On Clustering Bodies: Geometry and Polyhedral Approximation

Andreas Brieden · Peter Gritzmann

Abstract The present paper studies certain classes of closed convex sets in finite-dimensional real spaces that are motivated by their application to convex maximization problems, most notably, those evolving from geometric clustering. While these optimization problems are $\mathbb{NP}$-hard in general, polynomial-time approximation algorithms can be devised whenever appropriate polyhedral approximations of their related clustering bodies are available. Here we give various structural results that lead to tight approximations.

1 Introduction

The present paper studies structural problems for certain classes of closed convex sets in finite-dimensional real spaces that are motivated by optimization problems of the form

$$\max f(\mathbf{x})$$
$$G(\mathbf{x}, y) \leq 0,$$

where

$$f : \mathbb{R}^{R_1} \to \mathbb{R} \quad \text{convex}, \qquad G : \mathbb{R}^{R_1} \times \mathbb{R}^{R_2} \to \mathbb{R}^{R_3} \quad \text{affine}.$$

A. Brieden
Universität der Bundeswehr, Werner-Heisenberg-Weg 39, 85579 Neubiberg, Germany
e-mail: andreas.brieden@unibw.de

P. Gritzmann (✉)
Zentrum Mathematik, Technische Universität München, 80290 Munich, Germany
e-mail: gritzman@ma.tum.de

Note that G is of the form

$$G(\mathbf{x}, y) = M \begin{pmatrix} \mathbf{x} \\ y \end{pmatrix} + a,$$

where M is a real $R_3 \times (R_1 + R_2)$ matrix, and $a \in \mathbb{R}^{R_3}$; thus,

$$F = \left\{ (\mathbf{x}^T, y^T)^T : G(\mathbf{x}, y) \le 0 \right\}$$

is a polyhedron. As it is well known, in general such *convex maximization* problems admit exponentially many local maxima and are computationally hard; see, e.g., [8]. In fact, by [2] even the very special case

$$R_1 = n, \qquad R_2 = 0, \qquad R_3 = m, \qquad A \in \mathbb{R}^{m \times n}, \qquad b \in \mathbb{R}^m,$$

$$f(\mathbf{x}) = f(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^2, \qquad G(\mathbf{x}) = Ax - b,$$

of maximizing the square of the Euclidean norm over a polyhedron is already $\mathbb{NP}$-hard even if the polyhedron is a parallelotope.

Our problem is, however, much more general. Since f does not depend on y, we are, in effect, faced with the problem of maximizing a convex function over the orthogonal projection of the polyhedron F in $\mathbb{R}^{R_1 + R_2}$ on the space $\mathbb{R}^{R_1}$ of the variables $\mathbf{x}$. Since the number of facets of the projection of F may be exponential in the number of facets of F, our problem statement actually allows a compact encoding of feasible regions of an exponential number of linear inequalities.

Another example that is in spirit close to our main application is given by

$$R_1 = kd, \qquad R_2 = n, \qquad R_3 = 2kd + m,$$

$$x_1, \ldots, x_k \in \mathbb{R}^d, \qquad \mathbf{x} = (x_1^T, \ldots, x_k^T)^T, \qquad A_1, \ldots, A_k \in \mathbb{R}^{d \times n},$$

$$B \in \mathbb{R}^{m \times n}, \qquad b \in \mathbb{R}^m, \qquad f(\mathbf{x}) = \sum_{i=1}^{k} f_i(x_i),$$

$$G(\mathbf{x}, y) = \left(\ldots, (x_i - A_i y)^T, -(x_i - A_i y)^T, \ldots, (By - b)^T \right)^T.$$

Here a convex objective function that is separable in $x_1, \ldots, x_k$ is maximized over a polyhedron that is specified by inequalities involving variables y that are coupled with $x_1, \ldots, x_k$ by linear equations.

In this paper we focus on special classes of such problems that are basic in geometric clustering, a prime example for practical applications being the consolidation of farmland; see [4, 5]. (We will outline this class of examples in more detail at the end of this introduction.) While these problems are $\mathbb{NP}$-hard, polynomial-time approximation algorithms can be devised whenever appropriate approximations of f by piecewise affine functions are available.

Here we will take the geometric point of view leading to the problem of approximating certain closed convex sets by polyhedra. While such questions are well studied in fixed dimension with the number of facets of the approximation polyhedra

tending to infinity (see, e.g., [10]), different asymptotics become relevant when the dimension of the space is part of the input. Then the asymptotics involve the dimension tending to infinity while the number of facets of the approximating polyhedra is bounded by a polynomial in that dimension. In fact, it is precisely the number of facets of an approximating polyhedron that gives the number of linear programs that have to be solved in the corresponding approximation algorithms for the given convex maximization problem.

The functions of particular relevance here are defined on $\mathbb{R}^{kd}$ by means of norms in $\mathbb{R}^d$ and $\mathbb{R}^{\binom{k}{2}}$, respectively. In what follows, let for some $n \in \mathbb{N}$,

$$\| \cdot \|, \qquad \mathbb{B}, \qquad \mathbb{S}$$

denote a norm in $\mathbb{R}^n$, the unit ball, and unit sphere of the Minkowski space $\mathbb{M} = (\mathbb{R}^n, \| \cdot \|)$, respectively, i.e.,

$$\mathbb{B} = \{ x \in \mathbb{R}^n : \|x\| \le 1 \}, \qquad \mathbb{S} = \{ x \in \mathbb{R}^n : \|x\| = 1 \}.$$

The notation $\mathbb{B}^n$ and $\mathbb{S}^{n-1}$ is used in order to signify the dimension of the space. The most important cases are the Minkowski spaces with ℓ_p norms. For $x = (\xi_1, \ldots, \xi_n)^T \in \mathbb{R}^n$, the ℓ_p norm $\|x\|_{(p)}$ is defined for $1 \le p \le \infty$ as follows:

$$\|x\|_{(p)} = \left(\sum_{i=1}^{n} |\xi_i|^p \right)^{1/p} \quad (1 \le p < \infty), \qquad \|x\|_{(\infty)} = \max_{1 \le i \le n} |\xi_i|.$$

The Minkowski space $\mathbb{R}^n$ endowed with an ℓ_p norm will be denoted by $\mathbb{R}^n_{(p)}$, its unit ball and unit sphere by $\mathbb{B}_{(p)}$ and $\mathbb{S}_{(p)}$ respectively. Note that the norms $\| \cdot \|_{(1)}$ and $\| \cdot \|_{(\infty)}$ are polytopal, i.e., their unit balls are polytopes; $\mathbb{B}_{(1)}$ is the regular cross polytope, while $\mathbb{B}_{(\infty)}$ is the standard cube. Other norms of particular importance are ellipsoidal norms whose unit balls are images of the Euclidean unit ball $\mathbb{B}_{(2)}$ under linear transformations.

Often, we will be dealing with different norms in different spaces simultaneously. We will then use a lower index $\diamond$ for distinction and write $\| \cdot \|$ or $\| \cdot \|_\diamond$. Similarly, $\mathbb{B}$ and $\mathbb{B}_\diamond$ will indicate the corresponding unit balls.

In the following, let $\| \cdot \|$ be an arbitrary norm on $\mathbb{R}^d$, and let $\| \cdot \|_\diamond$ be a norm on $\mathbb{R}^{\binom{k}{2}}$ that is monotone i.e.,

$$0 \le x \le y \quad \Rightarrow \quad \|x\|_\diamond \le \|y\|_\diamond,$$

where, as usual, the inequalities on the left side are meant componentwise. Note that each ℓ_p norm is monotone, but, for instance, general ellipsoidal norms are not. Take, as an example, the norm in $\mathbb{R}^2$ whose unit ball is the ellipsoid with axes

$$[-2, 2] \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad [-1, 1] \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Then the point $(2, 2)^T$ has a smaller norm than its projection $(2, 0)^T$ on the first coordinate axis.

The objective functions f that will be considered here are then defined for $x_1, \ldots, x_k \in \mathbb{R}^d$ and $\mathbf{x} = (x_1^T, \ldots, x_k^T)^T$ by

$$f(\mathbf{x}) = f(x_1, \ldots, x_k) = \left\| \left(\|x_1 - x_2\|, \|x_1 - x_3\|, \ldots, \|x_{k-1} - x_k\| \right)^T \right\|_\diamond .$$

To be more specific, let

$$R_1 = kd, \qquad r = \binom{k}{2}, \qquad \tau_{i,j} = \|x_i - x_j\| \quad (1 \le i < j \le k),$$

and let $t = (\tau_1, \ldots, \tau_r)^T \in \mathbb{R}^r$ be the vector with the coordinates $\tau_{i,j}$ arranged in lexicographically increasing order of the index pairs (i, j). The fact that the lth coordinate corresponds to the pair (i, j) will sometimes be indicated by writing $l = l(i, j)$. Also, whenever we write

$$\begin{pmatrix} x_1 - x_2 \\ x_1 - x_3 \\ \vdots \\ x_{k-1} - x_k \end{pmatrix} \in \mathbb{R}^{dr}, \qquad \begin{pmatrix} \|x_1 - x_2\| \\ \|x_1 - x_3\| \\ \vdots \\ \|x_{k-1} - x_k\| \end{pmatrix} \in \mathbb{R}^r,$$

it is always this underlying lexicographical order we are tacitly referring to. To exclude the most trivial cases, we will always assume that $d \ge 1$ and $k \ge 2$.

For $\| \cdot \|_\diamond = \| \cdot \|_{(p)}$ with $1 \le p < \infty$ and $\| \cdot \|_\diamond = \| \cdot \|_{(\infty)}$, we are in fact confronted with the objective functions $f(\mathbf{x}) = (f_p(\mathbf{x}))^{1/p}$ and $f(\mathbf{x}) = f_\infty(\mathbf{x})$, where

$$f_p(\mathbf{x}) = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \|x_i - x_j\|^p, \qquad f_\infty(\mathbf{x}) = \max_{i,j=1,\ldots,k} \|x_i - x_j\|,$$

respectively. Also, when $\kappa_1, \ldots, \kappa_k > 0$, the weighted objective function

$$\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \kappa_i \kappa_j \|x_i - x_j\|^2$$

corresponds to the axis oriented ellipsoidal norm $\| \cdot \|_\diamond$ with half axes lengths $\sqrt{1/(\kappa_i \kappa_j)}$.

In geometric clustering, such objective functions are used to model real-world problems where clusters are built with respect to geometric proximity but under certain polyhedral balancing constraints. A major example where the results of the present paper are utilized is that of the consolidation of farmland based on modern land-lease initiatives. Here k is the number of farmers in a certain region, who cultivate a larger number m of lots scattered across a wider agricultural region. The goal is to redistribute the lots so as to reduce distances of the lots that are cultivated by the same farmer; see [4, 5] for details and [3] for some further background information. Naturally, the total size of each farmer's land should not change much in the course of redistribution, neither should its value or other possibly relevant parameters. While these constraints can be reasonably well modelled by linear inequalities, the objective

is to maximize a convex function of the kind introduced in the previous paragraph. Here the points $x_1, \ldots, x_k$ correspond to the centers of gravity of all lots assigned to each of the k farmers, respectively, and the objective is simply to "push them apart."

As we are taking here a geometric point of view, the prime objects of our study will be the level sets

$$C = \left\{ \mathbf{x} = \left(x_1^T, \ldots, x_k^T\right)^T \in \mathbb{R}^{kd} : f(\mathbf{x}) \leq 1 \right\}$$
$$= \left\{ \mathbf{x} \in \mathbb{R}^{kd} : \left\| \left(\|x_1 - x_2\|, \|x_1 - x_3\|, \ldots, \|x_{k-1} - x_k\| \right)^T \right\|_\diamond \leq 1 \right\},$$

the so-called *clustering bodies*. These and related convex sets are studied with a view towards their applications in new practical models for clustering under balancing constraints. However, their geometry is remarkably rich and makes the clustering bodies an interesting class of geometric objects in their own right. For instance, certain zonotopes fall into this class, but also bodies in some $\mathbb{R}^{dk}$, k of whose d-dimensional coordinate sections are Euclidean balls while "diagonal" k-sections from a certain d-parameter class are duals of permutahedra.

In the following, the "generic notation" C will be used in the general case, i.e., when the occurring norms are not particularly specialized. Of course, after first studying the structure of general clustering bodies, we will later derive additional results for some specific cases, most notably those where $\| \cdot \|_\diamond = \| \cdot \|_{(1)}$ while $\| \cdot \|$ is still arbitrary or where both norms $\| \cdot \|$ and $\| \cdot \|_\diamond$ are certain ℓ_p norms.

As to the connection to our optimization problem, note that, for the given instance, the task can be interpreted as to find the minimal nonnegative λ such that the corresponding polyhedron F is contained in $\lambda C \times \mathbb{R}^{R_2}$ or, alternatively, λC contains the orthogonal projection of F on $\mathbb{R}^{R_1} \times \{0\}^{R_2}$. See [9] for a survey on containment problems.

The paper is organized as follows. Section 2 studies the geometry of general clustering bodies C, while Sects. 3 and 4 derive additional results particularly when the determining norms are specified to certain ℓ_p norms. Section 5 considers polynomial-time polyhedral approximations.

2 General Geometric Structure

In the following, we study the geometry of the clustering bodies C. We begin with two preparatory lemmas that deal with the basic underlying set

$$K = \left\{ \mathbf{v} = (v_1^T, \ldots, v_r^T)^T : \left\| \left(\|v_1\|, \ldots, \|v_r\| \right)^T \right\|_\diamond \leq 1 \right\},$$

which, by the first lemma, is a union of certain Cartesian products.

Lemma 2.1 *With* $t = (\tau_1, \ldots, \tau_r)^T$,

$$K = \bigcup_{t \in \mathbb{B}_\diamond \cap [0, \infty[^r} (\tau_1 \mathbb{S} \times \cdots \times \tau_r \mathbb{S}).$$

Proof Let $\mathbf{v} = (v_1^T, \ldots, v_r^T)^T \in K$, $\tau_i = \|v_i\|$, $s_i \in \mathbb{S}$ with $\tau_i s_i = v_i$ for $i = 1, \ldots, r$, and $t = (\tau_1, \ldots, \tau_r)^T$. Then $t \in \mathbb{B}_\diamond$ and $v_i \in \tau_i \mathbb{S}$; hence, $\mathbf{v} \in \tau_1 \mathbb{S} \times \cdots \times \tau_r \mathbb{S}$.

Conversely, let $t \in \mathbb{B}_\diamond \cap [0, \infty[^r$, $s_i \in \mathbb{S}$, and $v_i = \tau_i s_i$ for $i = 1, \ldots, r$. Then

$$\left\| \left(\|v_1\|, \ldots, \|v_r\| \right)^T \right\|_\diamond = \left\| (\tau_1, \ldots, \tau_r)^T \right\|_\diamond = \|t\|_\diamond \le 1.$$

Hence, $\mathbf{v} = (v_1^T, \ldots, v_r^T)^T \in K$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

By Lemma 2.1 it is by no means clear whether K is convex. As we will see later, this is not always the case. However, the following lemma shows that K is convex whenever $\| \cdot \|_\diamond$ is monotone; in particular, the function $g : \mathbb{R}^{dr} \to \mathbb{R}$ defined by

$$g(\mathbf{v}) = g(v_1, \ldots, v_r) = \left\| \left(\|v_1\|, \ldots, \|v_r\| \right)^T \right\|_\diamond$$

is then a norm.

Lemma 2.2 *Let $\| \cdot \|_\diamond$ be monotone. Then K has the following properties*:

(1) *K is symmetric about the origin, compact, convex, and* $\mathrm{int}(K) \neq \emptyset$.
(2) *Let $S \subset \mathbb{R}^d$ and $T \subset \mathbb{R}^r$ be such that*

$$\mathbb{B} = \mathrm{conv}(S), \qquad \mathbb{B}_\diamond \cap [0, \infty[^r = \mathrm{conv}(T).$$

Then, with $t = (\tau_1, \ldots, \tau_r)^T$,

$$K = \mathrm{conv}\left(\bigcup_{t \in T} (\tau_1 S \times \cdots \times \tau_r S) \right).$$

(3) *K is a polytope if and only if $\mathbb{B}$ and $\mathbb{B}_\diamond \cap [0, \infty[^r$ are polytopes.*

Proof To prove (1) we show that g is a norm. Of course, g is positive definite and homogeneous. To verify the triangle inequality, observe that, by the monotonicity of $\| \cdot \|_\diamond$, we have

$$g(\mathbf{v} + \mathbf{w}) = \left\| \left(\|v_1 + w_1\|, \ldots, \|v_r + w_r\| \right)^T \right\|_\diamond$$

$$\le \left\| \left(\|v_1\| + \|w_1\|, \ldots, \|v_r\| + \|w_r\| \right)^T \right\|_\diamond \le g(\mathbf{v}) + g(\mathbf{w}).$$

Hence, K is the unit ball associated with the norm g.

(2) Since K is convex, the inclusion "$\supset$" is clear.

Now, let $\mathbf{v} = (v_1^T, \ldots, v_r^T)^T \in K$. Then, of course, $(\|v_1\|, \ldots, \|v_r\|)^T \in \mathbb{B}_\diamond \cap [0, \infty[^r$. Hence, by Caratheodory's theorem, there exist $t_0, \ldots, t_r \in T$ and $\lambda_0, \ldots, \lambda_r \in [0, 1]$ with $\sum_{i=0}^r \lambda_i = 1$ such that

$$\left(\|v_1\|, \ldots, \|v_r\| \right)^T = \sum_{i=0}^r \lambda_i t_i.$$

For $i = 0, \ldots, r$, let $t_i = (\tau_{i,1}, \ldots, \tau_{i,r})^T$, and let $j \in \{1, \ldots, r\}$. Then,

$$v_j \in \sum_{i=0}^{r} \lambda_i \tau_{i,j} \mathbb{B}.$$

Hence, there exist $s_{j,0}, \ldots, s_{j,d} \in S$ and $\mu_{j,0}, \ldots, \mu_{j,d} \in [0, 1]$ with $\sum_{l=0}^{d} \mu_{j,l} = 1$ such that

$$v_j = \sum_{i=0}^{r} \lambda_i \tau_{i,j} \sum_{l=0}^{d} \mu_{j,l} s_{j,l} = \sum_{i=0}^{r} \sum_{l=0}^{d} \lambda_i \mu_{j,l} \tau_{i,j} s_{j,l}.$$

Using the fact that $\sum_{l=0}^{d} \mu_{j,l} = 1$ for each $j = 1, \ldots, r$, this implies

$$\begin{pmatrix} v_1 \\ \vdots \\ v_r \end{pmatrix} = \sum_{i=0}^{r} \sum_{l_1=0}^{d} \cdots \cdot \sum_{l_r=0}^{d} \lambda_i \left(\prod_{j=1}^{r} \mu_{j,l_j} \right) \begin{pmatrix} \tau_{i,1} s_{1,l_1} \\ \vdots \\ \tau_{i,r} s_{r,l_r} \end{pmatrix}.$$

Since

$$\sum_{i=0}^{r} \sum_{l_1=0}^{d} \cdots \cdot \sum_{l_r=0}^{d} \lambda_i \prod_{j=1}^{r} \mu_{j,l_j} = \sum_{i=0}^{r} \lambda_i \left(\sum_{l_1=0}^{d} \mu_{1,l_1} \left(\cdots \cdot \left(\sum_{l_r=0}^{d} \mu_{r,l_r} \right) \cdots \right) \right) = 1,$$

this is the desired convex combination of vectors of the specified form.

(3) If $\mathbb{B}$ and $\mathbb{B}_\diamond \cap [0, \infty[^r$ are polytopes, then there exist finite sets $S \subset \mathbb{R}^d$ and $T \subset \mathbb{R}^r$ such that

$$\mathbb{B} = \mathrm{conv}(S), \qquad \mathbb{B}_\diamond \cap [0, \infty[^r = \mathrm{conv}(T).$$

Hence, by (2), K is a polytope. So suppose now that K is a polytope. Then, according to Lemma 2.1, K is the convex hull of finitely many, say m, vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of the form

$$\mathbf{v}_i = \left(\tau_{i,1} s_{i,1}^T, \ldots, \tau_{i,r} s_{i,r}^T \right)^T,$$

where for $i = 1, \ldots, m$ and $j = 1, \ldots, r$,

$$s_{i,j} \in \mathbb{S}, \qquad t_i = (\tau_{i,1}, \ldots, \tau_{i,r})^T \in \mathbb{B}_\diamond \cap [0, \infty[^r.$$

Let $s \in \mathbb{B}$. Further, let $t \in \mathbb{B}_\diamond \cap [0, \infty[^r$ be such that its first coordinate τ is maximal. Then, of course, $\tau > 0$, and, by the monotonicity of $\| \cdot \|_\diamond$ and Lemma 2.1, in particular, $(\tau s^T, 0, \ldots, 0)^T \in K$. Thus,

$$\tau s \in \mathrm{conv}\{\tau_{1,1} s_{1,1}, \ldots, \tau_{m,1} s_{m,1}\}.$$

Hence,

$$\mathbb{B} \subset \mathrm{conv}\left\{ \frac{\tau_{1,1}}{\tau} s_{1,1}, \ldots, \frac{\tau_{m,1}}{\tau} s_{m,1} \right\} \subset \mathbb{B};$$

thus, $\mathbb{B}$ is a polytope.

Now, let $t = (\tau_1, \ldots, \tau_r)^T \in \mathbb{B}_\diamond$ with $t \geq 0$. Further, let $s_1, \ldots, s_r \in \mathbb{S}$. Then, according to Lemma 2.1, $\mathbf{v} = (\tau_1 s_1^T, \ldots, \tau_r s_r^T)^T \in K$, and there are $\lambda_1, \ldots, \lambda_m \in [0, 1]$ with $\sum_{i=1}^m \lambda_i = 1$ such that $\mathbf{v} = \sum_{i=1}^m \lambda_i \mathbf{v}_i$. In particular, we have for $j = 1, \ldots, r$,

$$\tau_j = \|\tau_j s_j\| = \left\| \sum_{i=1}^m \lambda_i \tau_{i,j} s_{i,j} \right\| \leq \sum_{i=1}^m \lambda_i \tau_{i,j} \|s_{i,j}\| = \sum_{i=1}^m \lambda_i \tau_{i,j};$$

hence,

$$t \leq \sum_{i=1}^m \lambda_i t_i.$$

Therefore,

$$\mathbb{B}_\diamond \cap [0, \infty[^r \subset \mathrm{conv}\big(\{t_1, \ldots, t_m\}\big) +]-\infty, 0]^r.$$

Since $\| \cdot \|_\diamond$ is monotone, this implies the asserted polytopality. $\qquad\square$

Next we derive some simple properties of C. Let Π_k denote the symmetric group on k-elements. We will regard $\pi \in \Pi_k$ as a permutation on whatever underlying set of k elements is involved.

Lemma 2.3 *Let $\| \cdot \|_\diamond$ be monotone. Then C has the following properties*:

(1) $0 \in C$ *and C is convex and closed.*
(2) *C is invariant under translations of $\mathbb{R}^d$, i.e.,*

$$f(x_1, \ldots, x_k) = f(a + x_1, \ldots, a + x_k)$$

for $a \in \mathbb{R}^d$.
(3) *C is invariant under all (affine or surjective) isometries of $\mathbb{R}^d$, i.e.,*

$$f(x_1, \ldots, x_k) = f\big(\sigma(x_1), \ldots, \sigma(x_k)\big)$$

for any affine or surjective mapping $\sigma : \mathbb{R}^d \to \mathbb{R}^d$ with $\|\sigma(x)\| = \|x\|$ for all $x \in \mathbb{R}^d$. In particular, $C = -C$.
(4) *If $\| \cdot \|_\diamond$ is invariant under permutations of coordinates, then C is invariant under permutations of $x_1, \ldots, x_k$, i.e.,*

$$f(x_1, \ldots, x_k) = f(x_{\pi(1)}, \ldots, x_{\pi(k)})$$

for any permutation $\pi \in \Pi_k$.
(5) *The lineality space of C is the d-dimensional subspace*

$$\mathrm{ls}(C) = \big\{ \mathbf{x} = (x_1^T, \ldots, x_k^T)^T : x_1 = \cdots = x_k \big\}.$$

(6) *Let*

$$L := \left\{ \mathbf{x} = (x_1^T, \ldots, x_k^T)^T : \sum_{i=1}^k x_i = 0 \right\}, \qquad B := C \cap L.$$

Then $L^\perp = \mathrm{ls}(C)$, and B is centrally symmetric, compact, and convex.

Proof (1) Of course, $f(0) = 0$. Further, note that $f(\mathbf{x})$ is the composition of the function $g : \mathbb{R}^{dr} \to \mathbb{R}$ defined by

$$g(\mathbf{v}) = g(v_1, \ldots, v_r) = \left\| \left(\|v_1\|, \ldots, \|v_r\| \right)^T \right\|_\diamond$$

already before the previous lemma and the linear function induced by the matrix consisting of r blocks of the form $(0_d, \ldots, 0_d, I_d, 0_d, \ldots, 0_d, -I_d, 0_d, \ldots, 0_d)$, where I_d and 0_d denote the $d \times d$ identity and zero matrices, respectively.

Since, by Lemma 2.2, g is convex, the function f is convex on $\mathbb{R}^{dk}$ and thus continuous. So, as the preimage of the closed interval $]-\infty, 1]$, or, which is the same here, of $[0, 1]$, under f, C is closed and convex.

(2) and (4) are obvious. (3) is clear for affine isometries; and by the Mazur–Ulam Theorem, surjective isometries are affine; see, e.g., [1, 14.1].

(5) Let $S = \{\mathbf{x} : x_1 = \cdots = x_k\}$. Clearly, $S \subset \mathrm{ls}(C)$. Suppose now that $\mathbf{x} \notin S$. Then there are $i, j \in \{1, \ldots, k\}$ with $i < j$ and $x_i \neq x_j$. Let $y_{i,j} \in \mathbb{R}^r$ be the vector whose components are all 0 except for the lth component with $l = l(i, j)$, i.e., that corresponding to (i, j) which is $\|x_i - x_j\|$. By monotonicity it follows for $\lambda > 1/\|y_{i,j}\|_\diamond$, that $\lambda \mathbf{x} \notin C$, and hence $\mathbf{x} \notin \mathrm{ls}(C)$.

(6) Since the other parts are clear, it remains to show that B is bounded. Suppose that B is unbounded. Then, since $0 \in B$ and B is closed, there is a vector $\mathbf{x} \neq 0$ such that $[0, \infty[\mathbf{x} \subset B$. By central symmetry of B, this implies $\mathbb{R}\mathbf{x} \subset B$, i.e., $\mathbf{x} \in \mathrm{ls}(B)$. But then $\mathbf{x} \in L \cap \mathrm{ls}(C)$, which implies $\mathbf{x} = 0$, a contradiction. $\qquad \square$

Let us remark that it is the condition that $\|\cdot\|_\diamond$ is monotone which guarantees the convexity of C. As an example, let $d = 1$, $k = 3$, and let

$$C = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : \left\| \begin{pmatrix} |x_1 - x_2| \\ |x_1 - x_3| \\ |x_2 - x_3| \end{pmatrix} \right\|_\diamond \leq 1 \right\}.$$

Specifically, let $\|\cdot\|_\diamond$ be the ellipsoidal norm

$$\|\mathbf{v}\|_\diamond = \mathbf{v}^T E \mathbf{v} \quad \text{with } E = \frac{1}{5} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 7 & -7 \\ 1 & -7 & 8 \end{pmatrix}.$$

Then $(1, 1, 0)^T, (1, -1, 0)^T \in C$; in fact, the corresponding vectors $(|x_1 - x_2|, |x_1 - x_3|, |x_2 - x_3|)^T$ are $(0, 1, 1)^T$ and $(2, 1, 1)^T$, and have $\|\cdot\|_\diamond$-norms $1/5$ and 1. However,

$$\frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin C$$

since $\|(1, 1, 0)^T\|_\diamond = 6/5$. So, C is not convex.

Hence, in the following we will always assume that $\|\cdot\|_\diamond$ is monotone.

Note that, by Lemma 2.3, C defines a seminorm in $\mathbb{R}^{dk}$, while B is the unit ball of a norm in L.

Next, we describe the sets C via polarity. Recall first, that with each Minkowski space $\mathbb{M}$, there is an associated conjugate space $\mathbb{M}^*$ whose points are the linear functionals on $\mathbb{M}$. The norm $\|\cdot\|^*$ on $\mathbb{M}^*$ is defined by $\|y\|^* = \max_{x \in \mathbb{B}} y(x)$ for $y \in \mathbb{M}^*$; its unit ball and sphere are denoted by $\mathbb{B}^*$ and $\mathbb{S}^*$, respectively. (Note that, by the compactness of $\mathbb{B}$ and the continuity of $\langle \cdot, y \rangle$, the maximum is actually attained.) As an example, note that the spaces $\mathbb{R}^n_{(p)}$ and $\mathbb{R}^n_{(q)}$ are conjugate to each other, where for $p \in [1, \infty]$, the number q is defined by the condition

$$\frac{1}{p} + \frac{1}{q} = 1$$

(with the understanding that $1/\infty = 0$). The most important spaces of this sort are the self-conjugate Euclidean space $\mathbb{R}^n_{(2)}$ and the mutually conjugate spaces $\mathbb{R}^n_{(1)}$ and $\mathbb{R}^n_{(\infty)}$. Note that, in general, the monotonicity of $\|\cdot\|$ does not imply that $\|\cdot\|^*$ is monotone.

The usual bilinear form $\langle , \rangle$ on $\mathbb{M} \times \mathbb{M}^*$ is given by $\langle x, y \rangle = y(x)$. When the points x and y are expressed as coordinate vectors $x = (\xi_1, \ldots, \xi_n)^T$ and $y = (\eta_1, \ldots, \eta_n)^T$ in the standard conjugate bases, then, of course,

$$\langle x, y \rangle = \sum_{i=1}^{n} \xi_i \eta_i.$$

By polarity we have

$$\langle x, y \rangle \leq \|x\| \cdot \|y\|^*.$$

Note that for each $y \in \mathbb{S}^*$, there exists $x \in \mathbb{S}$ for which the equality is attained.

For $X \subset \mathbb{M}$, the polar X° is given by

$$X^\circ = \left\{ y \in \mathbb{M}^* : x \in X \Rightarrow \langle x, y \rangle \leq 1 \right\}.$$

Note that, in particular, $\mathbb{B}^* = \mathbb{B}^\circ$. Also, if X is polyhedral, then so is X°. Polars are defined analogously for subsets of $\mathbb{M}^*$ and, by identifying $(\mathbb{M}^*)^*$ with $\mathbb{M}$ in the usual way, are regarded as subsets of $\mathbb{M}$. We use the fact that if X is a closed convex set containing the origin, then $(X^\circ)^\circ = X$. Whenever appropriate, we will also identify $(\mathbb{R}^n, \|\cdot\|)^*$ with $(\mathbb{R}^n, \|\cdot\|^*)$. For more background information on convex analysis and specifically on polarity, see [11].

Now we consider the polar of the unit ball K of the Minkowski space $(\mathbb{R}^{dr}, g)$.

Lemma 2.4 *Let $\|\cdot\|_\diamond$ and $\|\cdot\|_\diamond^*$ be monotone. Then*

$$K^\circ = \left\{ \mathbf{w} = \left(w_1^T, \ldots, w_r^T \right)^T : \left\| \left(\|w_1\|^*, \ldots, \|w_r\|^* \right)^T \right\|_\diamond^* \leq 1 \right\}.$$

Proof Let

$$W = \left\{ \mathbf{w} = \left(w_1^T, \ldots, w_r^T \right)^T : \left\| \left(\|w_1\|^*, \ldots, \|w_r\|^* \right)^T \right\|_\diamond^* \leq 1 \right\}.$$

With the aid of polarity applied to $\| \cdot \|$ and $\| \cdot \|_\diamond$, we obtain

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{r} \langle v_i, w_i \rangle \leq \sum_{i=1}^{r} \| v_i \| \cdot \| w_i \|^*$$

$$= \left\langle \left(\| v_1 \|, \ldots, \| v_r \| \right)^T, \left(\| w_1 \|^*, \ldots, \| w_r \|^* \right)^T \right\rangle$$

$$\leq \left\| \left(\| v_1 \|, \ldots, \| v_r \| \right)^T \right\|_\diamond \cdot \left\| \left(\| w_1 \|^*, \ldots, \| w_r \|^* \right)^T \right\|_\diamond^*.$$

So, for $\mathbf{v} \in K$ and $\mathbf{w} \in W$, we have $\langle \mathbf{v}, \mathbf{w} \rangle \leq 1$. Hence, $W \subset K^\circ$.

Now, let $\mathbf{w} = (w_1^T, \ldots, w_r^T)^T \in K^\circ$. Applying the general fact that for each $y \in \mathbb{M}^*$, the equality in $\langle x, y \rangle \leq \| x \| \cdot \| y \|^*$ is attained for some $x \in \mathbb{S}$ to the two pairs of conjugate spaces $(\mathbb{R}^d, \| \cdot \|)$, $(\mathbb{R}^d, \| \cdot \|^*)$ and $(\mathbb{R}^r, \| \cdot \|)$, $(\mathbb{R}^r, \| \cdot \|^*)$, we see that there exists a vector $\hat{\mathbf{v}} = (\hat{v}_1^T, \ldots, \hat{v}_r^T)^T \in \mathrm{bd}(K)$ such that

$$\langle \hat{\mathbf{v}}, \mathbf{w} \rangle = \left\| \left(\| \hat{v}_1 \|, \ldots, \| \hat{v}_r \| \right)^T \right\|_\diamond \cdot \left\| \left(\| w_1 \|^*, \ldots, \| w_r \|^* \right)^T \right\|_\diamond^*.$$

Hence,

$$\left\| \left(\| w_1 \|^*, \ldots, \| w_r \|^* \right)^T \right\|_\diamond^* \leq 1,$$

i.e., $\mathbf{w} \in W$. $\square$

Now, let A denote the $dr \times dk$ matrix that maps $\mathbf{x} = (x_1^T, \ldots, x_k^T)^T$ to the vector of pairwise differences $x_i - x_j$, ordered again according to our convention, i.e., lexicographically increasing with respect to (i, j).

In the following, we will always use the notation u_i to denote the ith standard unit vector in whatever space $\mathbb{R}^n$ (with $i \leq n$) we are in.

With $I = I_d$, the matrix A is of the form

$$A = \begin{pmatrix} I & -I & 0 & \cdots & 0 & 0 \\ I & 0 & -I & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ I & 0 & 0 & \cdots & -I & 0 \\ I & 0 & 0 & \cdots & 0 & -I \\ 0 & I & -I & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & I & 0 & \cdots & -I & 0 \\ 0 & I & 0 & \cdots & 0 & -I \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I & -I \end{pmatrix}.$$

In order to emphasize the dimension d of the "component vectors" x_i, we will sometimes write A_d. In particular, A_1 will be the corresponding $r \times k$ matrix.

Note that A_d is the Kronecker product $A_1 \otimes I_d$ with each component $\alpha_{i,j}$ of A_1 replaced by $\alpha_{i,j} I_d$, i.e., A_d is obtained from A_1 by replacing each entry 0 by the $d \times d$ zero matrix 0_d, each entry 1 by I_d, and each entry -1 by $-I_d$.

In particular, A comprises a product structure with respect to $\mathbb{R}^d$. More precisely, let $x_i = (\xi_{i,1}, \ldots, \xi_{i,d})^T$ for $i = 1, \ldots, k$,

$$y_j = (\xi_{1,j}, \ldots, \xi_{k,j})^T$$

for $j = 1, \ldots, d$, and $\mathbf{y} = (y_1^T, \ldots, y_d^T)^T$. Then, up to a reordering of components, $A_d(x_1^T, \ldots, x_k^T)^T$ coincides with $((A_1 y_1)^T, \ldots, (A_1 y_d)^T)^T$. The lth row with $l = (i-1)d + j$ in the standard ordering (which corresponds to the block i with $i = 1, \ldots, r$ and the jth coordinate with $j = 1, \ldots, d$) is replaced by the mth row with $m = (j-1)r + i$. Hence this reordering is facilitated by the permutation matrix $P \in \mathbb{R}^{dr \times dr}$ whose $((j-1)r + i)$th row is $u_{(i-1)d+j}$. Then

$$P A \mathbf{x} = \begin{pmatrix} A_1 \\ \vdots \\ A_1 \end{pmatrix} \mathbf{y}.$$

Observe that

$$A^{-1} K = \{z : Az \in K\} = C.$$

Theorem 2.5

$$C^\circ = A^T K^\circ, \qquad B^\circ = A^T K^\circ + \mathrm{ls}(C).$$

Proof By polarity and Lemma 2.3,

$$B^\circ = (C \cap L)^\circ = C^\circ + L^\circ = C^\circ + L^\perp = C^\circ + \mathrm{ls}(C).$$

Hence it suffices to prove the first statement. We have

$$(A^T K^\circ)^\circ = \{x : y \in A^T K^\circ \Rightarrow \langle x, y \rangle \le 1\} = \{x : z \in K^\circ \Rightarrow \langle Ax, z \rangle \le 1\}$$
$$= A^{-1}\{w : z \in K^\circ \Rightarrow \langle w, z \rangle \le 1\} = A^{-1} K = C.$$

Since $A^T K^\circ$ is closed and contains the origin, we have $((A^T K^\circ)^\circ)^\circ = A^T K^\circ$, which completes the proof of the theorem. $\qquad \square$

The following corollary should be read with a view towards Lemma 2.2.

Corollary 2.6 *If K is polyhedral, then so are C and B.*

Proof The assertion follows directly from Theorem 2.5 in conjunction with the facts that the polar and the linear image of a polyhedron are themselves polyhedral. $\qquad \square$

3 Minkowski Sums and Permutahedra

In the following, we will build on the results for general clustering bodies of the previous section to gain further structural insight when $\| \cdot \|_\diamond$ is specialized to certain

ℓ_p norms. The specific situation of $\| \cdot \|_\diamond = \| \cdot \|_{(p)}$ will in the following be signified by the subscript p, i.e.,

$$C_p = \left\{ \mathbf{x} = \left(x_1^T, \ldots, x_k^T \right)^T \in \mathbb{R}^{kd} : \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \| x_i - x_j \|^p \leq 1 \right\}$$

for $1 \leq p < \infty$, and

$$C_\infty = \left\{ \mathbf{x} = \left(x_1^T, \ldots, x_k^T \right)^T \in \mathbb{R}^{kd} : \max_{i,j=1,\ldots,k} \| x_i - x_j \| \leq 1 \right\}.$$

(Note, that taking the pth root on the left side of the defining inequality for C_p does not change the set. Hence, C_p coincides indeed with C for $\| \cdot \|_\diamond = \| \cdot \|_{(p)}$.) In this section, $\| \cdot \|$ is still an arbitrary norm; Sect. 4 will also restrict $\| \cdot \|$.

As we will see, the most interesting case is that of $\| \cdot \|_\diamond = \| \cdot \|_{(1)}$. In particular, C_1° is the Minkowski sum of certain balls. More precisely, let for $l = 1, \ldots, r$,

$$\mathbb{B}_l^* = \left\{ \mathbf{w} = \left(w_1^T, \ldots, w_r^T \right)^T \in \mathbb{R}^{dr} : w_l \in \mathbb{B}^* \wedge w_i = 0 \, (i \neq l) \right\}.$$

Hence, $\mathbb{B}_l^*$ is just $\mathbb{B}^*$ embedded into the lth coordinate $\mathbb{R}^d$ of $\mathbb{R}^{dr}$. Further, set for $i, j \in \{1, \ldots, k\}$ with $i < j$,

$$\mathbb{D}_{i,j}^* = \left\{ \mathbf{y} = \left(y_1^T, \ldots, y_k^T \right)^T \in \mathbb{R}^{dk} : y_i, y_j \in \mathbb{B}^* \wedge (y_i + y_j = 0) \wedge y_r = 0 \, (r \neq i, j) \right\}.$$

Lemma 3.1 *Let $i, j \in \{1, \ldots, k\}$ with $i < j$, and $l = l(i, j)$. Then*

$$A^T \mathbb{B}_l^* = \mathbb{D}_{i,j}^*.$$

Proof The columns of the matrix A^T fall into r blocks of size $dk \times d$ of the form

$$(0_d, \ldots, 0_d, I_d, 0_d, \ldots, 0_d, -I_d, 0_d, \ldots, 0_d)^T.$$

In the lth such block, the entries I_d and $-I_d$ are at the positions i and j, respectively. Hence, for a vector $\mathbf{w} = (0, \ldots, 0, w_l^T, 0, \ldots, 0)^T \in \mathbb{B}_l^*$, we have

$$A^T \mathbf{w} = \left(0, \ldots, 0, w_l^T, 0, \ldots, 0, -w_l^T, 0, \ldots, 0 \right)^T \in \mathbb{D}_{i,j}^*.$$

The assertion follows now from that fact that every vector of $\mathbb{D}_{i,j}^*$ is obtained that way.   □

The following theorem characterizes the structure of C_1 and C_∞.

Theorem 3.2

$$C_1^\circ = A^T (\mathbb{B}^*)^r = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \mathbb{D}_{i,j}^*, \qquad C_\infty^\circ = \text{conv} \left(\bigcup_{i=1}^{k-1} \bigcup_{j=i+1}^{k} \mathbb{D}_{i,j}^* \right).$$

Proof By Theorem 2.5 and Lemmas 2.4 and 3.1, we have

$$C_1^\circ = A^T \left\{ \left(v_1^T, \ldots, v_r^T\right)^T : \sum_{i=1}^r \|v_i\| \le 1 \right\}^\circ$$

$$= A^T \left\{ \left(w_1^T, \ldots, w_r^T\right)^T : \max_{i=1,\ldots,r} \|w_i\|^* \le 1 \right\}$$

$$= A^T (\mathbb{B}^*)^r = A^T (\mathbb{B}_1^* + \cdots + \mathbb{B}_r^*) = A^T \mathbb{B}_1^* + \cdots + A^T \mathbb{B}_r^*$$

$$= \sum_{\substack{i,j=1 \\ i<j}}^k \mathbb{D}_{i,j}^*.$$

Similarly,

$$C_\infty^\circ = A^T \left\{ \left(v_1^T, \ldots, v_r^T\right)^T : \max_{i=1,\ldots,r} \|v_i\| \le 1 \right\}^\circ$$

$$= A^T \left\{ \left(w_1^T, \ldots, w_r^T\right)^T : \sum_{i=1}^r \|w_i\|^* \le 1 \right\} = A^T \operatorname{conv}(\mathbb{B}_1^* \cup \cdots \cup \mathbb{B}_r^*)$$

$$= \operatorname{conv}\left(A^T \mathbb{B}_1^* \cup \cdots \cup A^T \mathbb{B}_r^*\right) = \operatorname{conv}\left(\bigcup_{\substack{i=1,j \\ i<j}}^k \mathbb{D}_{i,j}^* \right),$$

completing the proof of the theorem.   □

Except for their specific embedding, all $\mathbb{D}_{i,j}^*$ are of the form

$$\hat{\mathbb{D}}^* = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} : x \in \mathbb{B}^* \right\}.$$

Hence, $\hat{\mathbb{D}}^*$ is the image of $\sqrt{2}\mathbb{B}^* \times \{0\}^d$ under the rotation

$$\frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix},$$

i.e., C_1° is the Minkowski sum and C_∞° the convex hull of balls $\sqrt{2}\mathbb{B}^*$ that are embedded into all possible "diagonal spaces."

In particular, if $\mathbb{B}$ is the Euclidean unit ball, then C_1° is the Minkowski sum of d-dimensional Euclidean balls "diagonally embedded" in $(2d)$-dimensional coordinate spaces. Also, if $\mathbb{B}^*$ is a zonotope, then so is C_1°.

It is clear that C_1 has d-dimensional sections similar to $\mathbb{B}$. In fact, we have the following simple theorem.

Theorem 3.3 *Let* $i_0 \in \{1, \ldots, k\}$ *and* $\mathbb{R}_{i_0}^{dk} = 0^{d(i_0-1)} \times \mathbb{R}^d \times 0^{d(k-i_0)}$. *Then*

$$C_1 \cap \mathbb{R}_{i_0}^{dk} = \frac{1}{k-1} \left(0^{d(i_0-1)} \times \mathbb{B} \times 0^{d(k-i_0)} \right).$$

Proof Let $\mathbf{x} \in \mathbb{R}^{dk}_{i_0}$. Then $\mathbf{x}$ is of the form $(0, \ldots, 0, x^T_{i_0}, 0, \ldots, 0)^T$, and hence,

$$f_1(\mathbf{x}) = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \|x_i - x_j\| = (k-1)\|x_{i_0}\|.$$

Thus,

$$f_1(\mathbf{x}) \le 1 \quad \Leftrightarrow \quad x_{i_0} \in \frac{1}{k-1}\mathbb{B},$$

which proves the assertion. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

In particular, C_1 contains the convex hull of $\|\cdot\|$-balls of radius $\frac{1}{k-1}$ embedded in each of the k coordinate $\mathbb{R}^d$'s.

The next lemma studies certain k-dimensional sections of C_1. Let for $x \in \mathbb{R}^d \setminus \{0\}$,

$$T_x = \mathbb{R}x \times \cdots \times \mathbb{R}x.$$

Note that $\mathbb{R}(x^T, \ldots, x^T)^T \subset T_x \cap \mathrm{ls}(C)$.

In the following, we will use the notation $x_{i,j}$ for $x \in \mathbb{R}^d$, $i = 1, \ldots, k-1$, and $j = i+1, \ldots, k$ to signify the vector

$$x_{i,j} = \left(0^T_{(i-1)d}, x^T, 0_{(j-i-1)d}, -x^T, 0^T_{(k-j)d}\right)^T,$$

and similarly

$$y_{i,j} = \left(0^T_{(i-1)d}, y^T, 0_{(j-i-1)d}, -y^T, 0^T_{(k-j)d}\right)^T$$

for $y \in (\mathbb{R}^d)^*$.

Theorem 3.4 *Let $x \in \mathbb{S}$. Then*

$$C_1 \cap T_x = T_x \cap \left(\frac{1}{\|x\|^2_{(2)}} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [-1,1]x_{i,j}\right)^\circ.$$

Proof By Theorem 3.2 and duality we have

$$(C_1 \cap T_x)^\circ = T_x^\circ + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \mathbb{D}^*_{i,j} = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left(\mathbb{D}^*_{i,j} + T_x^\perp\right).$$

Further,

$$T_x^\perp = \left\{\mathbf{y} = \left(y^T_1, \ldots, y^T_k\right)^T : \langle x, y_1\rangle = \cdots = \langle x, y_k\rangle = 0\right\}.$$

Now, let $y^* \in \mathbb{S}^*$ with $\langle x, y^*\rangle \ne 0$. For $y \in (\mathbb{R}^d)^*$, there exists a unique vector $z_y \in (\mathbb{R}^d)^*$ such that

$$y = \frac{\langle x, y\rangle}{\langle x, y^*\rangle} y^* + z_y, \qquad \langle x, z_y\rangle = 0.$$

Hence, using the notation introduced above, we have for $y \in \mathbb{B}^*$, $\mathbf{y} \in T_x^\perp$, and $i < j$,

$$y_{i,j} + \mathbf{y} \in \frac{\langle x, y \rangle}{\langle x, y^* \rangle} y_{i,j}^* + T_x^\perp \subset \frac{1}{\langle x, y^* \rangle} [-1, 1] y_{i,j}^* + T_x^\perp.$$

It follows that

$$\mathbb{D}_{i,j}^* + T_x^\perp = \frac{1}{\langle x, y^* \rangle} \{[-1, 1] y_{i,j}^* + \mathbf{y} : \mathbf{y} \in T_x^\perp \},$$

and thus,

$$(C_1 \cap T_x)^\circ = T_x^\perp + \frac{1}{\langle x, y^* \rangle} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [-1, 1] y_{i,j}^*.$$

Identifying again $(\mathbb{R}^d, \| \cdot \|)^*$ with $(\mathbb{R}^d, \| \cdot \|^*)$ and choosing $y^* = x/\|x\|^*$, we obtain

$$\langle x, y^* \rangle = \frac{\langle x, x \rangle}{\|x\|^*} = \frac{\|x\|_{(2)}^2}{\|x\|^*},$$

and hence, by polarity,

$$C_1 \cap T_x = T_x \cap \left(\frac{1}{\|x\|_{(2)}^2} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [-1, 1] x_{i,j} \right)^\circ,$$

as asserted. $\square$

As an example, consider the case $d = 1$ and $\mathbb{B} = [-1, 1]$. Then, by Theorem 3.4,

$$C_1^\circ = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [-1, 1](u_i - u_j),$$

where u_i denotes again the ith standard unit vector in $\mathbb{R}^k$.

Of course, in general dimensions, all sections $\mathbb{B} \cap \mathbb{R}x$ are line segments of the form $[-\alpha, \alpha]$ for some positive real α. This accounts for the additional factor in Theorem 3.4. Further, note that

$$T_x^\perp \subset \left(\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [-1, 1] x_{i,j} \right)^\circ.$$

Therefore, $C_1 \cap T_x$ is the "cross section" of the dual of a certain zonotope, which, as we will show, is a general permutahedron. While it is standard fare to derive the connection to the *standard permutahedron*

$$P_k = \operatorname{conv}\{ (\pi(1), \ldots, \pi(k))^T : \pi \in \Pi_k \}$$

directly, we will give the explicit statement and a direct independent proof in the next theorem. (In particular, Corollary 3.6 contains an alternative definition of the permutahedron as a certain zonotope.)

Theorem 3.5 *Let $x \in \mathbb{S}$. Then*

$$
C_1 \cap T_x = \bigcap_{\pi \in \Pi_k} \left\{ z = \left(v_1 x^T, \ldots, v_k x^T \right)^T : \sum_{i=1}^{k} (k + 1 - 2i) v_{\pi(i)} \leq 1 \right\}.
$$

Proof Let $\mathbf{x} = (x_1^T, \ldots, x_k^T)^T \in C_1 \cap T_x$ and $v_1, \ldots, v_k \in \mathbb{R}$ be such that $x_i = v_i x$ for $i = 1, \ldots, k$. Further, let $v_0 \leq \min_{i=1,\ldots,k} v_i$, and $y_i = (v_i - v_0)x$ for $i = 1, \ldots, k$. Then, for an arbitrary permutation $\pi \in \Pi_k$, we have

$$
\sum_{i=1}^{k} (k + 1 - 2i) v_{\pi(i)} = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} (v_{\pi(i)} - v_{\pi(j)})
$$

$$
= \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left((v_{\pi(i)} - v_0) - (v_{\pi(j)} - v_0) \right)
$$

$$
= \sum_{i=1}^{k-1} \sum_{i=j+1}^{k} \left(\|y_{\pi(i)}\| - \|y_{\pi(j)}\| \right) \leq \sum_{i=1}^{k-1} \sum_{i=j+1}^{k} \|y_{\pi(i)} - y_{\pi(j)}\|
$$

$$
= \sum_{i=1}^{k-1} \sum_{i=j+1}^{k} \|x_{\pi(i)} - x_{\pi(j)}\| \leq 1,
$$

showing the first inclusion "$\subset$".

To prove the reverse inclusion, let $v_1, \ldots, v_k \in \mathbb{R}$ be such that

$$
\sum_{i=1}^{k} (k + 1 - 2i) v_{\pi(i)} \leq 1
$$

for all $\pi \in \Pi_k$. Further, set $x_i = v_i x$ for $i = 1, \ldots, k$ and $z = (x_1^T, \ldots, x_k^T)^T$. Then, of course, $z \in T_x$. Now, let $\pi^* \in \Pi_k$ be such that

$$
v_{\pi^*(1)} \geq \cdots \geq v_{\pi^*(k)}.
$$

Then, in particular,

$$\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\|x_i - x_j\| = \sum_{i=1}^{k-1}\sum_{j=i+1}^{k}|v_i - v_j|\|x\| = \sum_{i=1}^{k-1}\sum_{j=i+1}^{k}|v_{\pi^*(i)} - v_{\pi^*(j)}|$$

$$= \sum_{i=1}^{k-1}\sum_{j=i+1}^{k}(v_{\pi^*(i)} - v_{\pi^*(j)}) = \sum_{i=1}^{k}(k+1-2i)v_{\pi^*(i)} \leq 1,$$

and hence, $z \in C_1$. This completes the proof.   □

Theorem 3.5 gives an explicit description by inequalities of $C_1 \cap T_x$ as the dual of a *general permutahedron*, which is a homothetic copy of the standard permutahedron P_k. Since P_k is a $(k-1)$-dimensional polytope in $\mathbb{R}^k$ which lies in the hyperplane of all points whose coordinate sum equals $k(k+1)/2$ and is symmetric with respect to

$$c = \frac{k+1}{2}(1, \ldots, 1)^T,$$

the combination of Theorems 3.4 and 3.5 gives, in particular, the following well-known description.

Corollary 3.6

$$\left(2(P_k - c)\right)^\circ = \left(\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}[-1,1](u_i - u_j)\right)^\circ$$

$$= \bigcap_{\pi \in \Pi_k}\left\{v = (v_1, \ldots, v_k) \in \mathbb{R}^k : \sum_{i=1}^{k}(k+1-2i)v_{\pi(i)} \leq 1\right\}.$$

Proof We have

$$\left(2(P_k - c)\right)^\circ = \left\{v \in \mathbb{R}^k : w \in 2P_k \Rightarrow v^T(w - 2c) \leq 1\right\}$$

$$= \bigcap_{\pi \in \Pi_k}\left\{v = (v_1, \ldots, v_k) \in \mathbb{R}^k : \sum_{i=1}^{k}2v_i\pi(i) - (k+1)\sum_{i=1}^{k}v_i \leq 1\right\}$$

$$= \bigcap_{\sigma \in \Pi_k}\left\{v = (v_1, \ldots, v_k) \in \mathbb{R}^k : \sum_{i=1}^{k}(2i - (k+1))v_{\sigma(i)} \leq 1\right\},$$

which proves the assertion.   □

Of course, the following result is another immediate corollary.

Corollary 3.7 *Let $x \in \mathbb{S}$, and let $h = h_x : \mathbb{R}^k \to \mathbb{R}^{kd}$ denote the linear map defined by*

$$h(v) = \begin{pmatrix} v_1 x \\ \vdots \\ v_k x \end{pmatrix} \quad \textit{for } v = (v_1, \ldots, v_k)^T \in \mathbb{R}^k.$$

Then

$$C_1 \cap T_x = h\big((2(P_k - c))^\circ\big).$$

Proof The assertion follows directly from Theorem 3.5 and Corollary 3.6.   □

4 On the Structure of A

In Sect. 3 we have used the special form of A to determine C° for some particular choices of $\| \cdot \|_\diamond$. We begin the present section by studying the matrix A in more detail and then further specialize the involved norms. The following lemma specifies A's singular value decomposition.

Lemma 4.1 *The spectrum of $A^T A$ consists of 0 and k; k is a $d(k-1)$-fold eigenvalue. Let the components of $D = (\delta_{i,j}) \in \mathbb{R}^{dr \times dk}$ be given by*

$$\delta_{i,j} = \begin{cases} \sqrt{k} & \textit{for } i = j = 1, \ldots, d(k-1), \\ 0 & \textit{else}; \end{cases}$$

let $Q = (q_1, \ldots, q_{dk}) \in \mathbb{R}^{dk \times dk}$ such that $q_1, \ldots, q_{d(k-1)}$ and $q_{d(k-1)+1}, \ldots, q_{dk}$ form an orthonormal bases of L and $\mathrm{ls}(C)$, respectively. Then there is an orthogonal matrix $\overline{Q} = (\overline{q}_1, \ldots, \overline{q}_{dr}) \in \mathbb{R}^{dr \times dr}$ such that

$$\overline{q}_i = \frac{1}{\sqrt{k}} A q_i \quad \textit{for } i = 1, \ldots, d(k-1)$$

and

$$A = \overline{Q} D Q.$$

Proof With $I = I_d$ and $0 = 0_d$, we have

$$A^T A = k \cdot \begin{pmatrix} I & 0 & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & I \end{pmatrix} - \begin{pmatrix} I & I & \cdots & I \\ I & I & \cdots & I \\ \vdots & \vdots & \ddots & \vdots \\ I & I & \cdots & I \end{pmatrix}$$

$$= k I_{dk} - (I, \ldots, I)^T (I, \ldots, I).$$

Now, note that

$$A^T A \mathbf{x} = \begin{cases} k\mathbf{x} & \text{for } \mathbf{x} \in L, \\ 0 & \text{for } \mathbf{x} \in \mathrm{ls}(C). \end{cases}$$

Hence, k and 0 are a $d(k-1)$-fold and a d-fold eigenvalue, respectively, with the corresponding eigenspaces L and $\mathrm{ls}(C)$.

There exist orthogonal matrices

$$Q \in \mathbb{R}^{dk \times dk}, \qquad \overline{Q} \in \mathbb{R}^{dr \times dr}$$

such that A has the singular value decomposition

$$A = \overline{Q} D Q.$$

The columns of $Q = (q_1, \ldots, q_{dk})$ and $\overline{Q} = (\overline{q}_1, \ldots, \overline{q}_{dr})$ constitute an orthonormal basis of eigenvectors of $A^T A$ and $A A^T$, respectively, that satisfy the relation

$$\overline{q}_i = \frac{1}{\sqrt{k}} A q_i \quad \text{for } i = 1, \ldots, d(k-1),$$

and all such corresponding choices are allowed. In particular, we can choose as columns of Q an arbitrary orthonormal basis of L expanded by an arbitrary orthonormal basis of $\mathrm{ls}(C)$. □

Note that we can choose Q and $\overline{Q}$ so that their determinants are $+1$. In conjunction with Theorem 2.5, Lemma 4.1 shows then that

$$C^\circ = A^T K^\circ = Q^T D^T \overline{Q}^T K^\circ$$

is geometrically obtained from K° by a sequence of operations beginning with a rotation, followed by a projection on the space of the first $d(k-1)$ coordinates of $\mathbb{R}^{dk}$, an expansion by a factor $\sqrt{k}$, and finally another rotation.

Of course, there are various lines of studies starting from here. It is natural to consider bodies K° which are invariant under $\overline{Q}^T$ or, even more appropriately, whose intersection with a suitable $\mathbb{R}^{d(k-1)}$ coincides with $\frac{1}{\sqrt{k}} D^T \overline{Q}^T K^\circ$. Also, one can start from a norm in $\mathbb{R}^{d(k-1)}$ and design "preimage norms" in $\mathbb{R}^{dr}$. We, however, will not do that here since other aspects are more important in our context. Rather, we will close this section with a few remarks on the situation where $\| \cdot \|_\diamond$ and $\| \cdot \|$ are specialized to certain ℓ_p norms.

If $p_1, p_2 \in [1, \infty]$ and

$$\| \cdot \|_\diamond = \| \cdot \|_{(p_1)}, \qquad \| \cdot \| = \| \cdot \|_{(p_2)},$$

we will signify C, B, and K by the subscripts p_1, p_2, i.e., we write C_{p_1,p_2}, B_{p_1,p_2}, and K_{p_1,p_2}. In particular, for $p_1, p_2 \in [1, \infty[$ and with $\mathbf{x} = (x_1^T, \ldots, x_k^T)^T$ and

$x_i = (\xi_{i,1}, \ldots, \xi_{i,d})^T$ for $i = 1, \ldots, k$,

$$C_{p_1,p_2} = \left\{ \mathbf{x} \in \mathbb{R}^{kd} : \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \left(\sum_{l=1}^{d} |\xi_{i,l} - \xi_{j,l}|^{p_2} \right)^{\frac{p_1}{p_2}} \leq 1 \right\}.$$

Clearly, the *fully ℓ_p case*, i.e., $p = p_1 = p_2$, is most natural; in particular,

$$K_{p,p} = \left\{ \mathbf{v} \in \mathbb{R}^{dr} : \|\mathbf{v}\|_{(p)} \leq 1 \right\} = \mathbb{B}_{(p)}^{dr}.$$

Among the clustering bodies $C_{p,p}$, the fully ℓ_∞ body $C_{\infty,\infty}$ has the simplest structure; in fact,

$$C_{\infty,\infty} = \left\{ \mathbf{x} = (x_1^T, \ldots, x_k^T)^T \in \mathbb{R}^{kd} : \max_{\substack{i=1,\ldots,k-1, j=i+1,\ldots,k \\ l=1,\ldots,d}} |\xi_{i,l} - \xi_{j,l}| \leq 1 \right\},$$

i.e., $C_{\infty,\infty}$ is the polyhedron given by the $dk(k-1)$ linear inequalities

$$-1 \leq \xi_{i,l} - \xi_{j,l} \leq 1$$

for $i = 1, \ldots, k-1, j = i+1, \ldots, k, l = 1, \ldots, d$. Of course, then $B_{\infty,\infty}$ is the polytope obtained from $C_{\infty,\infty}$ by adding the d additional homogeneous linear equations

$$\sum_{i=1}^{k} \xi_{i,l} = 0 \quad (l = 1, \ldots, d).$$

Let us now turn to the fully ℓ_1 case, i.e, $\| \cdot \|_\diamond$ and $\| \cdot \|$ are ℓ_1 norms. $\mathbb{B}_{(1)}$ is, of course, the well-known regular cross polytope. As we have seen, for $d = 1$, the corresponding clustering body is the dual of a permutahedron. This is the special case of the following result which identifies $C_{1,1}^\circ$ as a Cartesian product of permutahedra (up to the reordering of coordinates facilitated by the permutation matrix P introduced in Sect. 2).

Theorem 4.2

$$C_{1,1}^\circ = (k+1)(-1, \ldots, -1)^T + 2P(P_k \times \cdots \times P_k).$$

Proof Since $\mathbb{B}_{(1)}^* = \mathbb{B}_{(\infty)} = \sum_{l=1}^{d}[-1, 1]u_l$, we have

$$\hat{\mathbb{D}}^* = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} : \|x\|_{(\infty)} \leq 1 \right\} = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} : x \in \sum_{l=1}^{d}[-1, 1]u_l \right\}$$

$$= \sum_{l=1}^{d}[-1, 1](u_l - u_{d+l}).$$

(Recall that by our convention u_l denotes the lth standard unit vector independently of the space we are in. In the first two lines, $u_l \in \mathbb{R}^d$, while in the third line, $u_l \in \mathbb{R}^{2d}$.)

Hence, by Theorem 3.2 and Corollary 3.6,

$$C_{1,1}^{\circ} = A^T (\mathbb{B}_{(1)}^*)^r = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \mathbb{D}_{i,j}^* = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \sum_{l=1}^{d} [-1,1](u_{d(i-1)+l} - u_{d(j-1)+l})$$

$$= \sum_{l=1}^{d} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [-1,1](u_{d(i-1)+l} - u_{d(j-1)+l})$$

$$= P\big(2(P_k - c) \times \cdots \times 2(P_k - c)\big)$$

$$= (k+1)(-1, \ldots, -1)^T + 2P(P_k \times \cdots \times P_k). \qquad \square$$

The fully Euclidean case where $\| \cdot \|_\diamond$ and $\| \cdot \|$ are both ℓ_2 norms is a simple corollary of Lemma 4.1.

Theorem 4.3 $B_{2,2}$ *is congruent to* $\sqrt{k}\mathbb{B}^{d(k-1)}$.

Proof By Lemma 4.1, the corresponding body $C_{2,2}$ is the cylinder with "axis" ls(C) and "cross section" similar to the Euclidean ball of dimension $d(k-1)$ with the dilatation factor $\sqrt{k}$. $\qquad \square$

5 Approximation

In the following, we prove results on the approximation of C and C° by polyhedra. As is well known, polyhedra Q can be given as the feasible region of a finite number of linear inequalities or as the Minkowski sum of the convex hull of a finite point set and the conic hull of another finite point set. We will speak of an $\mathcal{H}$-presentation of Q in the first, and a $\mathcal{V}$-presentation in the second case; see [9]. For simplicity, we will also refer to Q as an $\mathcal{H}$-polyhedron or a $\mathcal{V}$-polyhedron, respectively. Since the number of facets of a polytope can be exponential in the number of its vertices and vice versa, the given presentation is highly relevant for algorithmic purposes. Of course, with a view towards polynomial-time approximation algorithms for the underlying convex maximization problems, approximations of C by $\mathcal{H}$-polyhedra Q with polynomially many facets or of C° by $\mathcal{V}$-polytopes with polynomially many vertices are our main focus. Once such an approximation is known, say by a (centrally symmetric) $\mathcal{H}$-polyhedron

$$Q = \big\{z : z^T q_1 \leq 1 \wedge \cdots \wedge z^T q_n \leq 1\big\},$$

the original convex maximization problem can be (approximately) solved by means of the n linear programs with objective functions $z \mapsto q_i^T z$ $(i = 1, \ldots, n)$.

We begin with a simple consequence of Theorem 2.5 that sharpens Corollary 2.6.

Lemma 5.1 *Let K be a polytope with m facets. Then the number of vertices of C° and the numbers of facets of C and B (as a polytope in L) are bounded above by m.*

Proof The results follow from Theorem 2.5 with the aid of polarity.   □

Of course, in general, the upper bound given in Lemma 5.1 is not tight. This is due to the projection part in the equation $C^\circ = A^T K^\circ = Q^T D^T \overline{Q}^T K^\circ$. For instance, in the fully ℓ_1 case, Lemma 5.1 gives the upper bound 2^{dr} for the number of facets of $C_{1,1}$, while by Theorem 4.2 we know that the correct number is $(k!)^d$, which is asymptotically much smaller. However, $(k!)^d$ is still exponential in d and in k, so even for rather small values of these parameters, it will not be possible to solve the underlying convex maximization problem exactly in practice since this would require the solution of $(k!)^d$ linear programs. Hence, even in such a polytopal case let alone for more general norms, we will have to resort to *polynomial approximations*, i.e., approximations of C or B by $\mathcal{H}$-polyhedra with only polynomially many facets. Equivalently, we may approximate C° by $\mathcal{V}$-polyhedra with only polynomially many vertices.

The following simple observation reduces this task to that of approximating K°.

Lemma 5.2 *Let* $\lambda, \mu > 0$, *let* S *be a* $\mathcal{V}$-*polytope such that* $\lambda S \subset K^\circ \subset \mu S$, *and set* $Q = A^T S$. *Then* Q *is a* $\mathcal{V}$-*polytope whose number of vertices is bounded by that of* S, *and we have*

$$\lambda Q \subset C^\circ \subset \mu Q.$$

Proof Applying Theorem 2.5, we obtain

$$\lambda Q = \lambda A^T S \subset A^T K^\circ = C^\circ \subset \mu A^T S = \mu Q. \qquad \square$$

In the fully ℓ_p case, we have $K = \mathbb{B}^{dr}_{(p)}$; hence, Lemma 5.2 can be combined with known results on the approximation of ℓ_p balls by polytopes. In fact, [6, 7] give deterministic and randomized algorithms which in many cases are even asymptotically optimal for that task. With q again defined by

$$\frac{1}{p} + \frac{1}{q} = 1,$$

we then obtain the error bounds

$$O\left(\frac{1}{\sqrt{dr}}\right) \quad (p = 1), \qquad O\left(\frac{(\log(dr))^{1/q}}{\sqrt{dr}}\right) \quad (1 < p < 2),$$

$$O\left(\left(\frac{\log(dr)}{dr}\right)^{\frac{1}{p}}\right) \quad (p \geq 2)$$

for our polynomial approximations of the specific balls. (Bounds for the convex maximization problem that involve pth powers of norms have, of course, to be adjusted.) Let us point out that, as for polynomial-time approximations of $\mathbb{B}^{dr}_{(p)}$ (and in the appropriate model of computation), the given bound is tight for $p \geq 2$; see [7] for more details and further results on the polynomial-time approximation of various other geometric functionals.

Note that approximation in $\mathbb{R}^{dr}$ ignores part of the potential that lies in the specific structure of clustering bodies. For instance, since $r = k(k-1)/2$, in the fully ℓ_2 case, we only obtain an error bound of $O((\log(dk)/(dk^2))^{1/2})$ along these lines. Theorem 4.3, on the other hand, leads to the bound

$$O\left(\left(\frac{\log(dk)}{dk}\right)^{\frac{1}{2}}\right)$$

and hence reduces the impact of k from linear to just square root.

As we have already remarked after Lemma 5.1, a similar effect occurs in the fully ℓ_1 case. Hence, in the following, we give improved approximations by utilizing the specific structural properties of clustering bodies. We concentrate on the case of B_p for $p \in [1, \infty]$ and begin with $p = \infty$.

Theorem 5.3 *Let $\lambda, \mu > 0$, and let S be a polytope with m vertices such that $\lambda S \subset \mathbb{B}^* \subset \mu S$. Then there is a polytope Q with mr vertices such that*

$$\lambda Q \subset C_\infty^\circ \subset \mu Q.$$

Proof By Theorem 3.2 we have

$$C_\infty^\circ = \mathrm{conv}\left(\bigcup_{i=1}^{k-1} \bigcup_{j=i+1}^{k} \mathbb{D}_{i,j}^*\right),$$

where

$$\mathbb{D}_{i,j}^* = \left\{\mathbf{y} = \left(y_1^T, \ldots, y_k^T\right)^T \in \mathbb{R}^{dk} : y_i, y_j \in \mathbb{B}^* \wedge (y_i + y_j = 0) \wedge y_r = 0 \ (r \neq i, j)\right\}.$$

So, let $s_1, \ldots, s_m$ denote the vertices of S and set

$$T_{i,j} = \mathrm{conv}\{\mathbf{y} = \left(y_1^T, \ldots, y_k^T\right)^T \in \mathbb{R}^{dk} :$$
$$y_i \in \{s_1, \ldots, s_m\} \wedge (y_i + y_j = 0) \wedge y_r = 0 \ (r \neq i, j)\}$$

and

$$Q = \mathrm{conv}\left(\bigcup_{i=1}^{k-1} \bigcup_{j=i+1}^{k} T_{i,j}\right).$$

Then, of course, $\lambda T_{i,j} \subset \mathbb{D}_{i,j}^* \subset \mu T_{i,j}$, and hence,

$$\lambda Q \subset C_\infty^\circ \subset \mu Q.$$

Clearly, Q is a polytope with mr vertices. $\qquad\square$

In the special case of $\mathbb{B} = \mathbb{B}_p^d$, we then obtain from [6] that $C_{\infty,p}$ can be approximated by a polytope with polynomially many facets up to an error

$$O\left(\frac{1}{\sqrt{d}}\right) \quad (p = 1), \qquad O\left(\frac{(\log(d))^{1/q}}{\sqrt{d}}\right) \quad (1 < p < 2),$$

$$O\left(\left(\frac{\log(d)}{d}\right)^{\frac{1}{p}}\right) \quad (p \geq 2).$$

The next theorem gives an inequality that will lead to a close approximation of C_p once a close approximation of $\mathbb{B}$ is given.

We will make free use of a standard inequality between different ℓ_p norms of vectors $z \in \mathbb{R}^n$. More precisely,

$$\|z\|_{(q)} \leq \|z\|_{(p)} \leq n^{\frac{1}{p}-\frac{1}{q}} \|z\|_{(q)}$$

for $1 \leq p \leq q \leq \infty$ (with the understanding again that $1/\infty = 0$).

Theorem 5.4 *Let* $\mathbf{x} = (x_1^T, \ldots, x_k^T)^T \in L$ *and* $p \in [1, \infty[$. *Then*

$$\frac{k}{2}\left(\frac{k}{k-1}\right)^{p-1} \sum_{i=1}^{k} \|x_i\|^p \leq \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \|x_i - x_j\|^p \leq 2^{p-1}(k-1) \sum_{i=1}^{k} \|x_i\|^p.$$

Proof We have

$$k^p \sum_{i=1}^{k} \|x_i\|^p = \sum_{i=1}^{k} \|kx_i\|^p = \sum_{i=1}^{k} \|(k-1)x_i + x_i\|^p$$

$$= \sum_{i=1}^{k} \left\| (k-1)x_i - \sum_{\substack{j=1 \\ j\neq i}}^{k} x_j \right\|^p = \sum_{i=1}^{k} \left\| \sum_{\substack{j=1 \\ j\neq i}}^{k} (x_i - x_j) \right\|^p$$

$$\leq \sum_{i=1}^{k} \left(\sum_{\substack{j=1 \\ j\neq i}}^{k} \|x_i - x_j\| \right)^p.$$

Now, with $\tau_{i,j} = \|x_i - x_j\|$ and $t_i = (\tau_{i,1}, \ldots, \tau_{i,i-1}, \tau_{i,i+1}, \ldots, \tau_{i,k})^T \in \mathbb{R}^{k-1}$, we obtain from the standard inequalities between ℓ_1 and ℓ_p norms that, for all $i = 1, \ldots, k$,

$$\sum_{\substack{j=1 \\ j\neq i}}^{k} \|x_i - x_j\| = \|t_i\|_{(1)} \leq (k-1)^{1-\frac{1}{p}} \|t_i\|_{(p)}.$$

Also, for $s_{i,j} = (\|x_i\|, \|x_j\|)^T$,

$$\|s_{i,j}\|_{(1)} \leq 2^{1-\frac{1}{p}} \|s_{i,j}\|_{(p)}.$$

With the aid of these two inequalities, we can continue with

$$k^p \sum_{i=1}^{k} \|x_i\|^p \le (k-1)^{p-1} \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \ne i}}^{k} \|x_i - x_j\|^p \le (k-1)^{p-1} \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \ne i}}^{k} \left(\|x_i\| + \|x_j\|\right)^p$$

$$\le \left(2(k-1)\right)^{p-1} \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \ne i}}^{k} \left(\|x_i\|^p + \|x_j\|^p\right) = \left(2(k-1)\right)^p \sum_{i=1}^{k} \|x_i\|^p.$$

Hence, we have

$$k\left(\frac{k}{k-1}\right)^{p-1} \sum_{i=1}^{k} \|x_i\|^p \le \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \ne i}}^{k} \|x_i - x_j\|^p \le 2^p(k-1) \sum_{i=1}^{k} \|x_i\|^p,$$

which implies the assertion.   □

We can use Theorem 5.4 now to relate B_p to a convex body in the same space $\mathbb{R}^{dk}$ rather than to K_p in $\mathbb{R}^{dr}$. In fact, let

$$N_p = \left\{ \mathbf{x} = \left(x_1^T, \dots, x_k^T\right)^T : \left\| \left(\|x_1\|, \dots, \|x_k\|\right)^T \right\|_{(p)} \le 1 \right\}.$$

Note that N_p has the same structural properties as K_p; see Sect. 2.

In terms of approximations we obtain the following result.

Theorem 5.5 *Let* $\lambda, \mu > 0$, *let* S *be a polytope with* m *facets such that* $\lambda S \subset N_p \subset \mu S$, *and set*

$$Q = \frac{1}{k}\left(\frac{k-1}{2}\right)^{1/q} (S \cap L).$$

Then (as a polytope in L) Q *has at most* m *facets, and*

$$\lambda\left(\frac{k}{k-1}\right) Q \subset B_p \subset \mu 2Q.$$

Proof The nontrivial part of the assertion follows directly from Theorem 5.4.   □

Since the case $p = 1$ in Theorem 5.4 is of special relevance, we state the following corollary explicitly.

Corollary 5.6

$$\frac{1}{k-1}\left(\mathbb{B}_{(1)}^{dk} \cap L\right) \subset B_{1,1} \subset \frac{2}{k}\left(\mathbb{B}_{(1)}^{dk} \cap L\right).$$

Proof The assertion follows by applying Theorem 5.4 to $N_{1,1}$.   □

As a special case, we see that the dual of the standard permutahedron P_k can be approximated by a polytope with at most 2^k facets up to an error of less than 2.

Of course, now we can again use the results of [7] to approximate $B_{p,p}$ by a polytope with polynomially many facets up to an error

$$O\left(\frac{1}{\sqrt{dk}}\right) \quad (p=1), \qquad O\left(\frac{(\log(dk))^{1/q}}{\sqrt{dk}}\right) \quad (1<p<2),$$

$$O\left(\left(\frac{\log(dk)}{dk}\right)^{\frac{1}{p}}\right) \quad (p\geq 2).$$

As an example for the reductions we obtain, let us point out that for $k=13$, P_k° has $6,227,020,800$ facets. Its ℓ_1-ball approximation has an error less than 2 and only $8,192$ facets, while the final approximation error of at most $2\sqrt{k}\approx 7.21$ needs only 28 facets. Consequently, we obtain a provably good approximation for the optimum of the corresponding convex optimization problem over an arbitrary polytope by solving only 28 linear programs. (Naturally, for all practical purposes of the consolidation of farmland, the approximation error is much smaller than the given worst-case bound.)

References

1. Benyamini, Y., Lindenstrauss, J.: Geometric nonlinear functional analysis I. Am. Math. Soc. Coll. Publ., vol. 48 (2000)
2. Boedlaender, H.L., Gritzmann, P., Klee, V., van Leeuwen, J.: Computational complexity of norm-maximization. Combinatorica **10**, 203–225 (1990)
3. Borgwardt, S., Brieden, A., Gritzmann, P.: : Constrained minimum-k-star clustering and its application to the consolidation of farmland. Oper. Res. (2009, in print). http://dx.doi.org/10.1007/s12351-009-0041-y
4. Brieden, A., Gritzmann, P.: A quadratic optimization model for the consolidation of farmland by means of lend-lease agreements. In: Ahr, D., Fahrion, R., Oswald, M., Reinelt, G. (eds.) Operations Research Proceedings 2003: Selected Papers of the International Conference on Operations Research (OR 2003), pp. 324–331. Springer, Heidelberg (2004)
5. Brieden, A., Gritzmann, P.: Voronoi properties of optimal clusterings (2009, submitted)
6. Brieden, A., Gritzmann, P., Kannan, R., Klee, V., Lovász, L., Simonovits, M.: Approximation of diameters: Randomization doesn't help. In: IEEE Symp. Found. Computer Sci. (FOCS'98), pp. 244–251 (1998)
7. Brieden, A., Gritzmann, P., Kannan, R., Klee, V., Lovász, L., Simonovits, M.: Deterministic and randomized polynomial-time approximation of radii. Mathematika **48**, 63–105 (2001)
8. Gritzmann, P., Klee, V.: Computational complexity of inner and outer j-radii of polytopes in finite-dimensional normed spaces. Math. Prog. **59**, 163–213 (1993)
9. Gritzmann, P., Klee, V.: On the complexity of some basic problems in computational convexity: I. Containment problems. Discrete Math. **136**, 129–174 (1994)
10. Gruber, P.M.: Aspects of approximation of convex bodies. In: Gruber, P.M., Wills, J.M. (eds.) Handbook of Convex Geometry, pp. 319–345. North-Holland, Amsterdam (1993)
11. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)