SIAM J. DISCRETE MATH.
Vol. 38, No. 4, pp. 2643–2664

KISSING POLYTOPES*

ANTOINE DEZA†, SHMUEL ONN‡, SEBASTIAN POKUTTA§, AND LIONEL POURNIN¶

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 bounds on this distance and discuss its exact computation. We also give similar bounds for disjoint rational polytopes whose binary encoding length is prescribed.

Key words. facial distance, vertex-facet distance, pyramidal width, alternating projections, distances in geometric lattices, lattice polytopes

MSC codes. 05B20, 51M15, 52B05, 52B20, 52C07

DOI. 10.1137/24M1640859

1. Introduction.

In general, the distance between two disjoint convex bodies P and Q contained in $\mathbb{R}^d$ can become 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 Braun, Pokutta, and Weismantel [4]. They provide an algorithm that computes a point in $P \cap Q$ when that intersection is nonempty or certifies that $P \cap Q$ is empty. In the latter case, the number of calls to a linear optimization oracle over P and Q required to certify 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 become.

Our study is related to the notion of *facial distance* considered by Peña and Rodríguez [11, section 2] and Gutman and Peña [8, 10]. It is also linked to the *vertex-facet distance* of a polytope investigated by Beck and Shtern [2] and to the closely

*Received by the editors February 21, 2024; accepted for publication (in revised form) July 15, 2024; published electronically October 9, 2024.
https://doi.org/10.1137/24M1640859
Funding: This article is based upon work partially supported by the National Science Foundation under grant DMS-1929284. The first author was partially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant program RGPIN-2020-06846. The second author was supported by a grant from the Israel Science Foundation and by the Dresner Chair at the Technion. The third author was supported by the DFG Cluster of Excellence MATH+ (EXC-2046/1, project 390685689) funded by the Deutsche Forschungsgemeinschaft (DFG).
†McMaster University, Hamilton, ON L8S 4L8, Canada (deza@mcmaster.ca).
‡Technion – Israel Institute of Technology, Haifa, 32000, Israel (onn@technion.ac.il).
§Zuse Institute Berlin, Berlin, Germany (pokutta@zib.de).
¶Université Paris 13, Villetaneuse, France (lionel.pournin@univ-paris13.fr).

related *pyramidal width* studied by Lacoste-Julien and Jaggi [9] and Rademacher and Shu [12]. We refer the reader to the survey by Braun et al. [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

$$\Phi(P) = \min\Big\{ d\big(F, \operatorname{conv}(\mathcal{V} \setminus F)\big) : F \in \mathcal{F} \Big\},$$

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 of P is defined as

$$\Delta(P) = \min\Big\{ d\big(\operatorname{aff}(F), \operatorname{conv}(\mathcal{V} \setminus F)\big) : F \in \overline{\mathcal{F}} \Big\}, \tag{1.1}$$

where $\overline{\mathcal{F}}$ is the set of the facets of P, as shown in [11, section 2]. Bounds have been given on the smallest possible vertex-facet distance of 0/1-simplices [1, 6]. In particular, Alon and Vũ [1, Theorem 3.2.2] show that

$$\frac{1}{\sqrt{2}^{d \log d - 2d + o(d)}} \leq \min \Delta(S) \leq \frac{1}{\sqrt{2}^{d \log d - 4d + o(d)}}, \tag{1.2}$$

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

The results of Vavasis on the complexity of quadratic optimization [14], generalized by Del Pia, Dey, and Molinaro in [5], imply as a special case that the squared distance between two rational polytopes is a rational number. Our work is concerned with 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*

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

We shall in fact prove a stronger bound (see Theorem 2.3) of which Theorem 1.1 is a consequence. 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 6.5). 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*

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

As above, Theorem 1.2 follows from a stronger bound (see Theorem 3.2). We also give an upper bound on the smallest possible distance between two rational polytopes whose binary encoding length is prescribed (see Theorem 6.6).

By its definition, the facial distance of a polytope is a distance between two polytopes. Inversely, the distance between two disjoint polytopes P and Q is the distance between two of their faces that belong to distinct 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,*

$$\frac{1}{(kd)^{2d}} \leq \min \Phi(P) \leq \frac{1}{\left(k\sqrt{d}\right)^{\sqrt{d}}}, \tag{1.3}$$

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 6.5 and 6.6.

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

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 notation and make 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 dentoe the unique face of Q that contains q in its relative interior. We remark that f_P and f_Q are contained in two parallel hyperplanes orthogonal to $p-q$. This situation is illustrated in Figure 1, where P and Q are two 0/1-polytopes, f_P is the diagonal of the unit cube, and f_Q is 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 from u^1-u^0 to $u^{\dim(f_P)}-u^0$ are linearly independent. Similarly,

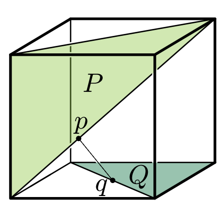

FIG. 1. *Two 0/1-polytopes P and Q and points p and q such that $d(P,Q)$ is equal to $d(p,q)$.*

pick a family from v^0 to $v^{\dim(f_Q)}$ of vertices of f_Q such that the vectors from $v^1 - v^0$ to $v^{\dim(f_Q)} - v^0$ are linearly independent. Consider the set

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

and extract from it a subset of linearly independent vectors from w^1 to w^r that span the same subspace of $\mathbb{R}^d$ as S. Observe that r is at most $d-1$ because these vectors are linearly independent and all of them are orthogonal to $p-q$. Further denote by w^0 the difference $u^0 - v^0$. Since f_P and f_Q are contained ino two parallel hyperplanes orthogonal to $p-q$, the scalar product $(x-y)\cdot(p-q)$, where x belongs to f_P and y to f_Q, does not depend on which points x and y are chosen in f_P and f_Q.

As a consequence, the equality

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

can be rewritten as

$$d(p,q) = w^0 \cdot \frac{(p-q)}{\|p-q\|}. \tag{2.1}$$

We will express the quotient in the right-hand side of (2.1) using the vectors w^i. In order to do that, consider the $r \times r$ matrix M whose term in row i and column j is $w^i \cdot w^j$, the column vector b whose coordinates are from $w^0 \cdot w^1$ to $w^0 \cdot w^r$, and the matrix M_i 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 from w^0 to w^r. Hence, there exist $r+1$ coefficients α_0 to α_r such that

$$p - q = \sum_{i=0}^{r} \alpha_i w^i. \tag{2.2}$$

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

$$\sum_{i=0}^{r} \alpha_i (w^i \cdot w^j) = 0. \tag{2.3}$$

Since $p-q$ is nonzero and orthogonal to the vectors from w^1 to w^r, it cannot be a linear combination of these vectors. It immediately follows that α_0 is nonzero, and we can denote, for each integer i such that $1 \le i \le r$,

$$\beta_i = -\frac{\alpha_i}{\alpha_0}. \tag{2.4}$$

With this notation, (2.3) can be rewritten as

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

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

$$M\beta=b,$$

where β is the column vector whose coordinates are from β_1 to β_r. Observe that M has rank r since the vectors w^i are linearly independent and that its determinant is therefore nonzero. As a consequence, according to Cramer's rule,

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

and, by (2.4), one can rewrite (2.2) as

$$\lambda(p-q)=\det(M)w^0-\sum_{i=0}^{r}\det(M_i)w^i, \tag{2.5}$$

where

$$\lambda=\frac{\det(M)}{\alpha_0}.$$

Finally, observe that (2.1) can be rewritten as

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

Combining this with (2.5) proves the lemma. ∎

Now observe that when P and Q are rational polytopes, the vectors from w^0 to w^r all have rational coordinates. In particular, we recover the following remark from Lemma 2.1. This remark is also a consequence of a more general result due to Vavasis [14] that was further improved in [5].

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 lower bound on the distance of P and 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*

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

Proof. According to Lemma 2.1,

$$d(P,Q)=\frac{w^0\cdot a}{\|a\|}, \tag{2.6}$$

where

$$a=\det(M)w^0-\sum_{i=1}^{r}\det(M_i)w^i. \tag{2.7}$$

Assuming that P and Q are lattice (d,k)-polytopes, the vectors from w^0 to w^r have integer coordinates. It then follows from (2.7) 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 (2.6) is then at least 1. As a consequence,

$$d(P,Q) \geq \frac{1}{\|a\|}. \tag{2.8}$$

Since both P and Q are lattice (d,k)-polytopes, the vectors w^i are all 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,

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

for all i. Moreover, the same inequality holds when M_i is replaced by M in the left-hand side. Plugging this into (2.7) yields

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

It follows that

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

and according to (2.8),

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

Finally, recall that r is at most $d-1$. Hence, this implies

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

which completes the proof. ∎

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*

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

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

3. Upper bounds. In this section, k is fixed, and we consider two positive integers σ and δ. Most of this section is devoted to proving the following theorem.

THEOREM 3.1. *If δ is at least 4, then there exist two lattice (d,k)-polytopes P and Q, where d is equal to $\delta(\sigma+1)$, such that*

$$d(P,Q) \leq \frac{\sqrt{\delta\sigma}}{\big(k(\delta-1)\big)^{\sigma}}. \tag{3.1}$$

Before we explicitly build two polytopes that allow us to establish Theorem 3.1, we prove the main result of the section as a consequence of this theorem.

THEOREM 3.2. *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*

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

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

$$d \geq 8^{\frac{1}{1-\beta}}, \tag{3.2}$$

and denote

$$\begin{cases} \sigma = \lfloor d^\beta \rfloor \text{ and} \\ \delta = \left\lfloor \dfrac{d}{\sigma+1} \right\rfloor. \end{cases} \tag{3.3}$$

Observe that σ is at least 1. In addition, (3.2) can be rewritten as

$$d \geq 8d^\beta.$$

As d^β is at least 1, it follows that

$$d \geq 4d^\beta + 4,$$

and as a consequence, δ is at least 4.

According to Theorem 3.1, under these conditions on σ and δ, there exist two lattice $(\delta(\sigma+1),k)$-polytopes P and Q such that

$$d(P,Q) \leq \frac{\sqrt{\delta\sigma}}{\big(k(\delta-1)\big)^\sigma}. \tag{3.4}$$

However, by (3.3), 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 (3.4) by their expressions as functions of d and β yields

$$d(P,Q) \leq \frac{\sqrt{\lfloor d^\beta \rfloor \left\lfloor \frac{d}{\lfloor d^\beta \rfloor + 1} \right\rfloor}}{k^{\lfloor d^\beta \rfloor} \left(\left\lfloor \frac{d}{\lfloor d^\beta \rfloor + 1} \right\rfloor - 1 \right)^{\lfloor d^\beta \rfloor}}. \tag{3.5}$$

Now observe that the right-hand side of (3.5) 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 (3.5) is less than

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

for any large enough d, as desired. ∎

Taking α equal to $1/2$ in the statement of Theorem 3.2 results in Theorem 1.2. In turn, Theorem 1.2 implies the upper bound stated by Theorem 1.3 on the smallest possible facial distance of a lattice (d,k)-polytope as follows.

THEOREM 3.3. *For any positive k and large enough d,*

$$\min \Phi(P) \le \frac{1}{\left(k\sqrt{d}\right)^{\sqrt{d}}},$$

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

Proof. Consider the two polytopes P and Q provided by Theorem 1.2. If P and Q are contained in distinct parallel hyperplanes, both P and Q are faces of $\mathrm{conv}(P\cup Q)$. Hence, by definition of the facial distance,

$$\Phi\big(\mathrm{conv}(P\cup Q)\big) \le d(P,Q),$$

which implies the desired bound. If P and Q are not contained in parallel hyperplanes, this bound can still be derived from the observation that the distance between P and Q is also the distance between a face F of P and a face G of Q contained in two parallel hyperplanes, both of which are still lattice (d,k)-polytopes. These faces are identified by considering a point p in P and a point q in Q whose distance is equal to $d(P,Q)$. These points belong to the relative interior of a face F of P and a face G of Q, each of which is contained in a hyperplane orthogonal to $p-q$, as desired. ◻

From now on, we denote $\delta(\sigma+1)$ by d. Let us proceed to build two lattice (d,k)-polytopes P and Q that allow us to prove Theorem 3.1.

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

$$\overline{x}_i = x_{\lfloor (i-1)/\delta\rfloor+1}$$

for every integer i. Equivalently,

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

Denote by P the convex hull of the lattice points x contained in the hypercube $[0,k]^d$ that satisfy $\overline{a}\cdot x = 0$. Likewise, denote by Q the convex hull of 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 (3.1), 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

$$M_{\mathsf{P}} = \begin{bmatrix} 0 & A & A & \cdots & A \\ 0 & B & C & \cdots & C \\ 0 & 0 & B & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & C \\ 0 & 0 & \cdots & 0 & B \end{bmatrix},$$

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

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

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

Note that u^s is a point in P because $\overline{a}{\cdot}u^s$ is equal to 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 of the entries in the first row as follows:

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

PROPOSITION 3.5. *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 $1 \leq s \leq \delta$, and denote by v^j the lattice point in $[0,k]^d$ whose coordinates are

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

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 from v^1 to v^δ. ∎

For any integer i such that $0 \leq i \leq \sigma$, 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

$$p = \left(1 - \theta\frac{\big(k(\delta-1)\big)^\sigma - 1}{\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^\sigma}\right)p^0 + \sum_{i=1}^{\sigma}\frac{\theta}{\big(k(\delta-1)\big)^i}p^i \tag{3.6}$$

and

$$q = \left(\frac{k(\delta-1)+1}{k(\delta-1)} - \theta\frac{\big(k(\delta-1)\big)^\sigma - 1}{\big(k(\delta-1)-1\big)\big(k(\delta-1)\big)^\sigma}\right)q^0 + \sum_{i=1}^{\sigma}\frac{\theta(1+(-1)^i)}{\big(k(\delta-1)\big)^i}q^i, \tag{3.7}$$

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 3.6. *If δ is at least 4, then p is a convex combination of the columns of M_{P}, and q is a convex combination of the columns of M_{Q}.*

Proof. It suffices to show that the coefficients in the right-hand sides of (3.6) and (3.7) are nonnegative and sum to 1. Assume that δ is at least 4. In that case, θ is nonzero, and its inverse is expressed as

$$\frac{1}{\theta}=\frac{1}{k(\delta-1)+1}\left(k(\delta-1)+\frac{1}{\big(k(1-\delta)\big)^{\sigma-1}}\right). \tag{3.8}$$

Since σ is positive,

$$\left|\frac{1}{\big(k(1-\delta)\big)^{\sigma-1}}\right|\leq 1. \tag{3.9}$$

It follows from (3.8) and (3.9) that $1/\theta$, and therefore θ are positive numbers. Hence, all the coefficients in the right-hand sides of (3.6) and (3.7) are nonnegative, except possibly for the coefficient of p^0 in (3.6) and the coefficient of q^0 in (3.7). However, observe that according to (3.8),

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

and, as a consequence,

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

It follows that the coefficient of p^0 in the right-hand side of (3.6) 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 (3.6) is positive when k is at least 4. Likewise, the coefficient of q^0 in the right-hand side of (3.7) 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 (3.6) sum to 1, and the coefficients in the right-hand side of (3.7) sum 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 (3.7) also sum to 1. ∎

We are now ready to bound the distance between P and Q. Note that the following theorem immediately implies Theorem 3.1.

THEOREM 3.7. *If δ is at least 4, then*

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

Proof. According to Propositions 3.4, 3.5, and 3.6, the points $\overline{p}$ and $\overline{q}$ are contained in P and Q, respectively. As a consequence,

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

Now observe that, by construction,

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

Hence, it suffices to show that

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

By (3.6) and (3.7), 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

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

it follows that the first coordinate of $q-p$ is equal to 0. According to (3.6) and (3.7) again, for any integer j satisfying $1\leq j\leq\sigma$,

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

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

$$|q_{j+1} - p_{j+1}| \leq \frac{1}{(\delta-1)\Big(\big(k(\delta-1)\big)^{\sigma}-1\Big)} \leq \frac{1}{\big(k(\delta-1)\big)^{\sigma}},$$

and therefore,

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

which completes the proof ∎

4. Special cases. From now on, we denote the smallest possible distance between two disjoint lattice (d,k)-polytopes by $\varepsilon(d,k)$. In this section, we focus on certain relevant special cases. The upper bounds stated in section 3 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 5 that this bound on $\varepsilon(d,1)$ is tight when d is equal to 2 or 3.

LEMMA 4.1. *For any d at least 2,*

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

Proof. Let P be the diagonal of $[0,1]^d$ 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 $\mathbb{R}^d$ such that one of the first $d-1$ coordinates of x is equal to 1 and all of its other coordinates are equal to 0. Note that P and Q are disjoint as the only point in P whose last coordinate is equal to 0 is the origin of $\mathbb{R}^d$. The point p of $\mathbb{R}^d$ with all coordinates 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. The construction is illustrated in Figure 2. ∎

We complement Lemma 4.1 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 4.2. *For any k and d at least 2,*

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

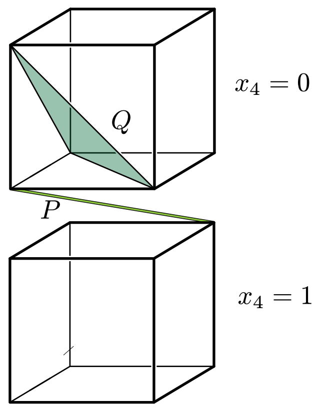

FIG. 2. *The construction of Lemma* 4.1 *when* d *is equal to* 4. *The cube at the top is the facet of* $[0,1]^4$ *made of the points* x *such that* $x_4 = 0$, *and the cube at the bottom is the opposite facet.*

Proof. Let P denote the point of $\mathbb{R}^d$ with all coordinates equal to 1. Further denote by Q the $(d-1)$-dimensional simplex whose vertices are the origin of $\mathbb{R}^d$ and all the points such that one of the first $d-1$ coordinates is equal to $k-1$ and all of the other coordinates are equal to k. Observe that the barycenter of the facet of Q that does not contain the origin is the point x such that the last coordinate of x is k and all of its other coordinates are $k - 1/(d-1)$. Now denote

$$\lambda = \frac{(d-1)(dk-1)}{1+(d-1)k(dk-2)},$$

and consider the point q equal to λx. Note that $0 \leq \lambda \leq 1$ when both k and d are at least 2, and, in that case, q is contained in Q. In addition,

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

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

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

As a consequence, 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 therefore suffices to observe that

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

when $k \geq 2$ 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 subset of $\mathbb{R}^d$ is defined as the dimension of its affine hull.

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

	k					
d	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	$\sqrt{58}$					

LEMMA 4.3. *For any two disjoint lattice* (d,k)*-polytopes* P *and* Q*,*

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

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 a 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 using, if needed, 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 dimensions of $\pi(P)\cup\pi(Q)$ and $P\cup Q$ coincide. In particular,

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

By induction,

$$d\big(\pi(P),\pi(Q)\big)\geq\varepsilon\Big(\dim\big(\pi(P)\cup\pi(Q)\big),k\Big)=\varepsilon\big(\dim(P\cup Q),k\big). \tag{4.1}$$

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

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

and combining this with (4.1) proves the lemma. ∎

We will see in section 5 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 4.3 is that this holds independently on the dimension of the ambient space.

THEOREM 4.4. *The smallest possible distance between two disjoint line segments whose vertices belong to* $\{0,1\}^d$ *is* $1/\sqrt{6}$.

5. 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 strategy is to enumerate all possible pairs of disjoint lattice (d,k)-polytopes. Let us give some properties that allow us to reduce the search space.

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

THEOREM 5.1. *$\varepsilon(d,k)$ 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 $\phi:\mathbb{R}^{d-1}\rightarrow\mathbb{R}^d$ 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

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

Consider a point p in P and a point q in Q whose distance is equal to $\varepsilon(d-1,k)$. By construction, both q and $\phi(p)$ belong to Q'. Let λ be a number in $[0,1]$, 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 shows 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. ∎

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

THEOREM 5.2. *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 whose distance is $\varepsilon(d,k)$. Among all such pairs of polytopes, we choose P and Q in such a way that the sum of the number of their vertices is as small as possible. We shall prove that P and Q then satisfy assertions (ii) and (iii) in the statement of the theorem.

Consider a point p in P and a point q in Q such that $d(p,q)$ is equal to $d(P,Q)$. According to 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 such 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 be 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 also a simplex, and q lies in its relative interior, which proves assertion (ii).

Let us now turn our attention to assertion (iii). 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 4.3, the distance between P and Q is at least $\varepsilon(d-1,k)$, which contradicts Theorem 5.1 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 u in $S_P\cup S_Q$ in such a way that these numbers collectively satisfy

$$\begin{cases} \displaystyle\sum_{u\in S_P}\alpha_u u=p, \\ \displaystyle\sum_{u\in S_P}\alpha_u=1, \end{cases}$$

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 and a vertex v_Q of Q. As P and Q are simplices, the sets

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

and

$$S'_Q=\left\{u-v_Q : u\in S_Q\backslash\{v_Q\}\right\}$$

are linearly independent. Further observe that all the vectors they contain are orthogonal to $p-q$. As a consequence, 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 S'_Q also sum to at least d, and the intersection of these subspaces has dimension at least one. Consider a nonzero point x in that intersection. This point can be expressed as a linear combination of the points from S'_P; that is, 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 nonzero, 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 this notation,

$$\sum_{u\in S_P}\gamma_u=0 \tag{5.1}$$

and

$$\sum_{u\in S_P}\gamma_u u=x. \tag{5.2}$$

Likewise, one can associate each point u in S_Q with a number γ_u such that (5.1) and (5.1) 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 the resulting affine combination of S_P all remain nonnegative. Likewise, $q - \lambda x$ still belongs to Q. Further observe that the distance between $p - \lambda x$ and $q - \lambda x$ is still $\varepsilon(d,k)$. 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 $\varepsilon(d,k)$, and proves assertion (iii).

Finally, observe that the affine hulls of P and Q are contained in two hyperplanes of $\mathbb{R}^d$ orthogonal to $p - q$. These hyperplanes are disjoint because they are parallel, and one of them contains p while the other contains q. Hence, (iv) holds. ∎

Using Theorem 5.2, one can compute $\varepsilon(d,k)$. This is done in practice by generating all the subsets of at most d points from $\{0,1,\dots,k\}^d$, by computing the dimension of their affine hull, and by discarding the subsets such that this dimension is less by at least 2 than the number of points they contain. Then, all the pairs of the remaining subsets whose dimensions sum to $d-1$ are considered, and the distance of their convex hulls is computed. The requirement that the dimensions of the considered pairs sum to $d-1$ can be enforced without loss of generality thanks to Theorem 5.2. This procedure can be further sped up by computing up to the symmetries of $[0,k]^d$. This allowed us to determine the values of $\varepsilon(d,k)$, whose inverses are reported in Table 1.

Let us describe two lattice (d,k)-polytopes that achieve each of the values of $\varepsilon(d,k)$ reported in Table 1. The distance of the origin of $\mathbb{R}^2$ to the diagonal of $[0,1]^2$ that does not contain the origin is equal to $\varepsilon(2,1)$. 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 $(k,k-1)$. These configurations are depicted at the top of Figure 3 when k is equal to 1, 2, or 3.

Pairs of line segments whose distance are $\varepsilon(3,1)$, $\varepsilon(3,2)$, and $\varepsilon(3,3)$ are shown at the bottom of Figure 3. As already mentioned, $\varepsilon(3,1)$ is achieved by a diagonal of

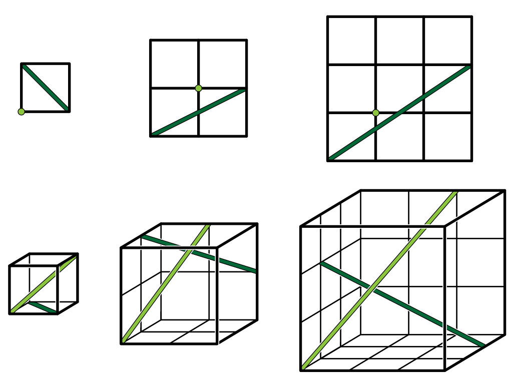

FIG. 3. *Two lattice polytopes P and Q such that $d(P,Q)$ is equal to $\varepsilon(2,k)$ (top) and to $\varepsilon(3,k)$ (bottom) when k is equal to* 1*, to* 2*, or to* 3 *(from left to right).*

the cube $[0,1]^3$ and a diagonal of a square face. In addition, the line 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 $(1,1,0,0,0)$. In the configurations we have just described, $\varepsilon(d,1)$ is always achieved by a diagonal of the hypercube and a $(d-2)$-dimensional simplex. It should be noted that these simplices are not standard, and these configurations are therefore different from the ones that we used in the proof of Lemma 4.1.

6. Estimates in terms of encoding length. We finally turn our attention to bounding $d(P,Q)$ in the case when P and Q are rational polytopes. In practice, the parameter that quantifies the size of a rational polytope is its binary *encoding input data length.* This parameter is the number L of bits required to represent a rational polytope P using either a system of linear inequalities whose set of solutions is P, or a set of points whose convex hull is P. Note, however, that L depends on the choice of a representation for P and may grow large in the case when the representation is redundant. Therefore, we will follow the terminology from [13] and use the vertex and facet complexities of P for our analysis. Let us introduce these quantities. If α and β are two relatively prime integers such that β is positive, the *size* of α/β is

$$\text{size}\left(\frac{\alpha}{\beta}\right) = 1 + \lceil \log_2(|\alpha|+1) \rceil + \lceil \log_2(\beta+1) \rceil.$$

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

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

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 can be described as

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

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

$$\text{size}(a^i) + \text{size}(b_i) \leq \varphi(P).$$

The following is proven in [13] (see Theorem 10.2 therein).

THEOREM 6.1. *If P is rational, then $\nu(P) \leq 4d^2\varphi(P)$ and $\varphi(P) \leq 4d^2\nu(P)$.*

The size of a matrix can be defined in the same spirit as the size of a vector; that is, if M is a matrix with rational coefficients, then size(M) is the number of coefficients in M plus the sum of the sizes of these coefficients.

The following statement is Theorem 3.2 from [13].

THEOREM 6.2. *If M is a square matrix with rational coefficients, then*

$$\text{size}\big(\det(M)\big) \le 2\,\text{size}(M).$$

We now state two propositions that allow us to bound the sizes of rational numbers or vectors. The first one is given as Exercise 1.3.5 in [7].

PROPOSITION 6.3. *If a and b are two vectors from $\mathbb{Q}^d$, then*

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

and

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

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

PROPOSITION 6.4. *If x is a positive rational number, then*

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

We are ready to give lower bounds on the distance of two disjoint rational polytopes P and Q in terms of their binary encoding input data length. In the statement of the following theorem and its proof, we denote

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

and

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

THEOREM 6.5. *If P and Q are disjoint rational polytopes, then*

$$d(P,Q) \ge \frac{8}{2^{4\nu(P,Q)(2d)^4}} \tag{6.1}$$

and

$$d(P,Q) \ge \frac{8}{2^{4\varphi(P,Q)(2d)^6}}. \tag{6.2}$$

Proof. In this proof, we consider the vectors from w^0 to w^r as well as the matrices M and M_1 to M_r that were associated to P and Q at the beginning of section 2. Recall that the vectors from 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 6.3 that

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

for every integer i satisfying $0 \le i \le r$. In turn, for any two integers i and j satisfying $0 \le i \le j \le r$, it follows from Proposition 6.3 that

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

and by Theorem 6.2,

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

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 of the vector from $\mathbb{R}^{r+1}$ whose coordinates are $\det(M)$ and $\det(M_1)$ to $\det(M_r)$ with the vector whose coordinates are w_i^0 and $-w_i^1$ to $-w_i^r$. Therefore, by Proposition 6.3,

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

and

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

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

$$\text{size}(a) \leq 16d^4\nu(P,Q) \tag{6.3}$$

and

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

In turn, according to Lemma 2.1 and Proposition 6.4,

$$d(P,Q) \geq \frac{4}{2^{32d^4\nu(P,Q)}\|a\|}. \tag{6.4}$$

It also follows from (6.3) and Proposition 6.3 that

$$\text{size}\big(\|a\|^2\big) \leq 64d^4\nu(P,Q)$$

and therefore, by Proposition 6.4,

$$\|a\|^2 \leq \frac{2^{64d^4\nu(P,Q)}}{4}. \tag{6.5}$$

The desired lower bound on the distance of P and Q in terms of $\nu(P,Q)$ is obtained by combining the inequalities (6.4) and (6.5). Finally, recall that Theorem 6.1 allows us to upper bound $\nu(P,Q)$ by a function of $\varphi(P,Q)$. Using this bound on $\nu(P,Q)$ in the denominator of the right-hand side of (6.1) proves (6.2). ∎

We can upper bound the smallest possible distance of two disjoint rational polytopes in terms of the same parameters. Such bounds can be easily derived from Theorem 3.2. It should be noted that these bounds no longer depend on k or d. In particular, the following theorem is obtained using two 0/1-polytopes whose dimension gets arbitrarily large, and even though we do not use the dependence on k, their distance decreases exponentially fast with their binary encoding length.

THEOREM 6.6. *For any number α in the interval $]0,1[$ and any positive number N, there exist two disjoint rational polytopes P and Q such that both $\nu(P,Q)$ and $\varphi(P,Q)$ are at least N and the distance of P and Q satisfies*

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

and

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

Proof. The theorem is proved by rewriting Theorem 3.2 in terms of the binary encoding input data length of P and Q. Indeed, consider a positive integer N, and assume that d is at least N and large enough for Theorem 3.2 to hold when k is equal to 1. Consider the two polytopes P and Q provided by the theorem in that case. Recall that by definition of the vertex and facet complexities, both $\nu(P,Q)$ and $\varphi(P,Q)$ are at least d. Hence, both str at least N, as desired. Moreover, both $\nu(P)$ and $\nu(Q)$ are at most $4d$ as the coordinates of the vertices of P and Q are 0 or 1, and the sizes of these two numbers are 2 and 3. Hence, $\nu(P,Q)$ is at most $4d$ and $\varphi(P,Q)$ is at most $16d^3$ according to Theorem 6.1, which can be rewritten as

$$d \geq \frac{\nu(P,Q)}{4}$$

and

$$d \geq \left(\frac{\varphi(P,Q)}{16}\right)^{\frac{1}{3}}.$$

Bounding d using these two inequalities in the denominator of the upper bound on $d(P,Q)$ from Theorem 3.2 proves the theorem. ◻

Acknowledgments. Part of this work was performed while the authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, 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.

REFERENCES

[1] N. Alon and V. H. Vũ, *Anti-Hadamard matrices, coin weighing, threshold gates, and indecomposable hypergraphs*, J. Combin. Theory Ser. A, 79 (1997), pp. 133–160.

[2] A. Beck and S. Shtern, *Linearly convergent away-step conditional gradient for non-strongly convex functions*, Math. Program., 164 (2017), pp. 1–27.

[3] G. Braun, A. Carderera, C. W. Combettes, H. Hassani, A. Karbasi, A. Mokhtari, and S. Pokutta, *Conditional Gradient Methods*, arXiv:2211.14103, 2022.

[4] G. Braun, S. Pokutta, and R. Weismantel, *Alternating Linear Minimization: Revisiting von Neumann's Alternating Projections*, arXiv:2212.02933, 2022.

[5] A. Del Pia, S. S. Dey, and M. Molinaro, *Mixed-integer quadratic programming is in NP*, Math. Program., 162 (2017), pp. 225–240.

[6] R. L. Graham and N. J. A. Sloane, *Anti-Hadamard matrices*, Linear Algebra Appl., 62 (1984), pp. 113–137.

[7] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer, 1993.

[8] D. H. Gutman and J. F. Peña, *The condition number of a function relative to a set*, Math. Program., 188 (2021), pp. 255–294.

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

[10] J. F. Peña, *Generalized Conditional Subgradient and Generalized Mirror Descent: Duality, Convergence, and Symmetry*, arXiv:1903.00459, 2019.

[11] J. F. Peña and D. Rodríguez, *Polytope conditioning and linear convergence of the Frank–Wolfe algorithm*, Math. Oper. Res., 44 (2018), pp. 1–18.

[12] L. Rademacher and C. Shu, *The smoothed complexity of Frank–Wolfe methods via conditioning of random matrices and polytopes*, Math. Stat. Learn., 5 (2022), pp. 273–310.

[13] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, 1998.

[14] S. A. Vavasis, *Quadratic programming is in NP*, Inform. Process. Lett., 36 (1990), pp. 73–77.