal{P}_A^p)| \leq \frac{\delta}{2}|V(\mathcal{P}_A^p)|$ and $\delta = n \cdot p$ is polynomially bounded by the input size L, it is left for us to show that both $t(Adj)$ and $t(f)$ are polynomially bounded in L. Clearly the hardest computation in these evaluations is to check whether a given neighbor candidate of a vertex is actually a neighbor vertex. By Theorem 2.6, this is an LP problem whose input size is polynomially bounded by L, and thus is polynomially solvable in L. Finally, the reverse search algorithm stores only two vertices at once, hence the space complexity of the algorithm coincides with that of implementing Adj and f, which can be, by the polynomial solvability of LP, proportional to the input size. This completes the proof. $\square$

5 A lower bound theorem for partition polytopes

In the previous section, we provided an output sensitive algorithm that enumerates all extreme points of a partition polytope. When the polytope has much fewer vertices than the worst-case $v_{p,d}(n)$, our algorithm significantly outperforms any algorithm that runs in time polynomial in the size of the worst-case output. In this section we give a construction of partition polytopes with few vertices, demonstrating that the instances in which our algorithm is more efficient are abundant.

Recall that $l_{p,d}(n)$ denotes the minimal number of vertices of a maximal dimensional p-partition polytope $\mathcal{P}_A^p$, where A is a free set of n vectors in d-space. It is easy to see that the dimension of $\mathcal{P}_A^p$ is at most $d(p-1)$, as in every $d \times p$ matrix $A^\pi \in \mathcal{P}_A^p$ the sum of any row i is constant, $\Sigma_{j=1}^p \Sigma_{a \in \pi_j} a_i = \Sigma_{a \in A} a_i$. If we do not require $\mathcal{P}_A^p$ to be of maximal dimension then, for every n, the partition polytope $\mathcal{P}_A^p$ of the n-set $A = \{j \cdot e_1 \ : \ j = 1, \dots, n\} \subset \mathbb{R}^d$ equals the partition polytope of the single-point set $\{\binom{n+1}{2} \cdot e_1\}$ hence $\mathcal{P}_A^p$ is the $(p-1)$-simplex and has p vertices independent of n. If we do require maximal dimensionality but we do not require A to be a free set then for every n, the partition polytope $\mathcal{P}_A^p$ of the nd-set $A = \{j \cdot e_i \ : \ j = 1, \dots, n, \ i = 1, \dots, d\}$ equals the partition polytope of the d-set $\{\binom{n+1}{2} \cdot e_1, \dots, \binom{n+1}{2} \cdot e_d\}$ hence $\mathcal{P}_A^p$ is the d-fold product of the

$(p-1)$-simplex and has p^d vertices independent of n. Consequently it is natural to have a lower bound theorem for free sets and maximal dimensional partition polytopes.

To prove Theorem 3, we first compute the maximal number of separable p-partitions of a point set that lies on a straight line. An ordered p-partition of a set of n points in the d-dimensional Euclidean space $\mathbb{R}^d$ is called *separable* [1] if the convex hulls of the p parts are pairwise disjoint.

Lemma 5.1 *Let B be a free set of t points on a straight line in d-space. The number of separable p-partitions of B is at most $p\,t^{p-1}$.*

Proof. Let $B = \{b_1, b_2, \ldots, b_t\}$ be a free set of t points on a straight line indexed in order along the line. It is easy to see that if π is a separable partition of B then any nonempty part of π consists of consecutive points, i.e. is of the form $\{b_i, \ldots, b_j\}$ for some $1 \le i \le j \le n$. Thus, for $1 \le m \le p$, the number of separable p-partitions having exactly m nonempty parts is $\binom{t-1}{m-1}\binom{p}{m}m!$. Substituting $k := m-1$, we find that the total number of separable p-partitions of B is

$$\sum_{k=0}^{p-1}\binom{t-1}{k}p\binom{p-1}{k}k! \;=\; p\sum_{k=0}^{p-1}(t-1)(t-2)\cdots(t-k)\binom{p-1}{k} \;<\; p\sum_{k=0}^{p-1}(t-1)^k\binom{p-1}{k} \;=\; p\,t^{p-1}\,. \;\Box$$

Proof of Theorem 3. First, note that if a set A contains a basis $\{a_1, a_2, \ldots, a_d\}$ of $\mathbb{R}^d$ then $\mathcal{P}_A^p$ is of maximal dimension $d(p-1)$: indeed, for $i = 1, \ldots, d$ and $j = 1, \ldots, p-1$ let $\pi^{i,j}$ be the partition having two nonempty parts $\pi_j^{i,j} := \{a_i\}$ and $\pi_p^{i,j} := A \setminus \{a_i\}$; then the $d \times p$ matrices corresponding to these partitions together with the p-partition $(\emptyset, \ldots, \emptyset, A)$ form an affine basis of $\mathcal{P}_A^p$.

Now, let $A := B \cup C$, where B is a set of $t = n-d+2$ points on an affine line, say $B := \{e_1 + i \cdot e_2 \;:\; 0 \le i \le n-d+1\}$, and C is a set of $d-2$ linearly independent vectors such that $B \cup C$ is a free set spanning $\mathbb{R}^d$, say $C := \{e_3, e_4, \ldots, e_d\}$, so that $\mathcal{P}_A^p$ is of maximal dimension. By Lemma 5.1, there are at most $p(n-d+2)^{p-1}$ separable p-partitions of B. Moreover the number of p-partitions of C is clearly at most p^{d-2}, hence the number of separable partitions of A is at most $p^{d-2} \cdot p(n-d+2)^{p-1}$. Since a p-partition π is separable whenever A^π is a vertex [7], $\mathcal{P}_A^p$ has at most that many vertices and the theorem follows. $\Box$

Theorem 3 implies that for every fixed d and p we have $l_{p,d}(n) \le O(n^{p-1})$ while the worst case number of vertices $v_{p,d}(n)$ obeys the much larger lower bound $v_{p,d}(n) \ge \Omega(n^{\lfloor \frac{d-1}{2} \rfloor p})$. Thus, even under the restriction of fixed dimension d and fixed number of parts p, our new output sensitive enumeration algorithm can be far superior to the previously known algorithms of [3, 10, 12].

References

[1] N. Alon and S. Onn, Separable partitions, *Discrete Applied Mathematics*, 91:39–51, 1999.

[2] A. Andrzejak and K. Fukuda, Optimization over k-set polytopes and efficient k-set enumeration, In *Proc. 6th International Workshop on Algorithms and Data Structures (WADS'99)*, LNCS 1663, pages 1–12. Springer-Verlag, 1999.

[3] S. Aviran and S. Onn, Momentopes and the vertex complexity of partition polytopes, *Discrete and Computational Geometry*, to appear.

[4] D. Avis and K. Fukuda, A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra, *Discrete and Computational Geometry*, 8:295–313, 1992.

[5] D. Avis and K. Fukuda, Reverse search for enumeration, *Discrete Applied Mathematics*, 65:21–46, 1996.

[6] R.G. Bland, New finite pivot rules for the simplex method, *Mathematics of Operations Research*, 2:103–107, 1977.

[7] E.R. Barnes, A.J. Hoffman and U.G. Rothblum, Optimal partitions having disjoint convex and conic hulls, *Mathematical Programming*, 54:69–86, 1992.

[8] V. Chvátal, *Linear Programming*, W.H.Freeman and Company, 1983.

[9] D. Granot and U.G. Rothblum, The Pareto set of the partition bargaining game, *Games and Economic Behavior*, 3:163–182, 1991.

[10] F.K. Hwang, S. Onn and U.G. Rothblum, A polynomial time algorithm for shaped partition problems, *SIAM Journal on Optimization*, 10:70–81, 1999.

[11] F.K. Hwang, S. Onn and U.G. Rothblum, Linear-shaped partition problems, *Operations Research Letters*, 26:159–163, 2000.

[12] S. Onn and L.J. Schulman, The vector partition problem for convex objective functions, *Mathematics of Operations Research*, 26:583–590, 2001.

[13] S. Onn and B. Sturmfels, Cutting corners, *Advances in Applied Mathematics*, 23:29–48, 1999.