McMaster University

Advanced Optimization Laboratory

Title:

Kissing Polytopes

Authors:

Antoine Deza, Shmuel Onn,
Sebastian Pokutta, and Lionel Pournin

AdvOL-Report No. 2023/1

May 2023, Hamilton, Ontario, Canada

KISSING POLYTOPES

Antoine Deza

McMaster University, Hamilton, Ontario, Canada
e-mail: deza@mcmaster.ca

Shmuel Onn

Technion – Israel Institute of Technology, Haifa, Israel
e-mail: onn@technion.ac.il

Sebastian Pokutta

Zuse Institute Berlin, Germany
e-mail: pokutta@zib.de

Lionel Pournin

Université Paris 13, Villetaneuse, France
e-mail: lionel.pournin@univ-paris13.fr

abstract

We investigate the following question: *how close can two disjoint lattice polytopes contained in a fixed hypercube be?* This question stems from various contexts where the minimal distance between such polytopes appears in complexity bounds of optimization algorithms. We provide nearly matching lower and upper bounds on this distance and discuss its exact computation. We also give similar bounds in the case of disjoint rational polytopes whose binary encoding length is prescribed.

1. Introduction

In general, the distance between two disjoint convex bodies P and Q contained in $\mathbb{R}^d$ can get arbitrarily small. However, this is no longer the case when P and Q satisfy certain constraints. For instance, if P and Q are two d-dimensional 0/1-polytopes, then they cannot be closer than a positive distance that only depends on d. This is due to the observation that, when d is fixed, there are finitely many such pairs of polytopes. Another relevant constraint that often arises in optimization algorithms is when P and Q are rational polytopes whose binary encoding length (as subsets of $\mathbb{R}^d$ satisfying a set of linear inequalities) is prescribed. Here, again, the smallest possible distance between P and Q is a positive number that depends on that encoding length and on d. Our goal is to estimate these minimal distances.

Our study stems from the complexity bounds established by Gábor Braun, Sebastian Pokutta, and Robert Weismantel [4]. In their article, an algorithm is provided that either computes a point in $P \cap Q$ when that intersection is non-empty or certifies that $P \cap Q$ is empty. In the latter case, the complexity of certifying that $P \cap Q$ is empty is

$$O\left(\frac{1}{d(P,Q)^2}\right)$$

and therefore, it is natural to ask how small $d(P,Q)$ can get.

It turns out that our study of the minimum distance between disjoint polytopes is related to the notion of *facial distance* considered by Javier Peña and Daniel Rodriguez [8, Section 2] and David Gutman and Javier Peña [5, 7]. It is also related to the notion of *vertex-facet distance* investigated by Amir Beck and Shimrit Shtern [2]. Another quantity linked to these three is the *pyramidal width* of a polytope, studied by Simon Lacoste-Julien and Martin Jaggi [6]. We refer the reader to the survey by Gábor Braun, Alejandro Carderera, Cyrille Combettes, Hamed Hassani, Amin Karbasi, Aryan Mokhtari, and Sebastian Pokutta [3] for an overview of these notions.

The facial distance is crucial in establishing linear convergence rates for conditional gradient methods over polytopes and naturally occurs in the complexity bounds. The facial distance of a polytope P is defined as

$$\Phi(P) = \min\Big\{d\big(F, \mathrm{conv}(\mathcal{V} \backslash F)\big) : F \in \mathcal{F}\Big\}, \tag{1}$$

where $\mathcal{V}$ denotes the vertex set of P and $\mathcal{F}$ the set of its proper faces. In other words, the facial distance of P is the minimal distance between any of its faces and the convex hull of its vertices not contained in that face. In contrast to our study, this notion considers a specific polytope P and decomposes it into its faces and their complements. The vertex-facet distance is measured in the special case when F is a facet of the considered polytope and is replaced in (1) with its affine hull. It can then be expressed as

$$\Delta(P) = \min\Big\{d\big(\operatorname{aff}(F), \operatorname{conv}(\mathcal{V}\backslash F)\big) : F \in \overline{\mathcal{F}}\Big\}, \tag{2}$$

where $\overline{\mathcal{F}}$ is the set of the facets of P, as shown in [8, Section 2]. When P is a 0/1-simplex, bounds have been given on the smallest possible vertex-facet distance [1, 10]. In particular Noga Alon and Văn Vũ show in [1] that

$$\frac{1}{\sqrt{2}^{d\log d-2d+o(d)}} \leqslant \min\ \Delta(S) \leqslant \frac{1}{\sqrt{2}^{d\log d-4d+o(d)}} \tag{3}$$

where the minimum is over all the d-dimensional 0/1-simplices S.

The results of Stephen Vavasis on the complexity of quadratic optimization [12], generalized by Alberto Del Pia, Santanu Dey, and Marco Molinaro in [9] imply as a special case that the squared distance between two rational polytopes is a rational number. Our work is concerned by providing bounds on how close such polytopes can be under the mentioned constraints.

Recall that a polytope whose vertices belong to the integer lattice $\mathbb{Z}^d$ is a *lattice polytope*. We will refer to a lattice polytope contained in the hypercube $[0,k]^d$ as a *lattice (d,k)-polytope*. In this article, we first provide a lower bound, as a function of d and k on the smallest possible distance between two disjoint lattice (d,k)-polytopes and then we complement these lower bounds with constructions that provide almost matching upper bounds.

In terms of lower bounds our main result is the following.

THEOREM 1.1: *If P and Q are disjoint lattice (d,k)-polytopes, then*

$$d(P,Q) \geqslant \frac{1}{(kd)^{2d}}.$$

We shall in fact prove a stronger bound (see Theorem 2.3) which Theorem 1.1 is a consequence of. We also prove a lower bound on the distance of two rational polytopes in terms of the dimension and their binary encoding length (see Theorem 3.6). Our main result regarding upper bounds in the following.

THEOREM 1.2: *Consider a positive integer k. For any large enough d, there exist two disjoint (d,k)-lattice polytopes P and Q such that*

$$d(P,Q) \leqslant \frac{1}{\big(k\sqrt{d}\big)^{\sqrt{d}}}.$$

As above, Theorem 1.2 follows from a stronger bound (see Theorem 4.1). We also give an upper bound on the smallest possible distance between two rational polytopes whose binary encoding length is prescribed (see Theorem 4.6).

By its definition, the facial distance of a polytope is a distance between two polytopes. Inversely, the distance between two polytopes P and Q is the distance between two of their faces that belong to parallel hyperplanes. In particular, $d(P,Q)$ is at least the facial distance of the convex hull of these two faces. As a consequence, our results provide bounds on the smallest possible facial distance of a lattice (d,k)-polytope in terms of d and k.

THEOREM 1.3: *For any positive k and large enough d,*

$$\frac{1}{(kd)^{2d}} \leqslant \min \Phi(P) \leqslant \frac{1}{\left(k\sqrt{d}\right)^{\sqrt{d}}} \tag{4}$$

where the minimum is over all the lattice (d,k)-polytopes P.

Similar bounds in the case of rational polytopes, in terms of their dimension and binary encoding length follow from Theorems 3.6 and 4.6.

Remark 1.4: While (3) and (4) have similar forms, neither of these statements implies the other. Indeed, on the one hand $\Delta(P)$ is the distance between a point and a hyperplane and on the other, $\Phi(P)$ is the distance between two polytopes of arbitrary dimensions. More precisely, the lower bound in (4) cannot be deduced from that in (3) and the same holds for the upper bounds.

We establish the announced lower bounds for lattice polytopes in Section 2 and for rational polytopes in Section 3. The upper bounds and the corresponding constructions, for both lattice and rational polytopes, are provided in Section 4. These upper bounds are only valid for all sufficiently large dimensions and we therefore provide bounds that hold in all dimensions in Section 5, where we also study the smallest possible distance of two lattice polytopes whose dimension is fixed independently on the dimension of the ambient space. We end the article with Section 6, that contains computational results and, in particular the exact value of the smallest possible distance between two disjoint lattice (d,k)-polytopes for certain d and k (see Table 1). In order to compute these distances, we prove in Section 6 that one can restrict to considering a well-behaved subset of the pairs of lattice (d,k)-polytopes.

2. Lower bounds

In this section P and Q are two fixed, disjoint polytopes contained in $\mathbb{R}^d$ and our goal is to prove Theorem 1.1. Let us first introduce some notations and give a few remarks. Since P and Q are compact subsets of $\mathbb{R}^d$, there exists a point p in P and a point q in Q whose distance is equal to $d(P,Q)$. Let f_P denote the unique face of P that contains p in its relative interior and f_Q the unique face of Q that contains q in its relative interior. This situation is illustrated in Figure 1, where P and Q are two 0/1-polytopes, f_P is the diagonal of the cube and f_Q a diagonal of one of its square faces.

We now consider $\dim(f_P)+1$ vertices of f_P, that we label by u^0 to $u^{\dim(f_P)}$ such that the vectors u^1-u^0 to $u^{\dim(f_P)}-u^0$ are linearly independent. Similarly, pick a family v^0 to $v^{\dim(f_Q)}$ of vertices of f_Q such that the vectors v^1-v^0 to $v^{\dim(f_Q)}-v^0$ are linearly independent. Consider the set

$$S=\Big\{u^i-u^0:1\leqslant i\leqslant \dim(f_P)\Big\}\cup\Big\{v^i-v^0:1\leqslant i\leqslant \dim(f_Q)\Big\}$$

and extract from it a subset of linearly independent vectors w^1 to w^r that span the same subspace of $\mathbb{R}^d$ than S. Further denote

$$w^0=u^0-v^0.$$

The equality

$$d(p,q)=\frac{(p-q)\cdot(p-q)}{\|p-q\|}$$

can be rewritten into

(5) $$d(p,q)=w^0\cdot\frac{(p-q)}{\|p-q\|}.$$

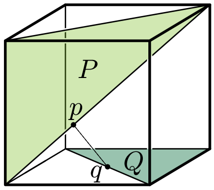

Figure 1. Two 0/1-polytopes P and Q and points p and q such that $d(P,Q)$ is equal to $d(p,q)$.

We will express the quotient in the right-hand side of (5) in terms of the vectors w^i. In order to do that, consider the matrix M whose term in row i and column j is $w^i \cdot w^j$, the column vector b whose coordinates are $w^0 \cdot w^1$ to $w^0 \cdot w^r$, and the column vector β whose coordinates are β_1 to β_r. Further denote by M_i the matrix obtained from M by replacing column i with b.

LEMMA 2.1: *The distance between P and Q satisfies*

$$d(P,Q) = w^0 \cdot \frac{a}{\|a\|}$$

where

$$a = \det(M) w^0 - \sum_{i=1}^{r} \det(M_i) w^i.$$

Proof. Observe that $p - q$ belongs to the space spanned by vectors w^0 to w^r. Hence, $\lambda(p-q)$ also belongs to that space for any positive λ. As a consequence, there exist $r+1$ coefficients α_0 to α_r such that

$$\lambda(p-q) = \sum_{i=0}^{r} \alpha_i w^i. \tag{6}$$

Let j be an integer such that $1 \leqslant j \leqslant r$. As $\lambda(p-q)$ is orthogonal to w^j,

$$\sum_{i=0}^{r} \alpha_i (w^i \cdot w^j) = 0. \tag{7}$$

Now observe that $\lambda(p-q)$ cannot be a linear combination of w^1 to w^r as otherwise λ would necessarily be equal to 0. Hence, α_0 is non-zero and we can denote, for each integer i such that $1 \leqslant i \leqslant r$,

$$\beta_i = -\frac{\alpha_i}{\alpha_0}.$$

With this notation, (7) can be rewritten into

$$\sum_{i=1}^{r} \beta_i (w^i \cdot w^j) = w^0 \cdot w^j$$

and the linear system obtained by letting j range between 1 and r is

$$M\beta = b.$$

Note that M has rank r since the vectors w^i are linearly independent. As a consequence, according to Cramer's rule,

$$\beta_i = \frac{\det(M_i)}{\det(M)}.$$

Picking $\alpha_0 = \det(M)$ therefore yields

$$\alpha_i = -\det(M_i)$$

when $1 \leqslant i \leqslant r$. This provides values for the coefficients α_0 to α_r that appear in (6). In particular, with these coefficients,

$$\lambda(p-q) = \det(M)w^0 - \sum_{i=0}^{r} \det(M_i)w^i. \tag{8}$$

Now observe that (5) can be rewritten into

$$d(p,q) = w^0 \cdot \frac{\lambda(p-q)}{\|\lambda(p-q)\|}.$$

Combining this with (8) yields proves the lemma. ∎

Now observe that, when P and Q are rational polytopes, then the vectors w^0 to w^r have rational coordinates. Therefore, we recover the following remark from Lemma 2.1. This remark is a consequence of a more general result due to Stephen Vavasis [12] that was further improved in [9].

Remark 2.2: If P and Q are rational polytopes then $d(P,Q)^2$ is rational.

We are now ready to prove the announced bound on on $d(P,Q)$ in the case when both P and Q are lattice polytopes.

THEOREM 2.3: *If P and Q are disjoint lattice (d,k)-polytopes, then*

$$d(P,Q) \geqslant \frac{1}{k^{2d-1}\sqrt{d}^{3d+2}}.$$

Proof. According to Lemma 2.1,

$$d(P,Q) = \frac{w^0 \cdot a}{\|a\|} \tag{9}$$

where

$$a = \det(M)w^0 - \sum_{i=1}^{r} \det(M_i)w^i. \tag{10}$$

Assuming that P and Q are lattice (d,k)-polytopes, the vectors w^0 to w^r have integers coordinates. It then follows from (10) that all the coordinates of

a are integers. By the assumption that P and Q are disjoint, the numerator in the right-hand side of (10) is at least 1. As a consequence,

$$d(P,Q) \geqslant \frac{1}{\|a\|}. \tag{11}$$

Since P and Q are lattice (d,k)-polytopes, all the w^i are contained in the hypercube $[-k,k]^d$. Hence the absolute value of each entry in the matrices M and M_i is at most dk^2 and, by Hadamard's inequality,

$$|\det(M_i)| \leqslant d^r k^{2r} r^{\frac{r}{2}}$$

for all i. Moreover, the same inequality holds by replacing M in the left-hand side by M_i. Plugging this into (10) yields

$$|a_i| \leqslant (r+1)d^r k^{2r+1} r^{\frac{r}{2}}.$$

It follows that

$$\|a\| \leqslant (r+1)d^{\frac{2r+1}{2}} k^{2r+1} r^{\frac{r}{2}}$$

and according to (11),

$$d(P,Q) \geqslant \frac{1}{(r+1)d^{\frac{2r+1}{2}} k^{2r+1} r^{\frac{r}{2}}}.$$

Finally, observe that $r \leqslant d-1$. Hence,

$$d(P,Q) \geqslant \frac{1}{d^{\frac{3d+2}{2}} k^{2d-1}}$$

as desired. ∎

Note that the distance between the origin of $\mathbb{R}^d$ and the $(d-1)$-dimensional standard simplex is equal to $1/\sqrt{d}$. It turns out that the distance between the origin and any lattice polytope contained in the positive orthant $[0,+\infty[^d$ but that does not contain the origin is at least this value.

LEMMA 2.4: *If P is a lattice polytope contained in $[0,+\infty[^d \backslash \{0\}$, then*

$$d(0,P) \geqslant \frac{1}{\sqrt{d}}.$$

Proof. Let p be a point in P such that $d(0,P) = d(0,p)$. Observe that all the vertices x of P satisfy $\|x\|_1 \geqslant 1$. As any point in P is a convex combination of vertices of P, it follows that $\|p\|_1 \geqslant 1$. However, by the Cauchy–Schwarz inequality $\|p\|_1 \leqslant \sqrt{d}\|p\|_2$, which proves the lemma. ∎

3. Lower bounds in terms of encoding length

We now turn our attention to bounding $d(P,Q)$ in the case when P and Q are rational polytopes, in terms of their binary encoding input data length. We first recall some definitions regarding this quantity, see [11]. If α and β are two relatively prime integers such that β is positive, the *size* of α/β is

$$\mathrm{size}\left(\frac{\alpha}{\beta}\right) = 1 + \big\lceil \log(|\alpha|+1)\big\rceil + \big\lceil \log(|\beta|+1)\big\rceil.$$

In turn, the size of a vector a from $\mathbb{R}^d$ with rational coordinates is

$$\mathrm{size}(a) = d + \sum_{i=1}^{d} \mathrm{size}(a_i).$$

In other words, the size of a vector with rational coordinates is the number of its coordinates plus the sum of the sizes of these coordinates.

If P is a rational polytope, then its vertices have rational coordinates and the *vertex complexity of* P is the smallest number $\nu(P)$ such that $\nu(P)$ is at least d and the size of any vertex of P is at most $\nu(P)$. Still under the assumption that P is a rational polytope, the *facet complexity* of P is the smallest number $\varphi(P)$ such that $\varphi(P)$ is at least d and there exists a family of vectors a^1 to a^n from $\mathbb{Q}^d$ and a family of rational numbers b_1 to b_n such that

$$P = \big\{x \in \mathbb{R}^d : \forall\, i \in \{1, \ldots, d\},\, a^i{\cdot}x \leqslant b_i\big\}$$

and for all i satisfying $1 \leqslant i \leqslant n$,

$$\mathrm{size}(a^i) + \mathrm{size}(b_i) \leqslant \varphi(P).$$

The following is proven in [11] (see Theorem 10.2 therein).

THEOREM 3.1: *If P is rational, then $\nu(P) \leqslant 4d^2\varphi(P)$ and $\varphi(P) \leqslant 4d^2\nu(P)$.*

The size of a matrix can be defined in the same spirit as the size of a vector: if M is a matrix with rational coefficients, then $\mathrm{size}(M)$ is the number of coefficients in M plus the sum of the sizes of these coefficients.

The following statement is the Theorem 3.2 from [11].

THEOREM 3.2: *If M is a square matrix with rational coefficients, then*

$$\mathrm{size}\big(\det(M)\big) \leqslant 2\,\mathrm{size}(M).$$

We now bound the size of a sum and a product of rational numbers.

Proposition 3.3: *If a is a vector from $\mathbb{Q}^d$, then*

$$\text{size}\left(\sum_{i=1}^{d} a_i\right) \leqslant 2\sum_{i=1}^{d}\text{size}(a_i) \tag{12}$$

and

$$\text{size}\left(\prod_{i=1}^{d} a_i\right) \leqslant \sum_{i=1}^{d}\text{size}(a_i). \tag{13}$$

Proof. First recall that the logarithm of a product of numbers is the sum of the logarithms of these numbers. Moreover, the ceiling of a sum of numbers is at most the sum of the ceilings of these numbers. The inequality (13) immediately follows from these two properties. Let us now prove (12).

Consider a vector a contained in $\mathbb{Q}^d$. There exist two relatively prime integers α and β such that the later is positive and

$$\sum_{i=1}^{d} a_i = \frac{\alpha}{\beta}.$$

For any integer i satisfying $1 \leqslant i \leqslant d$, further denote by γ_i and δ_i the two relatively prime integers such that δ_i is positive and

$$a_i = \frac{\gamma_i}{\delta_i}.$$

With these notations,

$$\sum_{i=1}^{d} \frac{\gamma_i}{\delta_i} = \frac{\alpha}{\beta} \tag{14}$$

and, as a consequence,

$$\beta \leqslant \prod_{i=1}^{d} \delta_i \leqslant \prod_{i=1}^{d} \big(\delta_i + 1\big). \tag{15}$$

Now observe that

$$\left|\sum_{i=1}^{d} \frac{\gamma_i}{\delta_i}\right| \leqslant \sum_{i=1}^{d} |\gamma_i| \leqslant \prod_{i=1}^{d}\Big(|\gamma_i| + 1\Big). \tag{16}$$

If follows from (14), (15), and (16) that

$$|\alpha| \leqslant \prod_{i=1}^{d}\Big(|\gamma_i| + 1\Big)\big(\delta_i + 1\big). \tag{17}$$

Finally, observe that

$$\text{size}\left(\frac{\alpha}{\beta}\right) \leqslant 1 + \lceil \log(2|\alpha|) \rceil + \lceil \log(2|\beta|) \rceil.$$

As a consequence,

$$\text{size}\left(\sum_{i=1}^{d} a_i\right) \leqslant 3 + \lceil \log\,|\alpha| \rceil + \lceil \log\,|\beta| \rceil.$$

As the ceiling of a sum is at most the sum of the ceilings of the summands, combining this with (15) and (17) proves the proposition. ∎

Proposition 3.3 allows to bound the size of a scalar product.

PROPOSITION 3.4: *If a and b are two vectors from $\mathbb{Q}^d$, then*

$$\text{size}(a{\cdot}b) \leqslant 2\,\text{size}(a) + 2\,\text{size}(b).$$

Proof. Consider two two vectors a and b in $\mathbb{Q}^d$. By Proposition 3.3,

$$\text{size}(a{\cdot}b) \leqslant 2\sum_{i=1}^{d} \text{size}(a_i b_i) \leqslant 2\sum_{i=1}^{d} \text{size}(a_i) + 2\sum_{i=1}^{d} \text{size}(b_i).$$

Since the size of a vector coincides with the sum of the sizes of its coordinates, this immediately results in the desired inequality. ∎

The following proposition whose proof is straightforward provides the smallest possible positive rational number with a given size.

PROPOSITION 3.5: *If x is a positive rational number, then*

$$\frac{4}{2^{\text{size}(x)}} \leqslant x \leqslant \frac{2^{\text{size}(x)}}{4}.$$

We can now state our lower bound on the distance of P and Q when these polytopes are rational, in terms of their binary encoding input data length. In the statement of the following theorem and its proof, we denote

$$\nu(P, Q) = \max\{\nu(P), \nu(Q)\}.$$

and

$$\varphi(P, Q) = \max\{\varphi(P), \varphi(Q)\}.$$

THEOREM 3.6: *If P and Q are disjoint rational polytopes, then*

$$d(P, Q) \geqslant \frac{8}{2^{4\nu(P,Q)(2d)^4}}$$

and

$$d(P,Q) \geqslant \frac{8}{2^{4\varphi(P,Q)(2d)^6}}.$$

Proof. In this proof, we consider the vectors w^0 to w^r as well as the matrices M and M_1 to M_r that have been associated to P and Q at the beginning of Section 2. Recall that the vectors w^0 to w^r are obtained by subtracting from one another two vertices of P, two vertices of Q, or a vertex of P and a vertex of Q. As a consequence, it follows from the first inequality in the statement of Proposition 3.3, that for every integer i satisfying $0 \leqslant i \leqslant r$,

$$\mathrm{size}\big(w^i\big) \leqslant 2\nu(P,Q).$$

In turn, for any two integers i and j satisfying $0 \leqslant i \leqslant j \leqslant r$, it follows from Proposition 3.4 that the size of $w^i{\cdot}w^j$ can be bounded as

$$\mathrm{size}\big(w^i{\cdot}w^j\big) \leqslant 4\nu(P,Q)$$

and by Theorem 3.2,

$$\mathrm{size}\big(\det(M)\big) \leqslant 8r^2\nu(P,Q).$$

In addition, the same inequality holds when replacing M by any of the matrices M_1 to M_r. Now consider the vector

$$a = \det(M)w^0 - \sum_{i=1}^{r}\det(M_i)w^i$$

and observe that a_i is the scalar product between the vector from $\mathbb{R}^{r+1}$ whose coordinates are $\det(M)$ and $\det(M_1)$ to $\det(M_r)$ with the one whose coordinates are w_i^0 and $-w_i^1$ to $-w_i^r$. Therefore, by Proposition 3.4,

$$\mathrm{size}(a) \leqslant 2d(8r^2+1)(r+1)\nu(P,Q)$$

and

$$\mathrm{size}\big(w^0{\cdot}a\big) \leqslant 4\big(d(8r^2+1)(r+1)+1\big)\nu(P,Q).$$

Now recall that r is at most $d-1$. Hence,

$$\mathrm{size}(a) \leqslant 16d^4\nu(P,Q) \tag{18}$$

and

$$\mathrm{size}\big(w^0{\cdot}a\big) \leqslant 32d^4\nu(P,Q).$$

In turn, according to Lemma 2.1 and Proposition 3.5,

$$d(P,Q) \geqslant \frac{4}{2^{32d^4\nu(P,Q)}\|a\|}. \tag{19}$$

Now, by (18), Propositions 3.4 and 3.5 yield

$$\|a\|^2 \leqslant \frac{2^{64d^4\nu(P,Q)}}{4}. \tag{20}$$

Combining (19) with (20) and using Theorem 3.1 completes the proof. ∎

4. Upper bounds

In this section, k is fixed and we further consider two positive integers σ and δ. Our aim is to build two lattice $(\delta(\sigma+1), k)$-polytopes P and Q such that

$$d(P,Q) \leqslant \frac{\sqrt{\delta\sigma}}{\big(k(\delta-1)\big)^{\sigma}}. \tag{21}$$

when δ is at least 4 (see Theorem 4.5). Before we build P and Q, let us state and prove the main result of the section, which is a consequence of (21).

THEOREM 4.1: *Consider a number α in $]0,1[$. For any large enough d, there exist two disjoint lattice (d,k)-polytopes P and Q such that*

$$d(P,Q) \leqslant \frac{1}{k^{d^\alpha} d^{(1-\alpha)d^\alpha}}.$$

Proof. Let β be a number in the interval $]\alpha, 1[$. Assume that

$$d \geqslant 8^{\frac{1}{1-\beta}} \tag{22}$$

and denote

$$\left\{\begin{array}{l} \sigma = \lfloor d^\beta \rfloor \text{ and} \\ \delta = \left\lfloor \dfrac{d}{\sigma+1} \right\rfloor. \end{array}\right. \tag{23}$$

Observe that σ is at least 1. In addition, it follows from (22) that

$$d \geqslant 4d^\beta + 4$$

and as a consequence, δ is at least 4.

As announced above (see also Theorem 4.5 below), under these conditions on σ and δ, there exist two lattice $(\delta(\sigma+1), k)$-polytopes P and Q such that

$$d(P,Q) \leqslant \frac{\sqrt{\delta\sigma}}{\big(k(\delta-1)\big)^{\sigma}}. \tag{24}$$

However, by (23), d is at least $(\sigma+1)\delta$. Therefore, P and Q are also lattice (d,k)-polytopes. Moreover, replacing σ and δ in the right-hand side of (24) by their expressions as functions of d and β yields

$$d(P,Q) \leqslant \frac{\sqrt{\lfloor d^\beta \rfloor \left\lfloor \frac{d}{\lfloor d^\beta \rfloor + 1} \right\rfloor}}{k^{\lfloor d^\beta \rfloor}\left(\left\lfloor \frac{d}{\lfloor d^\beta \rfloor + 1} \right\rfloor - 1\right)^{\lfloor d^\beta \rfloor}}. \tag{25}$$

Now observe that the right-hand side of (25) behaves like

$$\frac{\sqrt{d}}{k^{d^\beta} d^{(1-\beta)d^\beta}}$$

as d goes to infinity. Since α is less than β,

$$\frac{\sqrt{d}}{k^{d^\beta} d^{(1-\beta)d^\beta}} < \frac{1}{k^{d^\alpha} d^{(1-\alpha)d^\alpha}}$$

when d is large enough. Hence, the right-hand side of (25) is less than

$$\frac{1}{k^{d^\alpha} d^{(1-\alpha)d^\alpha}}$$

for any large enough d, as desired. ∎

It should be noted that, taking α equal to $1/2$ in the statement of Theorem 4.1 results in Theorem 1.2. From now we denote $\delta(\sigma+1)$ by d. Let us proceed to building the two announced lattice (d,k)-polytopes P and Q.

Denote by a the vector from $\mathbb{Z}^{\sigma+1}$ whose coordinate i is

$$a_i = \big(k(1-\delta)\big)^{i-1}.$$

A vector $\overline{x}$ in $\mathbb{R}^d$ can be built from any vector x in $\mathbb{R}^{\sigma+1}$ by taking

$$\overline{x}_i = x_r$$

for every integer i where $r = \lfloor (i-1)/\delta \rfloor + 1$. Equivalently,

$$\overline{x} = (\underbrace{x_1,\ldots,x_1}_{\delta \text{ times}}, \underbrace{x_2,\ldots,x_2}_{\delta \text{ times}}, \ldots, \underbrace{x_{\sigma+1},\ldots,x_{\sigma+1}}_{\delta \text{ times}}).$$

Denote by P the convex hull of the lattice points x contained in $[0,k]^d$ that satisfy $\overline{a}\cdot x = 0$. Likewise, denote by Q the convex hull the lattice points x in $[0,k]^d$ such that $\overline{a}\cdot x = 1$. In order to prove that P and Q satisfy the inequality (21), we will exhibit a point in P and a point in Q whose distance is at most the right-hand side of this inequality.

Consider the $(\sigma+1)\times(\sigma+1)$ matrix

$$M_P = \begin{bmatrix} 0 & A & A & \cdots & A \\ 0 & B & C & \cdots & C \\ 0 & 0 & B & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & C \\ 0 & 0 & \cdots & 0 & B \end{bmatrix}$$

where

$$\begin{cases} A = (\delta-1)k/\delta, \\ B = 1/\delta, \text{ and} \\ C = A + B. \end{cases}$$

Recall that we identify the points from $\mathbb{R}^{\sigma+1}$ to the vector of their coordinates. In particular, the columns of M_P are points from $\mathbb{R}^{\sigma+1}$.

PROPOSITION 4.2: *If x is a column of M_P, then $\overline{x}$ belongs to P.*

Proof. Let x be the column i of M_p. Observe that, if i is equal to 1, then $\overline{a}\cdot\overline{x} = 0$ and in particular, $\overline{x}$ belongs to P. Now assume that i is at least 2 and consider an integer s such that $1 \leqslant s \leqslant \delta$. Denote by u^s the lattice point in $[0,k]^d$ whose coordinates are given by

$$u_j^s = \begin{cases} k \text{ if } 1 \leqslant j \leqslant \delta(i-1) \text{ and } \big((j-1) \bmod \delta\big) + 1 \neq s, \\ 1 \text{ if } \delta < j \leqslant \delta i \text{ and } \big((j-1) \bmod \delta\big) + 1 = s, \text{ and} \\ 0 \text{ otherwise.} \end{cases}$$

Note that u^s is a point in P because $\overline{a}\cdot u^s = 0$. As the barycenter of the points u^s when s ranges from 1 to δ is precisely $\overline{x}$ this proves the proposition. ∎

Now consider the $(\sigma+1)\times(\sigma+1)$ matrix M_Q obtained from M_P by adding $1/\delta$ to all the entries of the first row:

$$M_Q = \begin{bmatrix} B & C & C & \cdots & C \\ 0 & B & C & \cdots & C \\ 0 & 0 & B & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & C \\ 0 & 0 & \cdots & 0 & B \end{bmatrix}.$$

PROPOSITION 4.3: *If x is a column of M_Q, then $\overline{x}$ belongs to Q.*

Proof. Let x be the column i of M_Q. Consider an integer s such that $1 \leqslant s \leqslant \delta$ and denote by v^j the lattice point in $[0,k]^d$ whose coordinates are

$$v_j^s = \begin{cases} k \text{ if } 1 \leqslant j \leqslant \delta(i-1) \text{ and } \big((j-1) \bmod \delta\big) + 1 \neq s, \\ 1 \text{ if } 1 \leqslant j \leqslant \delta i \text{ and } \big((j-1) \bmod \delta\big) + 1 = s, \text{ and} \\ 0 \text{ otherwise.} \end{cases}$$

By construction, $\overline{a}\cdot v^s = 1$ and v^s is a point in Q. The proposition then follows from the observation that $\overline{x}$ is the barycenter of the points v^1 to v^δ. ∎

For any integer i such that $0 \leqslant i \leqslant \sigma$, we denote the column $i+1$ of the matrix M_P by p^i and the column $i+1$ of M_Q by q^i.

Assume that δ is at least 3 and consider the points

$$p = \left(1 - \theta\frac{\big(k(\delta-1)\big)^\sigma - 1}{\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^\sigma}\right)p^0 + \sum_{i=1}^{\sigma}\frac{\theta}{\big(k(\delta-1)\big)^i}p^i \tag{26}$$

and

$$q = \left(\frac{k(\delta-1)+1}{k(\delta-1)} - \theta\frac{\big(k(\delta-1)\big)^\sigma - 1}{\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^\sigma}\right)q^0 + \sum_{i=1}^{\sigma}\frac{\theta(1+(-1)^i)}{\big(k(\delta-1)\big)^i}q^i \tag{27}$$

where

$$\theta = \frac{\big(k(1-\delta)-1\big)\big(k(1-\delta)\big)^{\sigma-1}}{\big(k(1-\delta)\big)^\sigma - 1}.$$

These points are defined as linear combinations of the columns of M_P and M_Q. If δ is at least 4, they are convex combinations of these columns.

PROPOSITION 4.4: *If δ is at least 4, then p and q are convex combinations of the columns of M_P and M_Q, respectively.*

Proof. It suffices to show that the coefficients in the right-hand sides of the equations (26) and (27) are non-negative and sum to 1. Assume that δ is at least 4. In that case, θ is non-zero and its inverse is

$$\frac{1}{\theta} = \frac{1}{k(\delta-1)+1}\left(k(\delta-1) + \frac{1}{\big(k(1-\delta)\big)^{\sigma-1}}\right). \tag{28}$$

Since σ is positive,

$$\left|\frac{1}{\big(k(1-\delta)\big)^{\sigma-1}}\right| \leqslant 1. \tag{29}$$

It follows from (28) and (29) that $1/\theta$ and therefore θ are positive numbers. Hence, all the coefficients in the right-hand sides of (26) and (27) are non-negative, except possibly for the coefficient of p^0 in (26) and the coefficient of q^0 in (27). However, observe that (28) implies

$$\frac{1}{\theta} \geqslant \frac{1}{k(\delta-1)+1}\left(k(\delta-1) - \frac{1}{\big(k(\delta-1)\big)^{\sigma-1}}\right)$$

and, as a consequence,

$$\theta \leqslant \frac{\big(k(\delta-1)+1\big)\big(k(\delta-1)\big)^{\sigma-1}}{\big(k(\delta-1)\big)^{\sigma}-1}.$$

It follows that the coefficient of p^0 in the right-hand side of (26) is at least

$$1 - \frac{k(\delta-1)+1}{\big(k(\delta-1)-1\big)k(\delta-1)}.$$

This expression is positive when $k(\delta-1)$ is greater than 2. Hence, the coefficient of p^0 in the right-hand side of (26) is positive when k is at least 4. Likewise, the coefficient of q^0 in the right-hand side of (27) is at least

$$\frac{1}{k(\delta-1)} + 1 - \frac{k(\delta-1)+1}{\big(k(\delta-1)-1\big)k(\delta-1)}$$

which is positive as well when $k(\delta-1)$ is greater than 2. Now observe that

$$\sum_{i=1}^{\sigma}\frac{1}{\big(k(\delta-1)\big)^{i}} = \frac{\big(k(\delta-1)\big)^{\sigma}-1}{\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^{\sigma}}.$$

Therefore, the coefficients in the right-hand side of (26) sum to 1 and the coefficients in the right-hand side of (27) to

$$\frac{k(\delta-1)+1}{k(\delta-1)} + \sum_{i=1}^{\sigma}\frac{\theta}{\big(k(1-\delta)\big)^{i}}.$$

Finally, observe that

$$\sum_{i=1}^{\sigma}\frac{\theta}{\big(k(1-\delta)\big)^{i}} = \theta\frac{\big(k(1-\delta)\big)^{\sigma}-1}{\big(k(1-\delta)-1\big)\big(k(1-\delta)\big)^{\sigma}} = \frac{1}{k(1-\delta)}.$$

Hence, the coefficients in the right-hand side of (26) also sum to 1. ∎

We are now ready to bound the distance between P and Q.

THEOREM 4.5: *If δ is at least 4, then*

$$d(P,Q) \leqslant \frac{\sqrt{\delta\sigma}}{\big(k(\delta-1)\big)^{\sigma}}.$$

Proof. According to Propositions 4.2, 4.3, and 4.4, the points $\overline{p}$ and $\overline{q}$ are contained in P and Q, respectively. Therefore,

$$d(P,Q) \leqslant d(\overline{p},\overline{q}).$$

Now observe that, by construction,

$$d(\overline{p},\overline{q}) = \sqrt{\delta}d(p,q).$$

Hence, it suffices to show that

$$d(p,q) \leqslant \frac{\sqrt{\sigma}}{\big(k(\delta-1)\big)^{\sigma}}.$$

By (26) and (27), the first coordinate of $q-p$ is

$$\begin{aligned} q_1 - p_1 = {} & \frac{k(\delta-1)+1}{\delta k(\delta-1)} - \theta\frac{\big(k(\delta-1)\big)^{\sigma}-1}{\delta\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^{\sigma}} \\ & +\frac{\theta}{\delta}\sum_{i=1}^{\sigma}\frac{1}{\big(k(\delta-1)\big)^{i}} + \frac{k(\delta-1)+1}{\delta}\sum_{i=1}^{\sigma}\frac{\theta}{\big(k(1-\delta)\big)^{i}}. \end{aligned}$$

However, since

$$\sum_{i=1}^{\sigma}\frac{1}{\big(k(\delta-1)\big)^{i}} = \frac{\big(k(\delta-1)\big)^{\sigma}-1}{\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^{\sigma}}$$

and

$$\sum_{i=1}^{\sigma}\frac{\theta}{\big(k(1-\delta)\big)^{i}} = \frac{1}{k(1-\delta)}$$

the first coordinate of $q-p$ is equal to 0. According to (26) and (27) again, for any integer j satisfying $1 \leqslant j \leqslant \sigma$,

$$q_{j+1} - p_{j+1} = \frac{1}{\delta}\frac{\theta}{\big(k(1-\delta)\big)^{j}} + \frac{k(\delta-1)+1}{\delta}\sum_{i=j+1}^{\sigma}\frac{\theta}{\big(k(1-\delta)\big)^{i}}$$

However,

$$\sum_{i=j+1}^{\sigma}\frac{1}{\big(k(1-\delta)\big)^{i}} = \frac{1-\big(k(1-\delta)\big)^{\sigma-j}}{\big(k(\delta-1)+1\big)\big(k(1-\delta)\big)^{\sigma}}$$

and as a consequence,

$$q_{j+1}-p_{j+1}=\frac{\theta}{\delta\big(k(1-\delta)\big)^{\sigma}}=\frac{k(\delta-1)+1}{\delta k(1-\delta)\Big(1-\big(k(1-\delta)\big)^{\sigma}\Big)}$$

This quantity can be bounded as

$$|q_j-p_j|\leqslant\frac{1}{(\delta-1)\Big(\big(k(\delta-1)\big)^{\sigma}-1\Big)}\leqslant\frac{1}{\big(k(\delta-1)\big)^{\sigma}}$$

and therefore,

$$d(p,q)\leqslant\frac{\sqrt{\sigma}}{\big(k(\delta-1)\big)^{\sigma}}$$

as desired. ∎

Theorem 4.1 can be rewritten in terms of the binary encoding input data length of P and Q. Indeed, observe that, setting k to 1, $\nu(P)$ and $\nu(Q)$ are both bounded by $4d$ as the coordinates of the vertices of P and Q are equal to 0 or to 1 and the sizes of these two numbers are 2 and 3. Hence $\nu(P,Q)$ is at most $4d$ as well and $\varphi(P,Q)$ at most $16d^3$ by Theorem 3.1. Therefore, we immediately obtain the following from Theorem 4.1.

THEOREM 4.6: *For any number α in $]0,1[$ there exist two disjoint rational polytopes P and Q with arbitrarily large $\nu(P,Q)$ and $\varphi(P,Q)$ such that*

$$d(P,Q)\leqslant\frac{1}{\left(\dfrac{\nu(P,Q)}{4}\right)^{(1-\alpha)\left(\frac{\nu(P,Q)}{4}\right)^{\alpha}}}$$

and

$$d(P,Q)\leqslant\frac{1}{\left(\dfrac{\varphi(P,Q)}{16}\right)^{\frac{1-\alpha}{3}\left(\frac{\varphi(P,Q)}{16}\right)^{\frac{\alpha}{3}}}}.$$

5. Special cases

From now on, $\varepsilon(d,k)$ denotes the smallest possible distance between two disjoint lattice (d,k)-polytopes. In this section, we focus on certain relevant special cases. The upper bounds stated in Section 4 imply that $\varepsilon(d,k)$ decreases exponentially fast with d but these bounds only hold when d is *large enough*. We will prove a different bound that holds for all d at least 2, according to which $\varepsilon(d,1)$ is at most inverse linear as a function of d. We shall see in Section 6 that this bound on $\varepsilon(d,1)$ is tight when d is equal to 2 or 3.

LEMMA 5.1: *For any d at least 2,*

$$\varepsilon(d,1) \leqslant \frac{1}{\sqrt{d(d-1)}}.$$

Proof. Let P be the diagonal of the hypercube $[0,1]^d$ that is incident to the origin of $\mathbb{R}^d$. Denote by Q the $(d-2)$-dimensional simplex whose vertices are the points x of $\mathbb{R}^d$ whose one of the first $d-1$ coordinates is equal to 1 and whose all other coordinates are equal to 0. Note that P and Q are disjoint as the only point of P whose last coordinate is equal to 0 is the origin of $\mathbb{R}^d$.

The point p of $\mathbb{R}^d$ whose all coordinates are equal to $1/d$ belongs to P. The centroid of Q is the point q whose last coordinate is 0 and whose other coordinates are all equal to $1/(d-1)$. Since

$$d(p,q) = \frac{1}{\sqrt{d(d-1)}},$$

this proves the lemma. ∎

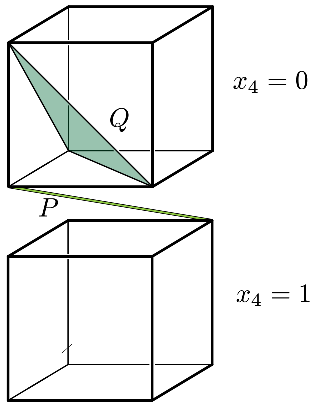

Figure 2. The construction of Lemma 5.1 when d is equal to 4. The cube at the top is the facet of the hypercube $[0,1]^4$ made up of the points x such that $x_4 = 0$ and the cube at the bottom is the opposite facet of $[0,1]^4$.

We complement Lemma 5.1 by showing that $\varepsilon(d,k)$ is at most inverse linear as a function of d and as a function of k for all d and k at least 2.

LEMMA 5.2: *For any k and d at least 2,*

$$\varepsilon(d,k) \leqslant \frac{1}{(d-1)k}.$$

Proof. Let P denote the point of $\mathbb{R}^d$ whose all coordinates are equal to 1. Denote by Q the $(d-1)$-dimensional simplex whose vertices are the origin of $\mathbb{R}^d$ and the points whose one of the first $d-1$ coordinates is equal to $k-1$ and whose all other coordinates are equal to k. Now consider the point q such that

$$q_i = 1 - \frac{k}{(d-1)k^2 + \big((d-1)k-1\big)^2}$$

when $1 \leqslant i \leqslant d-1$ and

$$q_d = 1 + \frac{(d-1)k-1}{(d-1)k^2 + \big((d-1)k-1\big)^2}.$$

This point is the convex combination of the vertices of Q where the coefficient of the origin is $1 - q_d/k$ and the coefficient of all the other vertices of Q is $q_d/(k(d-1))$. The distance of P and q is

$$d(P,q) = \frac{1}{\sqrt{(d-1)k^2 + \big((d-1)k-1\big)^2}}.$$

It suffices to observe that

$$\frac{1}{\sqrt{(d-1)k^2 + \big((d-1)k-1\big)^2}} \leqslant \frac{1}{(d-1)k}$$

when $k \geqslant 2$ in order to complete the proof. ∎

Let us now turn our attention to the case when the dimensions of P and Q are fixed independently on the dimension of the ambient space as, for example when P and Q are two line segments that live in a higher dimensional space.

We recall that the dimension of a non-necessarily convex subset of $\mathbb{R}^d$ is defined as the dimension of its affine hull.

LEMMA 5.3: *For any two disjoint lattice (d,k)-polytopes P and Q,*

$$d(P,Q) \geqslant \varepsilon\big(\dim(P \cup Q), k\big).$$

Proof. The proof is by induction on $d - \dim(P \cup Q)$. If this quantity is equal to 0, then the result is immediate. Let us assume that d is greater than the dimension of $P \cup Q$. In that case, there exists an hyperplane H of $\mathbb{R}^d$ that contains P and Q. Identify $\mathbb{R}^{d-1}$ with the subspace of $\mathbb{R}^d$ spanned by the first $d-1$ coordinates. We can assume that the vectors orthogonal to H do not belong to $\mathbb{R}^{d-1}$ by, if needed using an adequate permutation of the coordinates of $\mathbb{R}^d$. Now consider the orthogonal projection $\pi : \mathbb{R}^d \rightarrow \mathbb{R}^{d-1}$. Since the vectors orthogonal to H do not belong to $\mathbb{R}^{d-1}$, the restriction of π to H is a bijection between H and $\mathbb{R}^{d-1}$. Moreover, $\pi(\mathbb{Z}^d \cap H)$ is a subset of $\mathbb{Z}^{d-1}$. Hence, $\pi(P)$ and $\pi(Q)$ are two disjoint lattice $(d-1, k)$-polytopes and the dimension of $\pi(P) \cup \pi(Q)$ coincides with the dimension of $P \cup Q$. In particular,

$$d - 1 - \dim\big(\pi(P) \cup \pi(Q)\big) = d - \dim\big(P \cup Q\big) - 1.$$

By induction,

$$d\big(\pi(P), \pi(Q)\big) \geqslant \varepsilon\Big(\dim\big(\pi(P) \cup \pi(Q)\big), k\Big) = \varepsilon\big(\dim(P \cup Q), k\big). \tag{30}$$

Finally observe that the distance between two points in H is always at least the distance between their images by π. Therefore,

$$d(P, Q) \geqslant d\big(\pi(P), \pi(Q)\big)$$

and combining this with (30) proves the lemma. ∎

We will see in Section 6 that $\varepsilon(3, 1)$ is equal to $1/\sqrt{6}$ (see for instance Table 1) and that this distance is achieved between a diagonal of the cube $[0, 1]^3$ and a diagonal of one of its square faces. An immediate consequence of Lemma 5.3 is that this holds independently on the dimension of the ambient space.

THEOREM 5.4: *The smallest possible distance between two disjoint line segments whose vertices belong to* $\{0, 1\}^d$ *is* $1/\sqrt{6}$.

6. Computational aspects

In this section, we are interested in computing the explicit value of $\varepsilon(d, k)$, the smallest between two disjoint lattice (d, k)-polytopes. A brute-force strategy is to enumerate all possible pairs of disjoint lattice (d, k)-polytopes. Let us give some properties of that allow to reduce the search space.

By its definition, $\varepsilon(d, k)$ is a non-increasing function of d for all fixed k. We can prove the following stronger statement.

THEOREM 6.1: *$\varepsilon(d,k)$ is a decreasing function of d for all fixed k.*

Proof. Let us identify $\mathbb{R}^{d-1}$ with the subspace of $\mathbb{R}^d$ spanned by the first $d-1$ coordinates. Consider two lattice $(d-1,k)$-polytopes P and Q such that $d(P,Q)$ is equal to $\varepsilon(d-1,k)$. Now consider the map $\phi : \mathbb{R}^{d-1} \to \mathbb{R}^d$ such that $\phi(x)$ is the point of $\mathbb{R}^d$ obtained from x by changing its last coordinate to 1.

Now consider the lattice (d,k)-polytope

$$Q' = \text{conv}\big(\phi(P) \cup Q\big).$$

Denote by p and q a point in P and a point in Q whose distance is equal to $\varepsilon(d-1,k)$. By construction, both q and $\phi(p)$ belong to Q'. Now consider a number λ in the interval $[0,1]$ and denote by δ the squared distance between the points p and $\lambda\phi(p) + (1-\lambda)q$. It should be noted that δ coincides with $d(p,q)^2$ when λ is equal to 0. Observe that

$$\delta = (1-\lambda)^2 d(p,q)^2 + \lambda^2.$$

Differentiating this equality with respect to λ yields

$$\frac{\partial\delta}{\partial\lambda} = 2\lambda\Big(1 + d(p,q)^2\Big) - 2d(p,q)^2.$$

Note that this derivative is negative for all λ close enough to 0. In particular, one can find a value of λ such that δ is less than $d(p,q)^2$. As δ is the squared distance between p and a point in Q', this show that

$$d(P,Q') < d(p,q)$$

Since the right-hand side of this inequality is equal to $\varepsilon(d-1,k)$ and its left hand side is at least $\varepsilon(d,k)$, this proves the lemma. ∎

By the following theorem, in order to compute $\varepsilon(d,k)$ using brute force enumeration of all possible pairs of lattice (d,k)-polytopes, one only needs to consider pairs of disjoint simplices whose dimensions sum to $d-1$.

THEOREM 6.2: *There exist two lattice (d,k)-polytopes P and Q such that*

(i) $d(P,Q)$ *is equal to* $\varepsilon(d,k)$,
(ii) *both P and Q are simplices,*
(iii) $\dim(P) + \dim(Q)$ *is equal to* $d-1$, *and*
(iv) *the affine hulls of P and Q are disjoint.*

Proof. Consider two disjoint lattice (d,k)-polytopes P and Q such that $d(P,Q)$ is equal to $\varepsilon(d,k)$. Among all possible such pairs of polytopes, we choose P and Q in such a way that their number of vertices sum to the smallest possible number. We shall prove that a consequence of this choice is that P and Q satisfy assertions (ii) and (iii) in the statement of the lemma.

Consider a point p in P and a point q in Q such that $d(p,q)$ is equal to $d(P,Q)$. By Carathéodory's theorem, p is a convex combination of a set S_P of at most $\dim(P)+1$ affinely independent vertices of P. Moreover, we can choose S_P in such a way that all the points it contains have a positive coefficient in that convex combination. Equivalently, p lies in the relative interior of $\mathrm{conv}(S_P)$. In that case, $\varepsilon(d,k)$ is achieved as the distance between $\mathrm{conv}(S_P)$ and Q. It then follows from the above choice for P and Q that S_P must is precisely the vertex set of P. As a consequence, P is a simplex that contains p in its relative interior. By the same argument, Q is a simplex a well and q lies in its relative interior which proves assertion (ii).

Let us now turn our attention to assertion (iii). First observe that if $\dim(P)+\dim(Q)$ is less than $d-1$, then $\dim(P\cup Q)$ is at most $d-1$ and by Lemma 5.3, $d(P,Q)\geqslant\varepsilon(d-1,k)$, which would contradict Theorem 6.1 because $d(P,Q)$ is equal to $\varepsilon(d,k)$. This shows that $\dim(P)+\dim(Q)$ is at least $d-1$. Let us now show that the opposite inequality holds.

By convexity, one can associate a positive number α_u with each point in $S_P\cup S_Q$ in such a way that these numbers collectively satisfy

$$\left\{\begin{array}{l}\displaystyle\sum_{u\in S_P}\alpha_u u = p,\\ \displaystyle\sum_{u\in S_P}\alpha_u = 1,\end{array}\right.$$

and the same equalities hold when S_P is replaced by S_Q and p by q. Now consider a vertex v_P of P, a vertex v_Q of Q. As P and Q are simplices, the sets

$$S'_P=\Big\{u-v_P : u\in S_P\backslash\{v_P\}\Big\}$$

and

$$S'_Q=\Big\{u-v_Q : u\in S_Q\backslash\{v_Q\}\Big\}.$$

are linearly independent. Further observe that all the vectors they contain are orthogonal to $p-q$. Hence, these vectors collectively span a linear subspace M of $\mathbb{R}^d$ of dimension at most $d-1$. Assume for contradiction that the dimensions of P and Q sum to at least d. In that case, the dimensions of the subspaces of M spanned by S'_P and by S'_Q also sum to at least d and the intersection of these subspaces has dimension at least one. Let x be a non-zero point in that intersection. This point can be expressed as a linear combination of S'_P: one can associate each point u in $S_P\backslash\{v_P\}$ with a number β_u such that

$$\sum_{u\in S_P\backslash\{v_P\}}\beta_u(u-v_P)=x.$$

As x is non-zero, the coefficients in the left-hand side of this equality cannot all be equal to zero. For any u in S_P, denote $\gamma_u=\beta_u$ when $u\neq v_P$ and

$$\gamma_u=-\sum_{u\in S_P\backslash\{v_P\}}\beta_u$$

when $u=v_P$. With these notations,

$$\sum_{u\in S_P}\gamma_u=0 \tag{31}$$

and

$$\sum_{u\in S_P}\gamma_u u=x. \tag{32}$$

Likewise, one can associate each point u in S_Q with a number γ_u such that (31) and (31) still hold when replacing S_P by S_Q.

Now consider the number

$$\lambda=\min\left\{\frac{\alpha_u}{\gamma_u}:u\in S_P\cup S_Q,\,\gamma_u>0\right\}$$

It follows from this choice for λ that the point $p-\lambda x$ is still contained in P because the coefficients of is decomposition into an affine combination of S_P all remain non-negative. Likewise $q-\lambda x$ still belongs to Q. Further observe that the distance between $p-\lambda x$ and $q-\lambda x$ is still equal to $\varepsilon(d,k)$. However, also by our choice for λ, at least one of the coefficients in the expression of $p-\lambda x$ as a convex combination of S_P or in the expression of $q-\lambda x$ as a convex combination of S_Q must vanish. In other words, $\varepsilon(d,k)$ is achieved by a pair of disjoint lattice simplices whose combined number of vertices is less than that of P and Q. This contradicts the assumption that P and Q have the smallest combined number of vertices among the pairs of disjoint lattice (d,k)-polytopes whose distance is equal to $\varepsilon(d,k)$, which proves assertion (iii).

Finally, in order to prove (iv), observe that the affine hulls of P and Q are contained in two hyperplanes of $\mathbb{R}^d$ orthogonal to $p-q$. These two hyperplanes

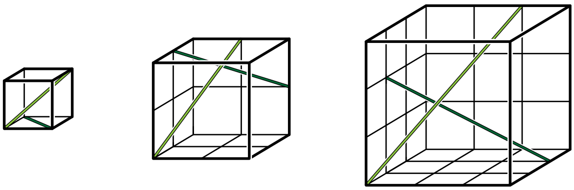

Figure 3. Two line segments P and Q whose distance is equal to $\varepsilon(3,k)$ when $1 \leqslant k \leqslant 3$ (from left to right).

are disjoint because they are parallel and one of them contains p while the other contains q. As a consequence, (iv) holds, as desired. ∎

Using Lemma 6.2, one can compute $\varepsilon(d,k)$ by considering all the pairs of lattice (d,k)-simplices whose dimensions sum to $d-1$. This procedure can be further sped up by doing the computation up to the symmetries of $[0,k]^d$. This allowed to determine the values of $\varepsilon(d,k)$ reported in Table 1.

Let us provide two lattice (d,k)-polytopes that achieve each of the values of $\varepsilon(d,k)$ reported in that table. The smallest possible distance between disjoint lattice $(2,1)$-polytopes is achieved by the origin of $\mathbb{R}^2$ and the diagonal of $[0,1]^2$ that doesn't contain the origin. The smallest possible distance between disjoint lattice $(2,2)$-polytopes is achieved by the point $(0,1)$ and the line segment with vertices $(0,0)$ and $(1,2)$. For all the other values of k considered in Table 1 in the two dimensional case, $\varepsilon(2,k)$ is achieved by the point $(1,1)$ and the line segment with vertices $(0,0)$ and $(k-1,k)$.

In three dimensions, line segments whose distance are $\varepsilon(3,1)$, $\varepsilon(3,2)$, and $\varepsilon(3,3)$ are shown in Figure 3. As already mentioned, $\varepsilon(3,1)$ is achieved by a diagonal of the cube $[0,1]^3$ and a diagonal of a square face. In addition, the line

d	k					
	1	2	3	4	5	6
2	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{13}$	5	$\sqrt{41}$	$\sqrt{61}$
3	$\sqrt{6}$	$5\sqrt{2}$	$\sqrt{299}$			
4	$3\sqrt{2}$					
5	$\sqrt{58}$					

Table 1. A few values of $1/\varepsilon(d,k)$.

segment with vertices $(0,0,0)$ and $(1,2,2)$ is at distance $\varepsilon(3,2)$ of the segment with vertices $(0,1,2)$ and $(2,2,1)$. Similarly, the line segment with vertices $(0,0,0)$ and $(2,3,3)$ is at distance $\varepsilon(3,3)$ from the segment with vertices $(0,1,2)$ and $(3,2,0)$. In four dimensions, $\varepsilon(4,1)$ is achieved between the diagonal of the hypercube $[0,1]^4$ incident to the origin and the triangle with vertices $(0,0,0,1)$, $(0,1,1,0)$, and $(1,0,1,0)$. In five dimensions, $\varepsilon(5,1)$ is achieved between the diagonal of the hypercube $[0,1]^5$ incident to the origin and the tetrahedron with vertices $(0,0,0,1,1)$, $(0,0,1,0,1)$, $(0,1,1,1,0)$, and $(1,1,0,0,0)$.

Acknowledgements. This article is based upon work partially supported by the National Science Foundation under Grant No. DMS-1929284 while the authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, Rhode Island, during the Discrete Optimization: Mathematics, Algorithms, and Computation semester program. The authors are grateful to Amitabh Basu, Santanu Dey, Yuri Faenza, Frédéric Meunier, and Dmitrii Pasechnik for useful discussions and for pointing out relevant references. The first author is partially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant program number RGPIN-2020-06846 and the second author by a grant from the Israel Science Foundation and by the Dresner Chair at the Technion.

References

[1] Noga Alon and Văn H. Vũ, *Anti-Hadamard matrices, coin weighing, threshold gates, and indecomposable hypergraphs*, Journal of Combinatorial Theory, Series A **79** (1997), 133–160.

[2] Amir Beck and Shimrit Shtern, *Linearly convergent away-step conditional gradient for non-strongly convex functions*, Mathematical Programming **164** (2017), 1–27.

[3] Gábor Braun, Alejandro Carderera, Cyrille W. Combettes, Hamed Hassani, Amin Karbasi, Aryan Mokhtari, and Sebastian Pokutta, *Conditional gradient methods*, arXiv:2211.14103 (2022).

[4] Gábor Braun, Sebastian Pokutta and Robert Weismantel, *Alternating linear minimization: revisiting von Neumann's alternating projections*, arXiv:2212.02933 (2022).

[5] David H. Gutman and Javier F. Peña, *The condition of a function relative to a polytope*, arXiv:1802.00271 (2018).

[6] Simon Lacoste-Julien and Martin Jaggi, *On the global linear convergence of Frank–Wolfe optimization variants*, Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS), 2015, pp. 496–504.

[7] Javier F. Peña, *Generalized conditional subgradient and generalized mirror descent: duality, convergence, and symmetry*, arXiv:1903.00459 (2019).

[8] Javier F. Peña and Daniel Rodríguez, *Polytope conditioning and linear convergence of the Frank–Wolfe algorithm*, Mathematics of Operations Research **44** (2018), no. 1, 1–18.

[9] Alberto Del Pia, Santanu S. Dey and Marco Molinaro, *Mixed-integer quadratic programming is in NP*, Mathematical Programming **162** (2017), 225–240.

[10] Neil J. A. Sloane Ronald L. Graham, *Anti-Hadamard matrices*, Linear Algebra and its Applications **62** (1984), 113–137.

[11] Alexander Schrijver, *Theory of linear and integer programming*, Wiley, 1998.

[12] Stephen A. Vavasis, *Quadratic programming is in NP*, Information Processing Letters **36** (1990), 73–77.