Sufficient enlargements of minimal volume for finite dimensional normed linear spaces

M.I. Ostrovskii

Department of Mathematics and Computer Science
St. John’s University
8000 Utopia Parkway
Queens, NY 11439, USA
e-mail: ostrovsm@stjohns.edu
Phone: (718)-990-2469
Fax: (718)-990-1650

April 22, 2008
Abstract. Let $B_Y$ denote the unit ball of a normed linear space $Y$. A symmetric, bounded, closed, convex set $A$ in a finite dimensional normed linear space $X$ is called a sufficient enlargement for $X$ if, for an arbitrary isometric embedding of $X$ into a Banach space $Y$, there exists a linear projection $P : Y \to X$ such that $P(B_Y) \subset A$. The main results of the paper: (1) Each minimal-volume sufficient enlargement is linearly equivalent to a zonotope spanned by multiples of columns of a totally unimodular matrix. (2) If a finite dimensional normed linear space has a minimal-volume sufficient enlargement which is not a parallelepiped, then it contains a two-dimensional subspace whose unit ball is linearly equivalent to a regular hexagon.

Keywords. Banach space, space tiling zonotope, sufficient enlargement for a normed linear space, totally unimodular matrix
1 Introduction

This paper is devoted to a generalization of the main results of [22], where similar results were proved in the dimension two. We refer to [22, 23] for more background and motivation.

1.1 Notation and definitions

All linear spaces considered in this paper will be over the reals. By a space we mean a normed linear space, unless it is explicitly mentioned otherwise. We denote by $B_X (S_X)$ the unit ball (sphere) of a space $X$. We say that subsets $A$ and $B$ of finite dimensional linear spaces $X$ and $Y$, respectively, are linearly equivalent if there exists a linear isomorphism $T$ between the subspace spanned by $A$ in $X$ and the subspace spanned by $B$ in $Y$ such that $T(A) = B$. By a symmetric set $K$ in a linear space we mean a set such that $x \in K$ implies $-x \in K$.

Our terminology and notation of Banach space theory follows [12]. By $B_n^p$, $1 \leq p \leq \infty$, $n \in \mathbb{N}$ we denote the closed unit ball of $\ell_p^n$. Our terminology and notation of convex geometry follows [27].

We use the term ball for a symmetric, bounded, closed, convex set with interior points in a finite dimensional linear space.

Definition 1 [18] A ball $A$ in a finite dimensional normed space $X$ is called a sufficient enlargement (SE) for $X$ (or of $B_X$) if, for an arbitrary isometric embedding of $X$ into a Banach space $Y$, there exists a projection $P : Y \to X$ such that $P(B_Y) \subset A$. A sufficient enlargement $A$ for $X$ is called a minimal-volume sufficient enlargement (MVSE) if $\text{vol} A \leq \text{vol} D$ for each SE $D$ for $X$.

It can be proved, using a standard compactness argument and Lemma 3 below, that minimal-volume sufficient enlargements exist for every finite dimensional space.

Recall that a real matrix $A$ with entries $-1, 0, \text{and} 1$ is called totally unimodular if all minors (that is, determinants of square submatrices) of $A$ are equal to $-1, 0, \text{or} 1$. See [25] and [29, Chapters 19–21] for a survey of results on totally unimodular matrices and their applications.

A Minkowski sum of finitely many line segments in a linear space is called a zonotope (see [3, 13, 14, 27, 28] for basic facts on zonotopes). We consider zonotopes that are sums of line segments of the form $I(x) = \{\lambda x : -1 \leq \lambda \leq 1\}$. For a $d \times m$ totally unimodular matrix with columns $\tau_i (i = 1, \ldots, m)$ and real numbers $a_i$ we consider the zonotope $Z$ in $\mathbb{R}^d$ given by

$$Z = \sum_{i=1}^{m} I(a_i \tau_i).$$

The set of all zonotopes that are linearly equivalent to zonotopes obtained in this way over all possible choices of $m$, of a rank $d$ totally unimodular $d \times m$ matrix, and of positive
numbers $a_i \ (i = 1, \ldots, m)$ will be denoted by $T_d$. Observe that each element of $T_d$ is $d$-dimensional in the sense that it spans a $d$-dimensional subspace. It is easy to describe all $2 \times m$ totally unimodular matrices and to show that $T_2$ is the union of the set of all symmetric hexagons and the set of all symmetric parallelograms.

The class $T_d$ of zonotopes has been characterized in several different ways, see [5, 6, 10, 15, 21, 31]. We shall use a characterization of $T_d$ in terms of lattice tiles. Recall that a compact set $K \subset \mathbb{R}^d$ is called a lattice tile if there exists a basis $\{x_i\}_{i=1}^d$ in $\mathbb{R}^d$ such that

$$\mathbb{R}^d = \bigcup_{m_1, \ldots, m_d \in \mathbb{Z}} \left( \left( \sum_{i=1}^d m_i x_i \right) + K \right),$$

and the interiors of the sets $(\sum_{i=1}^d m_i x_i) + K$ are disjoint. The set

$$\Lambda = \left\{ \sum_{i=1}^d m_i x_i : m_1, \ldots, m_d \in \mathbb{Z} \right\}$$

is called a lattice. The absolute value of the determinant of the matrix whose columns are the coordinates of $\{x_i\}_{i=1}^d$ is called the determinant of $\Lambda$ and is denoted $d(\Lambda)$, see [7, § 3].

**Theorem 1** [15], [6] A $d$-dimensional zonotope is a lattice tile if and only if it is in $T_d$.

It is worth mentioning that lattice tiles in $\mathbb{R}^d$ do not have to be zonotopes, see [32, 16, 17], and [33, Chapter 3].

### 1.2 Statements of the main results

The main result of [21] can be restated in the following way. (A finite dimensional normed space is called polyhedral if its unit ball is a polytope.)

**Theorem 2** A ball $Z$ is linearly equivalent to an MVSE for some $d$-dimensional polyhedral space $X$ if and only if $Z \in T_d$.

In [22] it was shown that for $d = 2$ the statement of Theorem 2 is valid without the restriction of polyhedrality of $X$. The main purpose of the present paper is to prove the same for each $d \in \mathbb{N}$. It is clear that it is enough to prove

**Theorem 3** Each MVSE for a $d$-dimensional space is in $T_d$.

Using Theorem 3 we show that spaces having non-parallelepipedal MVSE cannot be strictly convex or smooth. More precisely, we prove

**Theorem 4** Let $X$ be a finite dimensional normed linear space having an MVSE that is not a parallelepiped. Then $X$ contains a two-dimensional subspace whose unit ball is linearly equivalent to the regular hexagon.
Remarks. 1. Theorem 4 is a simultaneous generalization of [22, Theorem 4] (which is a special case of Theorem 4 corresponding to the case dim \(X = 2\)) and of [19, Theorem 7] (which states that each MVSE for \(\ell^2_2\) is a cube circumscribed about \(B^a_2\)).

2. The fact that \(X\) contains a two-dimensional subspace whose unit ball is linearly equivalent to a regular hexagon does not imply that \(X\) has an MVSE that is not a parallelepiped. A simplest example supporting this statement is \(\ell^3_\infty\).

2 Proof of Theorem 3

First we show that it is enough to prove the following lemmas. It is worth mentioning that our proof of Theorem 3 goes along the same lines as the proof of its two-dimensional version in [22]. The most difficult part of the proof is a \(d\)-dimensional version of the approximation lemma ([22, Lemma 2, p. 380]), it is the contents of Lemma 2 of the present paper. Also, a two-dimensional analogue of Lemma 1 is completely trivial.

**Lemma 1** Let \(T_n \subset \mathbb{R}^d\), \(n \in \mathbb{N}\) be such that \(T_n \in \mathcal{T}_d\), and \(\{T_n\}_{n=1}^\infty\) converges with respect to the Hausdorff metric to a \(d\)-dimensional set \(T\). Then \(T \in \mathcal{T}_d\).

**Remark.** If a sequence \(\{T_n\}_{n=1}^\infty \subset \mathcal{T}_d\) converges to a lower-dimensional set \(T\), the set \(T\) does not have to be in \(\mathcal{T}_{\text{dim}T}\). In fact, as it was already mentioned, \(\mathcal{T}_2\) is the set of all symmetric hexagons and parallelograms. On the other hand, it is easy to find a Hausdorff convergent sequence of elements of \(\mathcal{T}_3\) whose limit is an octagon.

**Lemma 2 (Main lemma)** For each \(d \in \mathbb{N}\) there exist \(\psi_d > 0\) and a function \(t_d : (0, \psi_d) \rightarrow (1, \infty)\) satisfying the conditions:

1. \(\lim_{\varepsilon \downarrow 0} t_d(\varepsilon) = 1\);

2. If \(Y\) is a \(d\)-dimensional polyhedral space, \(B\) is an MVSE for \(Y\), and \(A\) is an SE for \(Y\) satisfying

\[
\text{vol}A \leq (1 + \varepsilon)^d \text{vol}B
\]

for some \(0 < \varepsilon < \psi_d\), then \(A\) contains a ball \(\tilde{A}\) satisfying the conditions:

(a) \(d(\tilde{A}, T) \leq t_d(\varepsilon)\) for some \(T \in \mathcal{T}_d\), where by \(d(\tilde{A}, T)\) we denote the Banach–Mazur distance;

(b) \(\tilde{A}\) is an SE for \(Y\).

**Lemma 3** [22, Lemma 3] The set of all sufficient enlargements for a finite dimensional normed space \(X\) is closed with respect to the Hausdorff metric.
Proof of Theorem 3. (We assume that Lemmas 1 and 2 have been proved.) Let $X$ be a $d$-dimensional space and let $A$ be an MVSE for $X$. Let $\{\varepsilon_n\}_{n=1}^{\infty}$ be a sequence satisfying $\psi_d > \varepsilon_n > 0$ and $\varepsilon_n \downarrow 0$. Let $\{Y_n\}_{n=1}^{\infty}$ be a sequence of polyhedral spaces satisfying

$$
\frac{1}{1 + \varepsilon_n} B_X \subset B_{Y_n} \subset B_X. \quad (2)
$$

Then $A$ is an SE for $Y_n$. Let $B_n$ be an MVSE for $Y_n$. Then $(1 + \varepsilon_n)B_n$ is an SE for $X$. Since $A$ is a minimal-volume SE for $X$, we have

$$
\text{vol} A \leq \text{vol} ((1 + \varepsilon_n)B_n) = (1 + \varepsilon_n)^d \text{vol} B_n.
$$

By Lemma 2 for every $n \in \mathbb{N}$ there exists an SE $\tilde{A}_n$ for $Y_n$ satisfying $\tilde{A}_n \subset A$ and $d(\tilde{A}_n, T_n) \leq t_d(\varepsilon_n)$ (3) for some $T_n \in \mathcal{T}_d$.

The condition (2) implies that $(1 + \varepsilon_n)\tilde{A}_n$ is an SE for $X$.

The sequence $\{(1 + \varepsilon_n)\tilde{A}_n\}_{n=1}^{\infty}$ is bounded (all of its terms are contained in $(1 + \varepsilon_1)A$). By the Blaschke selection theorem [27, p. 50] the sequence $\{(1 + \varepsilon_n)\tilde{A}_n\}_{n=1}^{\infty}$ contains a subsequence convergent with respect to the Hausdorff metric. We denote its limit by $D$, and assume that the sequence $\{(1 + \varepsilon_n)\tilde{A}_n\}_{n=1}^{\infty}$ itself converges to $D$.

Observe that each $\tilde{A}_n$ contains $(1/(1 + \varepsilon_1))B_X$ and is contained in $A$. By (3) we may assume without loss of generality that $T_n$ are balls in $X$ satisfying

$$
\frac{1}{1 + \varepsilon_1} B_X \subset \tilde{A}_n \subset T_n \subset t_d(\varepsilon_n)\tilde{A}_n \subset t_d(\varepsilon_n)A. \quad (4)
$$

It is clear that $D$ is the Hausdorff limit of $\{\tilde{A}_n\}_{n=1}^{\infty}$. From (4) we get that $D$ is the Hausdorff limit of $\{T_n\}_{n=1}^{\infty}$. By Lemma 1 we get $D \in \mathcal{T}_d$.

By Lemma 3 the set $D$ is an SE for $X$. Since $(1 + \varepsilon_n)\tilde{A}_n \subset (1 + \varepsilon_n)A$, and $(1 + \varepsilon_n)A$ is Hausdorff convergent to $A$, we have $D \subset A$. On the other hand, $A$ is an MVSE for $X$, hence $D = A$ and $A \in \mathcal{T}_d$. ■

Proof of Lemma 1. By Theorem 1 the sets $T_n$ are lattice tiles. Let $\{\Lambda_n\}_{n=1}^{\infty}$ be lattices corresponding to these lattice tiles. Since volume is continuous with respect to the Hausdorff metric (see [27, p. 55]), the supremum $\sup_n \text{vol}(T_n)$ is finite. Since $T_n$ is a lattice tile with respect to $\Lambda_n$, the determinant of $\Lambda_n$ satisfies $d(\Lambda_n) = \text{vol}(T_n)$. (Although I have not found this result in the stated form, it is well known. It can be proved, for example, using the argument from [7, pp. 42–43, Proof of Theorem 2].) Hence $\sup_n d(\Lambda_n) < \infty$.

Since $T$ is $d$-dimensional, there exists $r > 0$ such that $rB_2^d \subset T$. Choosing a smaller
$r > 0$, if necessary, we may assume that $rB_2^n \subset T_n$ for each $n$. Therefore the lattices \( \{\Lambda_n\}_{n=1}^\infty \) satisfy the conditions of the selection theorem of Mahler (see, for example, [7, §17], where the reader can also find the standard definition of convergence for lattices). Hence the sequence \( \{\Lambda_n\}_{n=1}^\infty \) contains a subsequence which converges to some lattice $\Lambda$. It is easy to verify that $T$ tiles $\mathbb{R}^d$ with respect to $\Lambda$.

On the other hand, the number of possible distinct columns of a totally unimodular matrix with columns from $\mathbb{R}^d$ is bounded from above by $3^d$, because each entry is 0, 1, or $-1$. (Actually a much better exact bound is known, see [29, p. 299].) Using this we can show that $T$ is a zonotope by a straightforward argument. Also we can use the argument from [27, Theorem 3.5.2] and the observation that a convergent sequence of measures on the sphere of $\ell_2^d$, each of whom has a finite support of cardinality $\leq 3^d$, converges to a measure supported on $\leq 3^d$ points.

Thus, $T$ is a zonotope and a lattice tile. Applying Theorem 1 again, we get $T \in T_d$. ■

3 Proof of the Main Lemma

3.1 Coordinatization

**Proof of Lemma 2.** In our argument the dimension $d$ is fixed. Many of the parameters considered below depend on $d$, although we do not reflect this dependence in our notation.

Since $Y$ is polyhedral, we can consider $Y$ as a subspace of $\ell_\infty^m$. Let $P : \ell_\infty^m \to Y$ be a linear projection satisfying $P(B_\infty^m) \subset A$ (such a projection exists because $A$ is an SE). Let $\tilde{A} = P(B_\infty^m)$. It is easy to see that $\tilde{A}$ is an SE for $Y$. It remains to show that $\tilde{A}$ is close to some $T \in T_d$ with respect to the Banach–Mazur distance.

We consider the standard inner product on $\ell_\infty^m$. (The unit vector basis is an orthonormal basis with respect to this inner product.)

Let $\{\tilde{q}_1, \ldots, \tilde{q}_{m-d}\}$ be an orthonormal basis in ker $P$. Let $\{y_1, \ldots, y_d\}$ be an orthonormal basis in $Y$. Let $\tilde{q}_1, \ldots, \tilde{q}_d$ be such that $\{\tilde{q}_1, \ldots, \tilde{q}_d, \tilde{q}_1, \ldots, \tilde{q}_{m-d}\}$ is an orthonormal basis in $\ell_\infty^m$.

**Lemma 4 (Image Shape Lemma)** Let $P$ and $\tilde{q}_1, \ldots, \tilde{q}_d$ be as above. Denote by $\tilde{Q} = [\tilde{q}_1, \ldots, \tilde{q}_d]$ the matrix whose columns are $\tilde{q}_1, \ldots, \tilde{q}_d$. Let $z_1, \ldots, z_m$ be the columns of the transpose matrix $Q^T$. Then $P(B_\infty^m)$ is linearly equivalent to the zonotope $\sum_{i=1}^m I(z_i) \subset \mathbb{R}^d$.

**Proof.** It is enough to observe that:

(i) Images of $B_\infty^m$ under two linear projections with the same kernel are linearly equivalent. Hence, $P(B_\infty^m)$ is linearly equivalent to the image of the orthogonal projection with the kernel ker $P$.

(ii) The matrix $\tilde{Q}Q^T$ is the matrix of the orthogonal projection with the kernel ker $P$. ■
By Lemma 4 we may replace \( \tilde{A} \) by
\[
Z = \sum_{i=1}^{m} I(z_i)
\]
in the estimate (a) of Lemma 2.

Let \( M = \binom{m}{d} \). We denote by \( u_i \) (\( i = 1, \ldots, M \)) the \( d \times d \) minors of \([y_1, \ldots, y_d]\) (ordered in some way). We denote by \( w_i \) (\( i = 1, \ldots, M \)) the \( d \times d \) minors of \([\tilde{q}_1, \ldots, \tilde{q}_d]\) ordered in the same way as the \( u_i \). We denote by \( v_i \) (\( i = 1, \ldots, \binom{m}{m-d} = M \)) their complementary \((m-d) \times (m-d)\) minors of \([q_1, \ldots, q_{m-d}]\). Using the word *complementary* we mean that all minors are considered as minors of the matrix \([\tilde{q}_1, \ldots, \tilde{q}_d, q_1, \ldots, q_{m-d}]\), see [1, p. 76].

By the Laplacian expansion (see [1, p. 78])
\[
\det[y_1, \ldots, y_d, q_1, \ldots, q_{m-d}] = \sum_{i=1}^{M} \theta_i u_i v_i
\]
and
\[
\det[\tilde{q}_1, \ldots, \tilde{q}_d, q_1, \ldots, q_{m-d}] = \sum_{i=1}^{M} \theta_i w_i v_i
\]
for proper signs \( \theta_i \).

Since the matrix \([\tilde{q}_1, \ldots, \tilde{q}_d, q_1, \ldots, q_{m-d}]\) is orthogonal, we have
\[
\det[\tilde{q}_1, \ldots, \tilde{q}_d, q_1, \ldots, q_{m-d}] = \pm 1.
\]

We need the following result on compound matrices. (We refer to [1, Chapter V] for necessary definitions and background.)

A compound matrix of an orthogonal matrix is orthogonal (see [1, Example 4 on p. 94]).

This result implies, in particular, that the Euclidean norms of the vectors \( \{w_i\}_{i=1}^{M} \) and \( \{v_i\}_{i=1}^{M} \) in \( \mathbb{R}^M \) are equal to 1.

From (6) and (7) we get that either
(a) \( w_i = \theta_i v_i \) for every \( i \)

or
(b) \( w_i = -\theta_i v_i \) for every \( i \).

Without loss of generality, we assume that \( w_i = \theta_i v_i \) for all \( i \) (we replace \( q_1 \) by \( -q_1 \) if it is not the case).

We compute the volume of \( \tilde{A} \) and \( B \) with the normalization that comes from the Euclidean structure introduced above. It is well known (see [20, p. 318]) and is easy to verify that with this normalization
\[
\text{vol} \tilde{A} = \frac{2^d}{\sum_{i=1}^{M} \theta_i u_i v_i} \sum_{i=1}^{M} |v_i|
\]
and
\[ \text{vol } B = \frac{2^d}{\max_i |u_i|} \]
for each MVSE \( B \) for \( Y \).

**Remark.** After the publication of [20] I learned that the formula for the volume of a zonotope used in [20] can be found in [2, Appendix, Section VI].

Since \( \text{vol } \tilde{A} \leq \text{vol } A \), the inequality (1) implies that
\[ \max_i |u_i| \sum_{i=1}^M |v_i| \leq (1 + \varepsilon)^d \left| \sum_{i=1}^M \theta_i u_i v_i \right|. \]  
(8)

By (a) the inequality (8) can be rewritten as
\[ \max_i |u_i| \sum_{i=1}^M |w_i| \leq (1 + \varepsilon)^d \left| \sum_{i=1}^M u_i w_i \right|. \]  
(9)

We need the following two observations:
(i) \( 2^d \sum_{i=1}^M |w_i| \) is the volume of \( Z \) in \( \mathbb{R}^d \).
(ii) The vector \( \{u_i\}_{i=1}^M \) is what is called the Grassmann coordinates, or the Plücker coordinates of the subspace \( Y \subset \mathbb{R}^m \), see [9, Chapter VII] and [30, p. 42]. Recall that \( Y \) is spanned by the columns of the matrix \([y_1, \ldots, y_d]\). It is easy to see that if we choose another basis in \( Y \), the Grassman (Plücker) coordinates will be multiplied by a constant.

We denote by \( Z_\varepsilon (\varepsilon > 0) \) the set of all \( d \)-dimensional zonotopes in \( \mathbb{R}^d \) satisfying the condition (9) with an equality. More precisely, we define \( Z_\varepsilon \) as the set of those \( d \)-dimensional zonotopes \( Z \) in \( \mathbb{R}^d \) for which

(1) There exists \( m \in \mathbb{N} \) and a rank \( d \) matrix \( \tilde{Q} \) of size \( m \times d \) such that \( Z = \sum_{i=1}^m I(z_i) \), where \( z_i \in \mathbb{R}^d \), \( i = 1, \ldots, m \), are rows of \( \tilde{Q} \).

(2) There exists a rank \( d \) matrix \( Y \) of size \( m \times d \) such that, if we denote the \( d \times d \) minors of \( \tilde{Q} \) by \( \{w_i\}_{i=1}^\infty \), where \( M = \binom{m}{d} \), and the \( d \times d \) minors of \( Y \), ordered in the same way as the \( w_i \), by \( \{u_i\}_{i=1}^\infty \), then
\[ \max_i |u_i| \sum_{i=1}^M |w_i| = (1 + \varepsilon)^d \left| \sum_{i=1}^M u_i w_i \right|, \]  
(10)
and there is no \( Y \) for which
\[ \max_i |u_i| \sum_{i=1}^M |w_i| < (1 + \varepsilon)^d \left| \sum_{i=1}^M u_i w_i \right|. \]
Remarks. 1. It is clear that the zonotope property of being in $\mathcal{Z}_\varepsilon$ is invariant under changes of the system of coordinates.

2. We do not consider the class $\mathcal{Z}_0$ because, as it was shown in [21], this class is contained in $\mathcal{T}_d$.

Many objects introduced below depend on $Z$ and $\varepsilon$, although sometimes we do not reflect this dependence in our notation.

Let $Z \in \mathcal{Z}_\varepsilon$. We shall change the system of coordinates in $\mathbb{R}^d$ twice. First we introduce in $\mathbb{R}^d$ a new system of coordinates such that the unit (Euclidean) ball $B_d^2$ of $\mathbb{R}^d$ is the maximal volume ellipsoid in $Z$. From now on we consider the vectors $z_i$ introduced in Lemma 4 as vectors in $\mathbb{R}^d$ and not as $d$-tuples of real numbers.

It is easy to see that the support function of $Z$ is given by

$$h_Z(x) = \sum_{i=1}^{m} |\langle x, z_i \rangle|.$$ (11)

If the vectors in (11) are pairwise distinct, we let $\mu$ be the atomic measure on the unit (Euclidean) sphere $S$ whose atoms are given by $\mu(z_i/\|z_i\|) = \mu(-z_i/\|z_i\|) = \|z_i\|/2$. It is easy to see that

$$h_Z(x) = \int_S \langle x, z \rangle d\mu(z).$$ (12)

The defining formula for $\mu$ should be adjusted in the natural way if some of the vectors in (11) are equal.

Conversely, if $\mu$ is a nonnegative measure on $S$ supported on a finite set, then (12) is a support function of some zonotope (see [27, Section 3.5] for more information on this matter).

Dealing with subsets of $S$ we use the following terminology and notation. Let $x_0 \in S$, $r > 0$. The set $\Delta(x_0, r) := \{x \in S : \|x - x_0\| < r \text{ or } \|x + x_0\| < r\}$, where $\|\cdot\|$ is the $\ell_2$-norm, is called a cap. If $0 < r < \sqrt{2}$, then $\Delta(x_0, r)$ consists of two connected components. In such a case both $x_0$ and $-x_0$ will be considered as centers of $\Delta(x_0, r)$.

We are going to show that if $\varepsilon > 0$ is small, then the inequality (9) implies that all but a very small part of the measure $\mu$ is supported on a union of small caps centered at a set of vectors which are multiples of a set of vectors satisfying the condition: if we write their coordinates with respect to a suitably chosen basis, we get a totally unimodular matrix. Having such a set, it is easy to find $T \in \mathcal{T}_d$ which is close to $Z$ with respect to the Banach–Mazur distance, see Lemma 15.
For any two numbers \( \omega, \delta > 0 \) we introduce the set
\[
\Omega(\omega, \delta) := \{ x \in S : \mu(\Delta(x, \omega)) \geq \delta \}
\]
(recall that by \( S \) we denote the unit sphere of \( \ell_2^d \)). In what follows \( c_1(d), c_2(d), \ldots, C_1(d), C_2(d), \ldots \) denote quantities depending on the dimension \( d \) only. Since \( d \) is fixed throughout our argument, we regard them as constants.

First we find conditions on \( \omega \) and \( \delta \) under which the set \( \Omega(\omega, \delta) \) contains a normalized basis \( \{ e_i \}_{i=1}^d \) whose distance to an orthonormal basis can be estimated in terms of \( d \) only.

**Lemma 5** There exist \( 0 < c_1(d), C_1(d), C_2(d) < \infty \), such that for \( \omega \leq \frac{1}{\delta d} \) and \( \delta \leq c_1(d)\omega^{d-1} \) there is a normalized basis \( \{ e_i \}_{i=1}^d \) in the space \( \mathbb{R}^d \) satisfying the conditions:

(a) \( \mu(\Delta(e_i, \omega)) \geq \delta \).

(b) If \( \{ e_i \}_{i=1}^d \) is an orthonormal basis in \( \mathbb{R}^d \), then the operator \( N : \mathbb{R}^d \to \mathbb{R}^d \) given by \( N_{\alpha} = e_i \) satisfies \( \|N\| \leq C_1(d) \) and \( \|N^{-1}\| \leq C_2(d) \), where the norms are the operator norms of \( N, N^{-1} \) considered as operators from \( \ell_2^d \) into \( \ell_2^d \).

**Proof.** We need an estimate for \( \mu(S) \). Observe that if \( K_1 \) and \( K_2 \) are two symmetric zonotopes and \( K_1 \subset K_2 \), then \( \mu_1(S) \leq \mu_2(S) \) for the corresponding measures \( \mu_1 \) and \( \mu_2 \) (defined as even measures satisfying \( (12) \) with \( Z = K_1 \) and \( Z = K_2 \), respectively). To prove this statement we integrate the equality \( (12) \) with respect to \( x \) over the Haar measure on \( S \).

Now we use the assumption that \( B_2^d \) is the maximal volume ellipsoid in \( Z \). Let \( \sum_{i=1}^n \gamma_i x_i \otimes x_i \) be the F. John representation of the identity operator corresponding to \( Z \) (see [12, p. 46]). Then
\[
Z \subset \{ x : |\langle x, x_i \rangle| \leq 1 \forall i \in \{1, \ldots, n\} \}.
\]

Since \( x = \sum_{i=1}^n \langle x, x_i \rangle \gamma_i x_i \) for each \( x \in \mathbb{R}^d \), we have \( Z \subset \sum_{i=1}^n [\gamma_i x_i, \gamma_i x_i] \). Since \( \sum_{i=1}^n \gamma_i = d \), this implies \( \mu(S) \leq d \).

Using the well-known computation, which goes back to B. Grünbaum ([8, p. 462, (5.2)]), see, also, [11, pp. 94–95]) one can find estimates for \( \mu(S) \) from below, which imply \( \mu(S) \geq \sqrt{d} \). For our purposes the trivial estimate \( \mu(S) \geq 1 \) is sufficient (this estimate follows immediately from \( Z \supset B_2^d \), because this inclusion implies \( h_Z(x) \geq \|x\| \)).

We denote the normalized Haar measure on \( S \) by \( \eta \). It is well known that there exists \( c_2(d) > 0 \) such that
\[
\eta(\Delta(x, r)) \geq c_2(d) r^{d-1} \forall r \in (0, 1) \forall x \in S.
\]

Using a standard averaging argument and \( \mu(S) \geq 1 \), we get that there exists \( e_1 \in S \) such that
\[
\mu(\Delta(e_1, \omega)) \geq c_2(d)\omega^{d-1}.
\]
Consider the closed \( \left( \frac{1}{3d} + \omega \right) \)-neighborhood (in the \( \ell_2^d \) metric) of the line \( L_1 \) spanned by \( e_1 \). Let \( \Delta_1 \) be the intersection of this neighborhood with \( S \). Our purpose is to estimate \( \mu(S \setminus \Delta_1) \) from below. Let \( x \in S \) be orthogonal to \( e_1 \). Then
\[
1 \leq h_Z(x) \leq 1 \cdot \mu(S \setminus \Delta_1) + \left( \frac{1}{3d} + \omega \right) \cdot d,
\]
where the left-hand side inequality follows from the fact that \( Z \) contains \( B_{2}^{\ell_2} \). Therefore \( \mu(S \setminus \Delta_1) \geq 1 - \left( \frac{1}{3d} + \omega \right) d \).

We erase all measure \( \mu \) contained in \( \Delta_1 \), use a standard averaging argument again, and find a vector \( e_2 \) such that
\[
\mu(\Delta(e_2, \omega) \setminus \Delta_1) \geq c_2(d)\omega^{d-1} \left( 1 - \left( \frac{1}{3d} + \omega \right) d \right).
\]
Since \( \mu(\Delta(e_2, \omega) \setminus \Delta_1) > 0 \), the vector \( e_2 \) is not in the \( \frac{1}{3d} \)-neighborhood of \( L_1 \).

Let \( \Delta_2 \) be the intersection of \( S \) with the closed \( \left( \frac{1}{3d} + \omega \right) \)-neighborhood of \( L_2 = \operatorname{lin}\{e_1, e_2\} \) (that is, \( L_2 \) is the linear span of \( \{e_1, e_2\} \)). Let \( x \in S \) be orthogonal to \( L_2 \). Then
\[
1 \leq h_Z(x) \leq 1 \cdot \mu(S \setminus \Delta_2) + \left( \frac{1}{3d} + \omega \right) \cdot d,
\]
where the left-hand side inequality follows from the fact that \( Z \) contains \( B_{2}^{\ell_2} \). Therefore \( \mu(S \setminus \Delta_2) \geq 1 - \left( \frac{1}{3d} + \omega \right) d \).

Using the standard averaging argument in the same way as in the previous step we find a vector \( e_3 \) such that
\[
\mu(\Delta(e_3, \omega) \setminus \Delta_2) \geq c_2(d)\omega^{d-1} \left( 1 - \left( \frac{1}{3d} + \omega \right) d \right).
\]
Since \( \mu(\Delta(e_3, \omega) \setminus \Delta_2) > 0 \), the vector \( e_3 \) is not in the \( \frac{1}{3d} \)-neighborhood of \( L_2 \).

We continue in an obvious way. As a result we construct a normalized basis \( \{e_1, \ldots, e_d\} \) satisfying the conditions

(i) \( \mu(\Delta(e_i, \omega)) \geq c_2(d)\omega^{d-1} \left( 1 - \left( \frac{1}{3d} + \omega \right) d \right) \).

(ii) \( \operatorname{dist}(e_i, \operatorname{lin}\{e_j\}_{j=1}^{i-1}) \geq \frac{1}{3d} \), \( i = 2, \ldots, d \), where \( \operatorname{dist}(\cdot, \cdot) \) denotes the distance from a vector to a subspace.
If $\omega < \frac{1}{6d}$, the inequality (i) implies
\[
\mu(\Delta(e_i, \omega)) \geq \frac{1}{2} c_2(d) \omega^{d-1},
\]
and we get the estimate (a) of Lemma 5 with $c_1(d) = c_2(d)/2$.

To estimate $||N||$ and $||N^{-1}||$, we let $\{o_i\}_{i=1}^d$ be the basis obtained from $\{e_i\}$ using the Gram–Schmidt orthonormalization process. Let $N : \mathbb{R}^d \to \mathbb{R}^d$ be defined by $No_i = e_i$.

The estimate $||N|| \leq C_1(d)$ with $C_1(d) = \sqrt{d}$ follows because the vectors $\{e_i\}_{i=1}^d$ are normalized and the vectors $\{o_i\}_{i=1}^d$ form an orthonormal set.

To estimate $||N^{-1}||$ we observe that the matrix of $N$ with respect to the basis $\{o_i\}$ is
\[
N = \begin{pmatrix}
N_{11} & N_{12} & \cdots & N_{1d} \\
0 & N_{22} & \cdots & N_{2d} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & N_{dd}
\end{pmatrix},
\]
and that the inequality (ii) implies $N_{ii} \geq \frac{1}{3d}$. We have
\[
T = \begin{pmatrix}
N_{11} & 0 & \cdots & 0 \\
0 & N_{22} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & N_{dd}
\end{pmatrix},
\]
and
\[
D = \begin{pmatrix}
0 & N_{12} & \cdots & N_{1d} \\
N_{11} & 0 & \cdots & N_{2d} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & N_{dd}
\end{pmatrix},
\]

and $U = \begin{pmatrix}
0 & N_{12} & \cdots & N_{1d-1} \\
N_{11} & 0 & \cdots & N_{2d-1} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & N_{d-1,d-1}
\end{pmatrix}$. Therefore
\[
N^{-1} = (I + U)^{-1} D^{-1} = (I - U + U^2 - \cdots + (-1)^{d-1} U^{d-1}) D^{-1},
\]
the identity $(I + U)^{-1} = (I - U + U^2 - \cdots + (-1)^{d-1} U^{d-1})$ follows from the obvious equality $U^d = 0$. The definition of $U$ and $N_{ii} \geq \frac{1}{3d}$ imply that columns of $U$ are vectors with Euclidean norm at most $3d$, hence $||U|| \leq 3d^{\frac{d}{2}}$. Therefore the identity (14) implies the following estimate for $||N^{-1}||$:
\[
||N^{-1}|| \leq \frac{||U||^{d-1}}{||U|| - 1} \cdot ||D^{-1}|| \leq \frac{3d^{d\frac{d}{2}} - 1}{3d^\frac{d}{2} - 1} \cdot 3d.
\]
Denoting the right-hand side of this inequality by $C_2(d)$ we get the desired estimate. □
Remark. We do not need sharp estimates for \( c_1(d), C_1(d), \) and \( C_2(d) \) because \( d \) is fixed in our argument, and the dependence on \( d \) of the parameters involved in our estimates is not essential for our proofs.

We use the following notation: for a set \( \Gamma \subset S \) and a real number \( r > 0 \) we denote the set \( \{ x \in S : \inf\{||x - y|| : y \in \Gamma\} \leq r \} \) by \( \Gamma_r \).

**Lemma 6** Let \( c_2(d) \) be the constant from (13), then \( \mu(S \setminus ((\Omega(\omega, \delta))_\omega)) \leq \frac{\delta}{c_2(d)\omega^{d-1}} \).

**Proof.** Assume the contrary, that is, \( \mu(S \setminus ((\Omega(\omega, \delta))_\omega)) > \frac{\delta}{c_2(d)\omega^{d-1}} \). Then, using a standard averaging argument as in Lemma 5, we find a point \( x \) such that
\[
\mu(\Delta(x, \omega) \setminus ((\Omega(\omega, \delta))_\omega)) \geq c_2(d)\omega^{d-1} \cdot \frac{\delta}{c_2(d)\omega^{d-1}} = \delta.
\]
By the definition of \( \Omega(\omega, \delta) \) this implies \( x \in \Omega(\omega, \delta) \). On the other hand, since the set \( \Delta(x, \omega) \setminus ((\Omega(\omega, \delta))_\omega) \) is non-empty, it follows that \( x \notin \Omega(\omega, \delta) \). We get a contradiction. \( \blacksquare \)

### 3.2 Notation and definitions used in the rest of the proof

For each \( Z \in Z_\epsilon \) we apply Lemma 5 with \( \omega = \omega(\epsilon) = \epsilon^{4k} \) and \( \delta = \delta(\epsilon) = \epsilon^{4dk} \), where \( 0 < k < 1 \) is a number satisfying the conditions
\[
k < \frac{1}{6 + 4d^2} \quad \text{and} \quad k < \frac{1}{2d + 4d^2},
\]
we choose and fix such number \( k \) for the rest of the proof. It is clear that there is \( \Xi_0 = \Xi_0(d, k) > 0 \) such that the conditions \( \omega(\epsilon) \leq \frac{1}{6d} \) and \( \delta(\epsilon) \leq c_1(d)\omega(\epsilon)^d \) are satisfied for all \( \epsilon \in (0, \Xi_0) \), where \( c_1(d) \) is the constant from Lemma 5. In the rest of the argument we consider \( \epsilon \in (0, \Xi_0) \) only. Let \( \{e_i\}^d_{i=1} \) be one of the bases satisfying the conditions of Lemma 5 with the described choice of \( \omega \) and \( \delta \). Now we change the system of coordinates in \( \mathbb{R}^d \supset Z \) the second time. The new system of coordinates is such that \( \{e_i\}^d_{i=1} \) is its unit vector basis. We shall modify the objects introduced so far (\( \Omega, \mu, \) etc.) and denote their versions corresponding to the new system of coordinates by \( \tilde{\Omega}, \tilde{\mu}, \) etc. All these objects depend on \( Z, \epsilon, \) and the choice of \( \{e_i\}^d_{i=1} \).

We denote by \( \tilde{S} \) the Euclidean unit sphere in the new system of coordinates. We denote by \( \mathcal{N} : S \to \tilde{S} \) the natural normalization mapping, that is, \( \mathcal{N}(z) = z/||z|| \), where \( ||z|| \) is the Euclidean norm of \( z \) with respect to the new system of coordinates. The estimates for \( ||N|| \) and \( ||N^{-1}|| \) from Lemma 5 imply that the Lipschitz constants of the mapping \( \mathcal{N} \) and its inverse \( \mathcal{N}^{-1} : \tilde{S} \to S \) can be estimated in terms of \( d \) only.

We introduce a measure \( \tilde{\mu} \) on \( \tilde{S} \) as an atomic measure supported on a finite set and such that \( \tilde{\mu}(\mathcal{N}(z)) = \mu(z)||z|| \) for each \( z \in S \), where \( ||z|| \) is the norm of \( z \) in the new
system of coordinates. Using the definition of the zonotope \( Z \) it is easy to check that the function
\[
\hat{h}_Z(x) = \int_\tilde{S} |\langle x, \tilde{z} \rangle| d\tilde{\mu}(\tilde{z}),
\]
where \( \langle \cdot, \cdot \rangle \) is the inner product in the new coordinate system, is the support function of \( Z \) in the new system of coordinates.

We define \( \tilde{\Omega} = \tilde{\Omega}(\omega, \delta) \) as \( N(\Omega(\omega, \delta)) \). It is clear that \( e_i \in \tilde{\Omega} \). Everywhere below we mean coordinates in the new system of coordinates (when we refer to \( || \cdot ||, \Delta, \) etc).

The observation that \( \mathcal{N} \) and \( \mathcal{N}^{-1} \) are Lipschitz, with Lipschitz constants estimated in terms of \( d \) only, implies the following statements:

- There exist \( C_3(d), C_4(d) < \infty \) such that
  \[
  \tilde{\mu}(\tilde{S} \setminus (\tilde{\Omega}(\omega, \delta)C_3(d)\omega(\epsilon))) \leq C_4(d) \frac{\delta}{\omega^{d-1}}
  \]  
  (we use Lemma 6).
- There exist \( c_3(d) > 0 \) and \( C_5(d) < \infty \) such that
  \[
  \tilde{\mu}(\Delta(x, C_5(d)\omega)) \geq c_3(d)\delta \quad \forall x \in \tilde{\Omega}(\omega, \delta)
  \]  
  (we use the definitions of \( \Omega(\omega, \delta) \) and \( \tilde{\Omega}(\omega, \delta) \)).
- There exists a constant \( C_6(d) \) depending on \( d \) only, such that
  \[
  \text{vol}(Z) \leq C_6(d).
  \]

Let \( \tilde{Q} \) be the transpose of the matrix whose columns are the coordinates of \( z_i \) in the new system of coordinates. We denote by \( \tilde{w}_i \) (\( i = 1, \ldots, M \)) the \( d \times d \) minors of \( \tilde{Q} \) ordered in the same way as the \( w_i \). The vector \( \{\tilde{w}_i\}_{i=1}^M \) is a scalar multiple of \( \{w_i\}_{i=1}^M \). Therefore (10) implies
\[
\max_i |u_i| \sum_{i=1}^M |\tilde{w}_i| = (1 + \varepsilon)^d \left| \sum_{i=1}^M u_i \tilde{w}_i \right|.
\]
The volume of \( Z \) in the new system of coordinates is \( 2^d \sum_{i=1}^M |\tilde{w}_i| \).

3.3 Lemma on six large minors

To show that if \( \varepsilon > 0 \) is small, then the inequality (19) implies that all but a very small part of the measure \( \tilde{\mu} \) is supported “around” multiples of vectors represented by a totally unimodular matrix in some basis, we need the following lemma. It shows that the inequality (19) implies that the measure \( \tilde{\mu} \) cannot have non-trivial “masses” near \( (d + 2) \)-tuples of vectors satisfying certain condition.

**Lemma 7** Let \( \chi(\varepsilon) \), \( \sigma(\varepsilon) \), and \( \pi(\varepsilon) \) be functions satisfying the following conditions:
\( (1) \lim_{\varepsilon \to 0} \chi(\varepsilon) = \lim_{\varepsilon \to 0} \sigma(\varepsilon) = \lim_{\varepsilon \to 0} \pi(\varepsilon) = 0; \)

\( (2) \varepsilon = o((\chi(\varepsilon))^2(\sigma(\varepsilon))^d) \) as \( \varepsilon \downarrow 0; \)

\( (3) \pi(\varepsilon) = o(\chi(\varepsilon)) \) as \( \varepsilon \downarrow 0; \)

\( (4) \) There is a subset \( \Phi_0 \subset (0, \Xi_0) \) such that the closure of \( \Phi_0 \) contains 0, and for each \( \varepsilon \in \Phi_0 \) there exist \( Z \in \mathbb{Z}_\varepsilon \) and points \( x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4 \) in the corresponding \( \hat{S} \), such that

\[
\mu(\Delta(z, \pi(\varepsilon))) \geq \sigma(\varepsilon) \quad \forall z \in \{x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4\}. \tag{20}
\]

Let \( U_0 \) be the set of pairs \((\varepsilon, Z)\) in which \( \varepsilon \in \Phi_0 \) and \( Z \) satisfies the condition from (4).

Let \( \Phi_1 \subset \Phi_0 \) be the set of those \( \varepsilon \in \Phi_0 \) for which there exists \((\varepsilon, Z) \in U_0 \) such that the corresponding points \( x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4 \) satisfy the condition

\[
|\det(H_{\alpha,\beta})| \geq \chi(\varepsilon) \tag{21}
\]

for all matrices \( H_{\alpha,\beta} \) whose columns are the coordinates of \( \{x_1, \ldots, x_{d-2}, p_\alpha, p_\beta\}, \alpha, \beta \in \{1, 2, 3, 4\}, \alpha \neq \beta \), with respect to an orthonormal basis \( \{e_i\}_{i=1}^d \) in \( \mathbb{R}^d \). Then there exists \( \Xi_1 > 0 \) such that \( \Phi_1 \cap (0, \Xi_1) = \emptyset \).

\textbf{Proof.} We assume the contrary, that is, we assume that 0 belongs to the closure of \( \Phi_1 \). For each \( \varepsilon \in \Phi_1 \) we choose \( Z \in \mathbb{Z}_\varepsilon \) such that \((\varepsilon, Z) \in U_0 \) and the condition (20) is satisfied. We show that for sufficiently small \( \varepsilon > 0 \) this leads to a contradiction.

We consider the following perturbation of the matrix \( H_{\alpha,\beta} \): each column vector \( z \) in it is replaced by a vector from \( \Delta(z, \pi(\varepsilon)) \). We denote the obtained perturbation of the matrix \( H_{\alpha,\beta} \) by \( H_{\alpha,\beta}^p \). We claim that

\[
|\det(H_{\alpha,\beta}^p)| \geq \chi(\varepsilon) - d \cdot \pi(\varepsilon). \tag{22}
\]

To prove this claim we need the following lemma, which we state in a bit more general form than is needed now, because we shall need it later.

\textbf{Lemma 8} Let \( x_1, \ldots, x_d, z \in \mathbb{L}_2^d \) be such that \( \max_{2 \leq i \leq d} \|x_i\| \leq m \) and \( \|z - x_1\| \leq l \). Then

\[
|\det[z, x_2, \ldots, x_d] - \det[x_1, x_2, \ldots, x_d]| \leq l \cdot m^{d-1}.
\]

This lemma follows immediately from the volumetric interpretation of determinants.

To get the inequality (22) we apply Lemma 8 \( d \) times with \( m = 1 \) and \( l = \pi(\varepsilon) \).

Since \( Z \in \mathbb{Z}_\varepsilon \), it can be represented in the form \( Z = \sum_i I(z_i) \). First we complete our proof in a special case when the following condition is satisfied:

\((*)\) All vectors \( z_i \) whose normalizations \( z_i/\|z_i\| \) belong to the sets \( \Delta(z, \pi(\varepsilon)), z \in \{x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4\} \), have the same norm \( \tau \) and there are equal amounts of such vectors in
each of the sets $\Delta(z, \pi(\varepsilon))$, $z \in \{x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4\}$, we denote the common value of the amounts by $F$.

The inequality (20) implies

$$F \cdot \tau \geq \sigma(\varepsilon)$$

We denote by $\Lambda$ the set of all numbers $i \in \{1, \ldots, M\}$ satisfying the condition: the normalizations of columns of the minor $\tilde{w}_i$ form a matrix of the form $H_{\alpha, \beta}^p$, for some $\alpha, \beta \in \{1, 2, 3, 4\}$.

We need an estimate for $\sum_{i \in \Lambda} |\tilde{w}_i|$. The inequality (22) implies $|\tilde{w}_i| \geq \tau^d(\chi(\varepsilon) - d \cdot \pi(\varepsilon))$ for each $i \in \Lambda$.

On the other hand, the cardinality $|\Lambda|$ of $\Lambda$ is $6F^d$. In fact, there are $F^{d-2}$ ways to choose $z_i/||z_i||$ in the sets $\Delta(x_j, \pi(\varepsilon))$, $j = 1, \ldots, d-2$. There are $\binom{4}{2} = 6$ ways to choose two of the sets $\Delta(p_j, \pi(\varepsilon))$, $j = 1, 2, 3, 4$, and there are $F^2$ ways to choose one vector $z_i/||z_i||$ in each of them. Therefore $|\Lambda| = 6F^d$ and

$$\sum_{i \in \Lambda} |\tilde{w}_i| \geq 6F^d \tau^d(\chi(\varepsilon) - d \cdot \pi(\varepsilon)) \geq 6(\sigma(\varepsilon))^d(\chi(\varepsilon) - d \cdot \pi(\varepsilon)). \quad (23)$$

We assume for simplicity that $\max_i |u_i| = 1$ (if it is not the case, some of the sums below should be multiplied by $\max_i |u_i|$). The $u_i$ are defined above the equality (10). Then the condition (19) can be rewritten as

$$(1 + \varepsilon)^d \left| \sum_{i=1}^M u_i \tilde{w}_i \right| \geq \sum_{i \in \Lambda} |\tilde{w}_i| + \sum_{i \notin \Lambda} |\tilde{w}_i|. \quad (24)$$

On the other hand,

$$(1 + \varepsilon)^d \left| \sum_{i=1}^M u_i \tilde{w}_i \right| \leq (1 + \varepsilon)^d \left| \sum_{i \in \Lambda} u_i \tilde{w}_i \right| + (1 + \varepsilon)^d \sum_{i \notin \Lambda} |\tilde{w}_i|. \quad (25)$$

From (24) and (25) we get

$$(1 + \varepsilon)^d \left| \sum_{i \in \Lambda} u_i \tilde{w}_i \right| \geq \sum_{i \in \Lambda} |\tilde{w}_i| - ((1 + \varepsilon)^d - 1) \sum_{i \notin \Lambda} |\tilde{w}_i|. \quad (26)$$

As is well known, $2^d \sum_{i=1}^M |\tilde{w}_i|$ is the volume of $Z$, hence $\sum_{i=1}^M |\tilde{w}_i| \leq 2^{-d}C_6(d)$. Using this observation and the inequalities (23) and (26) we get

$$\left| \sum_{i \in \Lambda} u_i \tilde{w}_i \right| \geq \left( \frac{1}{(1 + \varepsilon)^d} - \frac{((1 + \varepsilon)^d - 1)C_6(d)2^{-d}}{6(\sigma(\varepsilon))^d(\chi(\varepsilon) - d \cdot \pi(\varepsilon))} \right) \sum_{i \in \Lambda} |\tilde{w}_i|. \quad (27)$$
(We use the fact that \( \chi(\varepsilon) - d \cdot \pi(\varepsilon) > 0 \) if \( \varepsilon > 0 \) is small enough.) The conditions (2) and (3) imply that there exists \( \psi > 0 \) such that

\[
\left( \frac{1}{1 + \varepsilon} - \frac{(1 + \varepsilon)^d - 1}{6(\sigma(\varepsilon))^d(\chi(\varepsilon) - d \cdot \pi(\varepsilon))} \right) > (1 - 0.04(\chi(\varepsilon) - d \cdot \pi(\varepsilon)))
\]

is satisfied if \( \varepsilon \in (0, \psi) \). The right-hand side is chosen in the form needed below.

Let \( \psi > 0 \) be such that the statement above is true. Then for \( \varepsilon \in (0, \psi) \) we have

\[
\left| \sum_{i \in \Lambda} u_i \tilde{w}_i \right| \geq (1 - 0.04(\chi(\varepsilon) - d \cdot \pi(\varepsilon))) \sum_{i \in \Lambda} |\tilde{w}_i|.
\]

(28)

Recall that \( u_i \) are \( d \times d \) minors of some matrix \([y_1, \ldots, y_d]\). We need the Plücker relations, see [9, p. 312] or [30, p. 42]. The result that we need can be stated in the following way: if \( \gamma_1, \ldots, \gamma_{d-2}, \kappa_1, \kappa_2, \kappa_3, \kappa_4 \) are indices of \( d + 2 \) rows of \([y_1, \ldots, y_d]\), then

\[
t_{1,2}t_{3,4} - t_{1,4}t_{3,2} + t_{2,4}t_{3,1} = 0,
\]

where \( t_{\alpha,\beta} \) is the determinant of the \( d \times d \) matrix whose rows are the rows of \([y_1, \ldots, y_d]\) with the indices \( \gamma_1, \ldots, \gamma_{d-2}, \kappa_\alpha, \) and \( \kappa_\beta \). Note that (29) can be verified by a straightforward computation (which is very simple if we make a suitable change of coordinates before the computation).

Now we show that (28) cannot be satisfied. Let \( \Psi \) be a set consisting of \( (d + 2) \) vectors \( z_{\kappa_1}, z_{\kappa_2}, z_{\kappa_3}, z_{\kappa_4}, z_{\gamma_1}, \ldots, z_{\gamma_{d-2}} \), formed in the following way. We choose vectors \( (z_{\kappa_i}/||z_{\kappa_i}||) \in \Delta(p_i, \pi(\varepsilon)), i = 1, 2, 3, 4, \) and choose vectors \( (z_{\gamma_i}/||z_{\gamma_i}||) \in \Delta(x_i, \pi(\varepsilon)), i = 1, \ldots, d-2. \) To each such selection there corresponds a set of 6 minors \( \tilde{w}_i \) of the form \( t_{\alpha,\beta} \det(H_{p_i}^{\alpha \beta}) \), we denote this set of six minors by \( \{\tilde{w}_i\}_{i \in M(\Psi)} \).

One of the immediate consequences of the Plücker relation (29) is that for any such \((d + 2)\)-tuple \( \Psi \)

\[
|u_i| \leq \frac{1}{\sqrt{2}} \quad \text{for some} \quad i \in M(\Psi).
\]

(30)

(Here we use the assumption that \( \max_i |u_i| = 1 \).

For each \( \Psi \) we choose one such \( i \in M(\Psi) \) and denote it by \( s(\Psi) \). The estimate (22) and the condition (*) imply that

\[
\tau^d \geq |\tilde{w}_i| \geq \tau^d(\chi(\varepsilon) - d \cdot \pi(\varepsilon))
\]

(31)

for every \( i \in \Lambda \).
Hence for every \((d + 2)\)-tuple \(\Psi\) of the described type we have

\[
\left| \sum_{i \in M(\Psi)} u_i \tilde{w}_i \right| \leq \sum_{i \in M(\Psi) \setminus \{s(\Psi)\}} \left| \tilde{w}_i \right| + \frac{1}{\sqrt{2}} \left| \tilde{w}_{s(\Psi)} \right| \leq \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right| - \frac{\sqrt{2} - 1}{\sqrt{2}} \left| \tilde{w}_{s(\Psi)} \right|
\]

\[
= \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right| \left( 1 - \frac{\sqrt{2} - 1}{\sqrt{2} \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right|} \right)
\leq \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right| \left( 1 - \frac{\sqrt{2} - 1}{\sqrt{2}} \tau^d (\chi(\varepsilon) - d \cdot \pi(\varepsilon)) \right)
\leq \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right| (1 - 0.04(\chi(\varepsilon) - d \cdot \pi(\varepsilon))).
\]

Thus

\[
\left| \sum_{i \in M(\Psi)} u_i \tilde{w}_i \right| < \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right| (1 - 0.04(\chi(\varepsilon) - d \cdot \pi(\varepsilon))).
\]

(32)

Recall that \(F\) is the number of vectors \(z_i\) corresponding to each of the sets \(\Delta(z, \pi(\varepsilon))\), \(z \in \{x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4\}\). Simple counting shows that for an arbitrary collection \(\{\Upsilon_i\}_{i \in \Lambda}\) of numbers we have

\[
\sum_{\Psi} \sum_{i \in M(\Psi)} \Upsilon_i = F^2 \sum_{i \in \Lambda} \Upsilon_i.
\]

Using (32) we get that

\[
F^2 \sum_{i \in \Lambda} \left| u_i \tilde{w}_i \right| = \sum_{\Psi} \sum_{i \in M(\Psi)} \left| u_i \tilde{w}_i \right| \leq \sum_{\Psi} \sum_{i \in M(\Psi)} u_i \tilde{w}_i \leq \sum_{i \in M(\Psi)} \left| \tilde{w}_i \right| (1 - 0.04(\chi(\varepsilon) - d \cdot \pi(\varepsilon)))
\]

\[
= F^2 \sum_{i \in \Lambda} \left| \tilde{w}_i \right| (1 - 0.04(\chi(\varepsilon) - d \cdot \pi(\varepsilon))).
\]

If \(\varepsilon \in (0, \psi)\), we get a contradiction with (28).

To see that the general case can be reduced to the case (*) we need the following observation:

Let \(\tau_1, \tau_2 > 0\) be such that \(\tau_1 + \tau_2 = 1\). We replace the row with the coordinates of \(z_j\) in \(\tilde{Q}\) by two rows, one of them is the row of coordinates of \(\tau_1 z_j\) and the other is the row of coordinates of \(\tau_2 z_j\). The zonotope generated by the rows of the obtained matrix coincides with \(Z\). In the matrix \([y_1, \ldots, y_d]\) we replace the \(j^{\text{th}}\) row by two copies of it. It is easy to see that if we replace the sequences \(\{u_i\}_{i=1}^M\) and \(\{\tilde{w}_i\}_{i=1}^M\) by sequences of \(d \times d\) minors of these new matrices, the condition (19) is still satisfied.
We can repeat this ‘cutting’ of vectors $z_j$ into ‘pieces’ with (19) still being valid.

Therefore, we may assume the following: among $z_j$ corresponding to each of the sets $\Delta(z, \pi(\varepsilon))$, $z \in \{x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4\}$ there exists a subset $\Phi(z, \pi(\varepsilon))$ consisting of vectors having the same length $\tau$, and such that the sum of norms of vectors from $\Phi(z, \pi(\varepsilon))$ is $\geq \sigma(\varepsilon)$, moreover, we may assume that the numbers of such vectors in the subsets $\Phi(z, \pi(\varepsilon))$ are the same for all $z \in \{x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4\}$.

Lemma 7 in this case can be proved using the same argument as before, but with $\Lambda$ being the set of those minors $\tilde{w}_i$ for which rows are from $\Phi(z, \pi(\varepsilon))$. Everything starting with the inequality (23) can be shown in the same way as before; only some constants will be changed (because we need to replace $\sigma(\varepsilon)$ by $\frac{\sigma(\varepsilon)}{2}$).

### 3.4 Searching for a totally unimodular matrix

Let $\rho(\varepsilon) = \varepsilon^k$, $\nu(\varepsilon) = \varepsilon^{3k}$. For a vector $s$ we denote its coordinates with respect to $\{e_i\}_{i=1}^d$ by $\{s_i\}_{i=1}^d$. (Here $k$ and $\{e_i\}_{i=1}^d$ are the same as in Section 3.2.)

**Lemma 9** If

$$k < \frac{1}{6 + 4d^2}, \quad (33)$$

then there exists $\Xi_2 > 0$ such that for $\varepsilon \in (0, \Xi_2)$, $s, t \in \hat{\Omega}(\omega(\varepsilon), \delta(\varepsilon))$, and $\alpha, \beta \in \{1, \ldots, d\}$, the inequality

$$\min\{|s_\alpha|, |s_\beta|, |t_\alpha|, |t_\beta|\} \geq \rho(\varepsilon), \quad (34)$$

implies

$$\left| \det \begin{pmatrix} s_\alpha & t_\alpha \\ s_\beta & t_\beta \end{pmatrix} \right| < \nu(\varepsilon). \quad (35)$$

**Proof.** Assume the contrary, that is, there exists a subset $\Phi_2 \subset (0, 1)$, having 0 in its closure and such that for each $\varepsilon \in \Phi_2$ there exist $Z \in Z$, $s, t \in \Omega(\omega(\varepsilon), \delta(\varepsilon))$ and $\alpha, \beta$ satisfying the condition (34), and such that

$$\left| \det \begin{pmatrix} s_\alpha & t_\alpha \\ s_\beta & t_\beta \end{pmatrix} \right| \geq \nu(\varepsilon).$$

We apply Lemma 7 with $\{x_1, \ldots, x_{d-2}\} = \{e_i\}_{i \neq \alpha, \beta}$, $\{p_1, p_2, p_3, p_4\} = \{e_\alpha, e_\beta, s, t\}$. Using a straightforward determinant computation we see that the condition (21) is satisfied with $\chi(\varepsilon) = \min\{1, \rho(\varepsilon), \nu(\varepsilon)\} = \varepsilon^{3k}$ (we consider $\varepsilon < 1$).

The inequality (17) implies that the condition (4) of Lemma 7 is satisfied with $\pi(\varepsilon) = C_5(d)\omega(\varepsilon) = C_5(d)\varepsilon^{4k}$ and $\sigma(\varepsilon) = c_3(d)\delta(\varepsilon) = c_3(d)\varepsilon^{4dk}$. It is clear that the conditions (2) and (3) of Lemma 7 are satisfied. To get (2) we use the condition (33). Applying Lemma 7, we get the existence of the desired $\Xi_2$. □
For each vector from \( \hat{\Omega}(\omega(\varepsilon), \delta(\varepsilon)) \) we define its top set as the set of indices of coordinates whose absolute values \( \geq \rho(\varepsilon) \).

The collection of all possible top sets is a subset of the set of all subsets of \( \{1, \ldots, d\} \), hence its cardinality is at most \( 2^d \). We create a collection \( \Theta(\omega(\varepsilon), \delta(\varepsilon)) \subset \hat{\Omega}(\omega(\varepsilon), \delta(\varepsilon)) \) in the following way: for each subset of \( \{1, \ldots, d\} \) which is a top set for at least one vector from \( \hat{\Omega}(\omega(\varepsilon), \delta(\varepsilon)) \), we choose one of such vectors; the set \( \Theta(\omega(\varepsilon), \delta(\varepsilon)) \) is the set of all vectors selected in this way.

In our next lemma we show that each vector from \( \hat{\Omega}(\omega(\varepsilon), \delta(\varepsilon)) \) can be reasonably well approximated by a vector from \( \Theta(\omega(\varepsilon), \delta(\varepsilon)) \). Therefore (as we shall see later), to prove Lemma 2 it is sufficient to find a “totally unimodular” set approximating \( \Theta(\omega(\varepsilon), \delta(\varepsilon)) \).

**Lemma 10** Let \( \rho(\varepsilon) \) and \( \nu(\varepsilon) \) be as above and let \( k \) and \( \Xi_2 \) be numbers satisfying the conditions of Lemma 9. Let \( \varepsilon \in (0, \Xi_2) \), \( Z \in \mathbb{Z}_\varepsilon \), and let \( s, t \in \hat{\Omega}(\omega(\varepsilon), \delta(\varepsilon)) \) be two vectors with the same top set \( \Sigma \). Then

\[
\min\{||t + s||, |t - s||\} \leq \sqrt{2 \nu(\varepsilon)\rho(\varepsilon)^2} + 4d\rho(\varepsilon)^2.
\]

(36)

**Proof.** Observe that if \( \rho(\varepsilon) = \varepsilon^k > \frac{1}{\sqrt{d}} \), the statement of the lemma is trivial. Therefore we may assume that \( \rho(\varepsilon) \leq \frac{1}{\sqrt{d}} \). In such a case \( \Sigma \) contains at least one element.

First we show that the signs of different components of \( s \) and \( t \) “agree” on \( \Sigma \) in the sense that either they are the same everywhere on \( \Sigma \), or they are the opposite everywhere on \( \Sigma \). In fact, assume the contrary, and let \( \alpha, \beta \in \Sigma \) be indices for which the signs “disagree”. Then, as is easy to check,

\[
\left| \det \begin{pmatrix} s_\alpha & t_\alpha \\ s_\beta & t_\beta \end{pmatrix} \right| = |s_\alpha||t_\beta| - |s_\beta||t_\alpha| \geq 2(\rho(\varepsilon))^2 > \nu(\varepsilon),
\]

and we get a contradiction. We consider the case when the signs of \( t_\alpha \) and \( s_\alpha \) are the same for each \( \alpha \in \Sigma \), the other case can be treated similarly (we can just consider \(-s\) instead of \( s \)).

We may assume without loss of generality that \( |t_\alpha| \geq |s_\alpha| \) for some \( \alpha \in \Sigma \). We show that in this case

\[
|t_\beta| \geq \left(1 - \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2}\right) |s_\beta|
\]

for all \( \beta \in \Sigma \). In fact, if \( |t_\beta| < \left(1 - \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2}\right) |s_\beta| \) for some \( \beta \in \Sigma \), then

\[
\nu(\varepsilon) > \left| \det \begin{pmatrix} s_\alpha & t_\alpha \\ s_\beta & t_\beta \end{pmatrix} \right| \geq |t_\alpha||s_\beta| - |s_\alpha||t_\beta| \geq |s_\alpha||s_\beta| \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2} \geq \nu(\varepsilon),
\]

21
a contradiction.

We have
\[
||t - s||^2 = ||t||^2 + ||s||^2 - 2 \langle t, s \rangle \leq 2 - 2 \sum_{\alpha \in \Sigma} \left( 1 - \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2} \right) s_\alpha^2 + 2 \sum_{\alpha \notin \Sigma} \rho(\varepsilon)^2 
\]
\[
\leq 2 \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2} + 4 \sum_{\alpha \notin \Sigma} \rho(\varepsilon)^2 \leq 2 \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2} + 4d\rho(\varepsilon)^2.
\]

Q.E.D. ■

Let \( \Theta(\omega(\varepsilon), \delta(\varepsilon)) = \{ b_j \}_{j=1}^d \), where \( J \leq 2^d \). We may and shall assume that \( \{ e_i(\varepsilon) \}_{i=1}^d \subset \Theta(\omega(\varepsilon), \delta(\varepsilon)) \) (see Lemma 5 and Section 3.2). We denote \( d \cdot 2^d \) by \( n \) and introduce \( d \cdot n \) functions: \( \varphi_1(\varepsilon), \ldots, \varphi_{d \cdot n}(\varepsilon) \), such that
\[
\varphi_1(\varepsilon) \geq \cdots \geq \varphi_{d \cdot n}(\varepsilon) = \rho(\varepsilon) = \varepsilon^k.
\]

(37)

(38)

We consider the matrix \( X \) whose columns are \( \{ b_j \}_{j=1}^d \). We order the absolute values of entries of this matrix in non-increasing order and denote them by \( a_1 \geq a_2 \geq \cdots \geq a_{d \cdot j} \). Let \( j_0 \) be the least index for which
\[
\varphi_{d \cdot j_0}(\varepsilon) > a_{j_0}.
\]

(39)

The existence of \( j_0 \) follows from \( \{ e_i(\varepsilon) \}_{i=1}^d \subset \Theta(\omega(\varepsilon), \delta(\varepsilon)) \). The definition of \( j_0 \) implies that \( a_j \geq \varphi_{d \cdot j}(\varepsilon) \) for \( j < j_0 \), hence \( a_j \geq \varphi_{d \cdot (j_0-1)}(\varepsilon) \) for \( j \leq j_0 - 1 \).

We replace all entries of the matrix \( X \) except \( a_1, \ldots, a_{j_0-1} \) by zeros and denote the obtained matrix by \( G = (G_{ij}), i = 1, \ldots, d, j = 1, \ldots, J \), and its columns by \( \{ g_j \}_{j=1}^d \). It is clear that
\[
||g_j - b_j|| \leq d \cdot \varphi_{d \cdot j_0}(\varepsilon).
\]

(40)

We form a bipartite graph \( G \) on the vertex set \( \{ 1, \ldots, d \} \cup \{ 1, \ldots, J \} \), where we use bars in \( 1, \ldots, d \) because these vertices are considered as different from the vertices \( 1, \ldots, d \), which are in the set \( \{ 1, \ldots, J \} \). The edges of \( G \) are defined in the following way: the vertices \( i \) and \( j \) are adjacent if and only if \( G_{ij} \neq 0 \). So there is a one-to-one correspondence between edges of \( G \) and non-zero entries of \( G \). We choose and fix a maximal forest \( F \) in \( G \). (We use the standard terminology, see, e. g. [29, p. 11].)

For each non-zero entry of \( G \) we define its level in the following way:

The level of entries corresponding to edges of \( F \) is 1.

For a non-zero entry of \( G \) which does not correspond to an edge in \( F \) we consider the cycle in \( G \) formed by the corresponding edge and edges of \( F \). We define the level of the entry as the half of the length of the cycle (recall that the graph \( G \) is bipartite, hence all cycles are even).
Observation. One of the classes of the bipartition has \( d \) vertices. Hence no cycle can have more than \( 2d \) edges, and the level of each vertex is at most \( d \).

To each entry \( G_{ij} \) of level \( f \) we assign a square submatrix \( G(ij) \) of \( G \) all other entries in which are of levels at most \( f-1 \). We do this in the following way. To entries corresponding to edges of \( F \) we assign the \( 1 \times 1 \) matrices containing these entries. For an entry \( G_{ij} \) which does not correspond to an edge in \( F \) we consider the corresponding edge \( e \) in \( G \) and the cycle \( C \) formed by \( e \) and edges of \( F \). Then we consider the entries in \( G \) corresponding to edges of \( C \) and the minimal submatrix in \( G \) containing all of these entries. Now we consider all edges in \( G \) corresponding to non-zero entries of this submatrix. We choose and fix in this set of edges a minimum-length cycle \( M \) containing \( e \). We define \( G(ij) \) as the minimal submatrix of \( G \) containing all entries corresponding to edges of \( M \). It is easy to verify that:

- \( G(ij) \) is a square submatrix of \( G \).
- Non-zero entries of \( G(ij) \) are in one-to-one correspondence with entries of \( M \).
- The expansion of the determinant of \( G(ij) \) according to the definition contains exactly two non-zero terms.
- All non-zero entries of \( G(ij) \) except \( G_{ij} \) have level \( \leq f-1 \).

Lemma 11 Let \( k < 1/(2d + 4d^2) \). If \( \varepsilon > 0 \) is small enough, then there exists a \( d \times J \) matrix \( \tilde{G} \) such that:

1. If some entry of \( G \) is zero, the corresponding entry of \( \tilde{G} \) is also zero.
2. The entries of level 1 of \( \tilde{G} \) are the same as for \( G \);
3. All other non-zero entries of \( \tilde{G} \) are perturbations of entries of \( G \) satisfying the following conditions:
   a. If \( G_{ij} \) is of level \( f \), then \( |G_{ij} - \tilde{G}_{ij}| < \varphi_{d,j_0-f+1}(\varepsilon) \).
   b. For each non-zero entry \( G_{ij} \) of level \( \geq 2 \) of \( G \) the determinant of the submatrix \( \tilde{G}(ij) \) of \( \tilde{G} \) corresponding to \( G(ij) \) is zero.

Proof. Let \( G_{ij} \) be an entry of level \( f \). Since, as it was observed above, all entries of \( G(ij) \) have level \( \leq f-1 \), we can prove the lemma by induction as follows.

1. We let \( \tilde{G}_{ij} = G_{ij} \) for all \( G_{ij} \) of level one.
2. Let \( f \geq 2 \). Induction hypothesis: We assume that for all entries \( G_{ij} \) of levels \( \ell(G_{ij}) \) satisfying \( 2 \leq \ell(G_{ij}) \leq f-1 \) we have found perturbations \( \tilde{G}_{ij} \) satisfying

   \[ |G_{ij} - \tilde{G}_{ij}| \leq \varphi_{d,j_0-\ell(G_{ij})+1}(\varepsilon), \]

   such that \( \det(\tilde{G}(ij)) = 0 \). (Note that this assumption is vacuous if \( f = 2 \).)

   Inductive step: Let \( G_{ij} \) be an entry of level \( f \). If \( \varepsilon > 0 \) is small enough we can find a number \( \tilde{G}_{ij} \) such that \( |\tilde{G}_{ij} - G_{ij}| \leq \varphi_{d,j_0-f+1}(\varepsilon) \) and \( \det(\tilde{G}(ij)) = 0 \). Observe that by

23
the induction hypothesis and the observation that all other entries of \(G(ij)\) have levels \(\leq f - 1\), all other entries of \(\tilde{G}(ij)\) have already been defined.

So let \(G_{ij}\) be an entry of level \(f\), and \(G(ij)\) be the corresponding square submatrix. Renumbering rows and columns of the matrix \(G\) we may assume that the matrix \(G(ij)\) looks like the one sketched below for some \(h \leq f\).

\[
\begin{pmatrix}
a_1 & 0 & \ldots & 0 & G_{ij} & 0 & 0 & \ldots & 0 & 1 & 0 & \ldots \\
b_1 & a_2 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 1 & \ldots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & a_{h-1} & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots \\
0 & 0 & \ldots & b_{h-1} & a_h & 0 & \ldots & 0 & 0 & 0 & \ldots \\
\ast & \ast & \ldots & \ast & \ast & 1 & 0 & \ldots & 0 & 0 & 0 & \ldots \\
\ast & \ast & \ldots & \ast & \ast & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\ast & \ast & \ldots & \ast & \ast & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots \\
\ast & \ast & \ldots & \ast & \ast & 0 & 0 & \ldots & 0 & 1 & 0 & \ldots \\
\end{pmatrix}
\]

(41)

We have assumed that we have already found entries \(\{\tilde{a}_n\}_{n=1}^h\) and \(\{\tilde{b}_n\}_{n=1}^{h-1}\) of \(\tilde{G}\) which are perturbations of \(\{a_n\}_{n=1}^h\) and \(\{b_n\}_{n=1}^{h-1}\). The entries 1 shown above are the only non-zero entries in their columns, therefore the corresponding edges of \(G\) should be in \(F\). Let us denote the perturbation of \(G_{ij}\) we are looking for by \(\tilde{G}_{ij}\). The condition (b) of Lemma 11 can be written as

\[
\prod_{n=1}^h \tilde{a}_n + (-1)^{h-1} \prod_{n=1}^{h-1} \tilde{b}_n \cdot \tilde{G}_{ij} = 0
\]

(42)

So it suffices to show that the number \(\tilde{G}_{ij}\), found as a solution of (42) satisfies \(|\tilde{G}_{ij} - G_{ij}| < \varphi_{d,j_0-f+1}(\varepsilon)\). To show this we assume the contrary. Since there are finitely many possibilities for \(j_0\) and \(f\), the converse can be described as existence of \(j_0\) and \(f\), such that there is a subset \(\Phi_3 \subset (0,1)\), whose closure contains 0, satisfying the condition:

For each \(\varepsilon \in \Phi_3\) there is \(Z \in \mathcal{Z}_\varepsilon\) such that after proceeding with all steps of the construction we get: all the conditions above are satisfied, but

\[
\prod_{n=1}^h \tilde{a}_n + (-1)^{h-1} \prod_{n=1}^{h-1} \tilde{b}_n \cdot G_{ij} > \varphi_{d,j_0-f+1}(\varepsilon) \prod_{n=1}^{h-1} |\tilde{b}_n|.
\]

(43)

We need to get from here an estimate for \(|\det(G(ij))|\) from below. To get it we observe that the inequality (43) is an estimate from below of the determinant of the matrix
\[ G'(ij) = \begin{pmatrix} \tilde{a}_1 & 0 & \ldots & 0 & G_{ij} \\ \tilde{b}_1 & \tilde{a}_2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \tilde{a}_{h-1} & 0 \\ 0 & 0 & \ldots & \tilde{b}_{h-1} & \tilde{a}_h \end{pmatrix}. \]

To get from here an estimate for \( \det(G(ij)) \) from below we observe the following: The \( \ell_2 \)-norm of each column of \( G_{ij} \) is \( \leq 1 \), the \( \ell_2 \)-distance between a column of \( G_{ij} \) and the corresponding column of \( G'(ij) \) is at most \( 2\varphi_{d_0-f+2}(\varepsilon) \). Hence the \( \ell_2 \)-norm of each column of \( G'(ij) \) is \( \leq 1 + 2\varphi_{d_0-f+2}(\varepsilon) \). Applying Lemma 8 \( h \) times we get

\[ |\det(G(ij))| \geq |\det(G'(ij))| - h \cdot 2\varphi_{d_0-f+2}(\varepsilon)(1 + 2\varphi_{d_0-f+2}(\varepsilon))^{h-1}. \]

The induction hypothesis implies

\[ |\tilde{b}_1| \geq \varphi_{d(j_0-1)}(\varepsilon) - \varphi_{d_0-f+2}(\varepsilon), \]
we get

\[ |\det(G(ij))| \geq \varphi_{d_0-f+1}(\varepsilon) \cdot (\varphi_{d(j_0-1)}(\varepsilon) - \varphi_{d_0-f+2}(\varepsilon))^{h-1} - h \cdot 2\varphi_{d_0-f+2}(\varepsilon)(1 + 2\varphi_{d_0-f+2}(\varepsilon))^{h-1}. \] (44)

Let us keep the notation \( \{g_j\}_{j=1}^J \) for columns of the matrix (41). We consider the following six \( d \times d \) minors of this matrix: the corresponding submatrices contain the columns \( \{g_2, \ldots, g_{h-1}, g_{h+1}, \ldots, g_d\} \), and two out of the four columns \( \{g_1, g_h, g_{d+1}, g_{d+2}\} \). Observe that \( g_{h+1} = e_{h+1}, \ldots, g_d = e_d, g_{d+1} = e_1, g_{d+2} = e_2 \).

The absolute values of the minors are equal to

\[ |\det G(ij)|, \prod_{n=2}^h a_n, \prod_{n=1}^{h-1} b_n, |a_1| \cdot \prod_{n=2}^{h-1} b_n, \prod_{n=2}^h b_n, \prod_{n=2}^{h-1} b_n. \] (45)

The first number in (45) was estimated in (44). All other numbers are at least \( (\varphi_{d(j_0-1)}(\varepsilon))^{h-1} \), it is clear that this number exceeds the number from (44).

We are going to use Lemma 7 with \( \{x_1, \ldots, x_{d-2}\} = \{\mathcal{N}(g_2), \ldots, \mathcal{N}(g_{h-1}), \mathcal{N}(g_{h+1}), \ldots, \mathcal{N}(g_d)\} \) and \( \{p_1, p_2, p_3, p_4\} = \{\mathcal{N}(g_1), \mathcal{N}(g_h), \mathcal{N}(g_{d+1}), \mathcal{N}(g_{d+2})\} \). (Recall that \( \mathcal{N}(z) = z/||z|| \).) Our definitions imply that \( ||b_j|| = 1 \) and \( ||g_j|| \leq 1 \), because \( g_j \) is obtained from \( \mathcal{b}_j \) by replacing some of the coordinates by zeros. Hence the inequality (44) and the remark above on the numbers (45) imply that the condition (21) is satisfied with

\[ \chi(\varepsilon) = \varphi_{d_0-f+1}(\varepsilon) \cdot (\varphi_{d(j_0-1)}(\varepsilon) - \varphi_{d_0-f+2}(\varepsilon))^{h-1} - h \cdot 2\varphi_{d_0-f+2}(\varepsilon)(1 + 2\varphi_{d_0-f+2}(\varepsilon))^{h-1}. \] (46)

The inequality (40), the inclusion \( \mathcal{b}_j \in \Omega(\omega(\varepsilon), \delta(\varepsilon)) \) and (17) imply that the condition (20) is satisfied with \( \pi(\varepsilon) = 2d \cdot \varphi_{d_0}(\varepsilon) + C_5(d)\omega(\varepsilon) \) and \( \sigma(\varepsilon) = c_3(d)\delta(\varepsilon) \). So it remains
to show that the condition (38) implies that the conditions (2) and (3) of Lemma 7 are satisfied.

By (38), (46), the inequality \(2 \leq h \leq f \leq d\), and the trivial observation that all functions \(\varphi_\alpha(\varepsilon)\) do not exceed 1 for \(0 \leq \varepsilon \leq 1\), we have

\[
(\varphi_{d_0-f+1}(\varepsilon))^d = O(\chi(\varepsilon)).
\]  

(47)

Now we verify the condition (3) of Lemma 7. The part (b) can be verified as follows. The conditions (37) and (38), together with \(f \geq 2\) and \(\omega(\varepsilon) = \varepsilon^{4k}\), imply that \(\pi(\varepsilon) = O(\varphi_{d_0}(\varepsilon)) = o((\varphi_{d_0-f+1}(\varepsilon))^d) = o(\chi(\varepsilon))\).

To verify the condition (2) of Lemma 7 it suffices to observe that (47) and (37) imply \(\rho(\varepsilon))^d = O(\chi(\varepsilon))\). Hence (2) is satisfied if \(2dk + 4d^2k < 1\). This inequality is among the conditions of Lemma 11. Hence we can apply Lemma 7 and get the conclusion of Lemma 11.

Now let \(\tilde{G}\) be an approximation of \(G\) by a matrix satisfying the conditions of Lemma 11. We use the same maximal forest \(F\) in \(G\) as above. It is easy to show (and the corresponding result is well known in the theory of matroids, see, for example, [24, Theorem 6.4.7]) that multiplying columns and rows of \(\tilde{G}\) by positive numbers we can make entries corresponding to edges of \(F\) to be equal to \(\pm 1\). Denote the obtained matrix by \(\hat{G}\).

Lemma 12 If \(\tilde{G}\) satisfies the conditions of Lemma 11, then \(\hat{G}\) is a matrix with entries \(-1, 0, \text{ and } 1\).

Proof. Assume the contrary, that is, there are entries \(\hat{G}_{ij}\) which are not in the set \((-1, 0, 1\}. \) Let \(\hat{G}_{ij}\) be one of such entries satisfying the additional condition: the level \(\ell(G_{ij})\) is the minimal possible among all entries \(\hat{G}_{ij}\) which are not in \((-1, 0, 1\}. \) Denote by \(\hat{G}(ij)\) the submatrix of \(\hat{G}\) which corresponds to \(G(ij)\).

Then, by observations preceding Lemma 11, the expansion of \(\det \hat{G}(ij)\) contains two non-zero terms: one of them is 1 or \(-1\), the other is \(\hat{G}_{ij}\) or \(-\hat{G}_{ij}\). Our assumptions imply that \(\det \hat{G}(ij) \neq 0\). This contradicts \(\det \hat{G}(ij) = 0\), because \(\hat{G}\) is obtained from \(\hat{G}\) using multiplications of columns and rows by numbers.

In Lemma 13 we show that for functions \(\varphi_\alpha(\varepsilon)\) chosen as above, the matrix \(\hat{G}\) should be totally unimodular for sufficiently small \(\varepsilon\). In Lemma 15 we show how to estimate the Banach–Mazur distance between \(Z\) and \(T_d\) in the case when \(\hat{G}\) is totally unimodular.

Lemma 13 If \(\varepsilon > 0\) is small enough, the matrix \(\hat{G}\) is totally unimodular.

Proof. The conclusion of Lemma 11 implies that each entry of \(\hat{G}\) is a \(\varphi_{d(j_0-1)+1}(\varepsilon)\)-approximation of an entry from \(G\). Therefore for small \(\varepsilon\) the absolute value of each non-zero entry of \(\hat{G}\) is at least \(\varphi_{d(j_0-1)}(\varepsilon)/2\). This implies the following observation.

Observation. Each \(d \times d\) minor of \(\hat{G}\) is a product of the corresponding minor of \(\hat{G}\) and a number \(\zeta\) satisfying \((\varphi_{d(j_0-1)}(\varepsilon)/2)^d \leq \zeta \leq 1\).
Proof. Consider a square submatrix \( \tilde{\mathbf{S}} \) in \( \tilde{G} \) and the corresponding submatrix \( \hat{\mathbf{S}} \) in \( \hat{G} \). If the corresponding minor is zero, there is nothing to prove. If it is non-zero, we reorder columns and rows of \( \tilde{\mathbf{S}} \) in such a way that all entries on the diagonal become non-zero, and do the same reordering with \( \hat{\mathbf{S}} \). Let \( r_i, c_j > 0 \) be such that after multiplying rows of \( \hat{\mathbf{S}} \) by \( r_i \) and columns of the resulting matrix by \( c_j \) we get \( \tilde{\mathbf{S}} \). Then

\[
\det(\tilde{\mathbf{S}}) = \det(\hat{\mathbf{S}}) \prod_i r_i \prod_j c_j.
\]

On the other hand, \( r_i c_i \geq \varphi_{d(j_0 - 1)}(\varepsilon)/2 \), because the diagonal entry of \( \hat{\mathbf{S}} \) is \( \pm 1 \), and the absolute value of the diagonal entry of \( \tilde{\mathbf{S}} \) is \( \geq \varphi_{d(j_0 - 1)}(\varepsilon)/2 \). The conclusion follows.

Lemma 14 Let \( \mathbf{D} \) be a \( d \times J \) matrix with entries \(-1, 0, \) and \( 1 \), containing a \( d \times d \) identity submatrix. If \( \mathbf{D} \) is not totally unimodular, then it has a square submatrix \( \mathbf{S} \) with \( |\det(\mathbf{S})| \geq 2 \). Let \( \mathbf{S} \) be of size \( h \times h \). Reordering columns and rows of \( \mathbf{D} \) (if necessary), we may assume that \( \mathbf{D} \) is of the form:

\[
\mathbf{D} = \begin{pmatrix} \mathbf{S} & 0 & I_h & * \\ * & I_{d-h} & 0 & * \end{pmatrix},
\]

where \( I_h \) and \( I_{d-h} \) are identity matrices of sizes \( h \times h \) and \((d - h) \times (d - h)\), respectively, 0 denote matrices with zero entries of the corresponding dimensions, and \( * \) denote matrices of the corresponding dimensions with unspecified entries.

We consider all matrices which can be obtained from \( \mathbf{D} \) by a sequence of the following operations:

- Addition or subtraction a row to or from another row,
- Multiplication of a column by \(-1\),

provided that after each such operation we get a matrix with entries \(-1, 0, \) and \( 1 \).

Among all matrices obtained from \( \mathbf{D} \) in such a way we select a matrix \( \hat{\mathbf{D}} \) which satisfies the following conditions:

1. Has all unit vectors among its columns;
2. Has the maximal possible number \( \xi \) of unit vectors among the first \( d \) columns.
Observe that $\xi < d$ because the operations listed above preserve the absolute value of the determinant and at the beginning the absolute value of the determinant formed by the first $d$ columns was $\geq 2$. Let $d_r$ be one of the first $d$ columns of $\hat{D}$ which is not a unit vector. Let $\{i_1, \ldots, i_t\}$ be indices of its non-zero coordinates. Then at least one of the unit vectors $e_{i_1}, \ldots, e_{i_t}$ is not among the first $d$ columns of $\hat{D}$ (the first $d$ columns of $\hat{D}$ are linearly independent). Assume that $e_{i_1}$ is not among the first $d$ columns of $\hat{D}$. We can try to transform $\hat{D}$ adding/subtracting the row number $i_1$ to/from rows number $i_2, \ldots, i_t$ (and multiplying the column number $r$ by $(-1)$, if necessary) into a new matrix $\hat{D}$ which satisfies the following conditions:

- Has among the first $d$ columns all the unit vectors it had before;
- Has $e_{i_1}$ as its column number $r$;
- Has all the unit vectors among its columns.

It is not difficult to verify that the only possible obstacle is that there exists another column $d_t$ in $\hat{D}$, such that for some $s \in \{2, \ldots, t\}$

$$\left| \det \left( \begin{array}{cc} D_{i_1r} & D_{i_1t} \\ D_{isr} & D_{islt} \end{array} \right) \right| = 2,$$

where by $D_{ij}$ we denote entries of $\hat{D}$. By the maximality assumption, a submatrix satisfying (48) exists.

It is easy to see that letting $\{\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4\} = \{d_r, d_s, e_{i_1}, e_{i_s}\}$, and $\{\hat{x}_1, \ldots, \hat{x}_{d-2} = \{e_1, \ldots, e_d\} \setminus \{e_{i_1}, e_{i_s}\}$, we get a set of columns of $\hat{D}$ satisfying the required condition.

Since the operations listed above preserve the absolute values of $d \times d$ minors, the corresponding columns of $D$ form the desired set. ■

**Remark.** Lemma 14 can also be obtained by combining known characterizations of regular and binary matroids, see [24] (we mean, first of all, Theorem 9.1.5, Theorem 6.6.3, Corollary 10.1.4, and Proposition 3.2.6).

We continue our proof of Lemma 13. Assume the contrary. Since there are finitely many possible values of $j_0$, there is $j_0$ and a subset $\Phi_4 \subset (0, 1)$, whose closure contains 0, satisfying the condition:

For each $\varepsilon \in \Phi_4$ there is $Z \in \mathbb{Z}_\varepsilon$ such that following the construction, we get the preselected value of $j_0$, and the obtained matrix $\hat{G}$ is not totally unimodular.

Since the entries of $\hat{G}$ are integers, the absolute values of the minors are at least one. We are going to show that the corresponding minors of $G$ are also ‘sufficiently large’, and get a contradiction using Lemma 7.

By the observation above the corresponding minors of $\hat{G}$ are at least $(\varphi_{d(j_0-1)}(\varepsilon)/2)^d$. The Euclidean norm of a column in $\hat{G}$ is at most $1 + d\varphi_{d(j_0-1)+1}(\varepsilon)$. Applying Lemma 8 $d$ times we get that the corresponding minor of $G$ are at least

$$(\varphi_{d(j_0-1)}(\varepsilon)/2)^d - d^2 \varphi_{d(j_0-1)+1}(\varepsilon) \cdot (1 + d\varphi_{d(j_0-1)+1}(\varepsilon))^{d-1}.$$
We are going to use Lemma 7 for \( x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4 \) defined in the following way. Let \( \hat{x}_1, \ldots, \hat{x}_{d-2}, \hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4 \) be the columns of \( G \) corresponding to the columns \( \hat{x}_1, \ldots, \hat{x}_{d-2}, \hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4 \) of \( \hat{G} \), and \( x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4 \) be their normalizations (that is, \( x_1 = \hat{x}_1/||\hat{x}_1|| \), etc). Since norms of columns of \( G \) are \( \leq 1 \), the condition (21) of Lemma 7 is satisfied with

\[
\chi(\varepsilon) = (\varphi_d(\varepsilon)/2)^d - d^2\varphi_d(1+\varepsilon) \cdot (1 + d\varphi_d(1+\varepsilon))^{d-1}.
\]

Now we recall that columns \( \{g_j\} \) of \( G \) satisfy (40) for some vectors \( b_j \in \Omega(\varepsilon, \delta) \). Hence the distance from \( x_1, \ldots, x_{d-2}, p_1, p_2, p_3, p_4 \) to the corresponding vectors \( b_j \) is \( \leq 2d\varphi_{d_0}(\varepsilon) \). By (17) the condition (20) is satisfied with

\[
\pi(\varepsilon) = 2d\varphi_{d_0}(\varepsilon) + C_3(d)\omega(\varepsilon)
\]

and

\[
\sigma(\varepsilon) = c_3(d)\delta(\varepsilon).
\]

The fact that the conditions (2) and (3) of Lemma 7 are satisfied is verified in the same way as at the end of Lemma 11, the only difference is that instead of (47) we have \((\varphi_d(\varepsilon)) = O(\chi(\varepsilon)) \). This does not affect the rest of the argument. Therefore, under the same condition on \( k \) as in Lemma 11 we get, by Lemma 7, that \( \hat{G} \) should be totally unimodular if \( \varepsilon > 0 \) is small enough. □

**Lemma 15** If \( \hat{G} \) is totally unimodular, then there exists a zonotope \( T \in \mathcal{T}_d \) such that

\[
d(Z, T) \leq t_d(\varepsilon),
\]

where \( t_d(\varepsilon) \) is a function satisfying \( \lim_{\varepsilon \downarrow 0} t_d(\varepsilon) = 1 \).

**Proof.** Observe that the matrix \( \hat{G} \) can be obtained from \( \hat{G} \) using multiplications of rows and columns by positive numbers. Hence, re-scaling the basis \( \{e_i\} \), if necessary, we get: columns of \( \hat{G} \) with respect to the re-scaled basis are of the form \( a_i\tau_i \), where \( \tau_i \) are columns of a totally unimodular matrix (see the definition of \( \mathcal{T}_d \) in the introduction).

We are going to approximate the measure \( \hat{\mu} \) by a measure \( \tilde{\mu} \) supported on vectors which are normalized columns of \( \hat{G} \). Recall that \( \hat{\mu} \) is supported on a finite subset of \( \hat{S} \).

The approximation is constructed in the following way. We erase the measure \( \hat{\mu} \) supported outside \( \Omega(\varepsilon, \delta) \). The total mass of the measure erased in this way is small by (16). As for the measure supported on \( B := \Omega(\varepsilon, \delta) \), we approximate each atom of it by the atom of the same mass supported on the nearest normalized column of \( \hat{G} \). We denote the nearest to \( z \in \text{supp}\hat{\mu} \) normalized column of \( \hat{G} \) by \( A(z) \). If there are several such columns, we choose one of them.

Now we estimate the distance from a point of \( \Omega(\varepsilon, \delta) \) to the nearest normalized column of \( \hat{G} \). The distance from this point to \( \hat{\Omega}(\varepsilon, \delta) \) is \( C_3(d)\omega(\varepsilon) \), the distance from a point from \( \hat{\Omega}(\varepsilon, \delta) \) to the point from \( \Theta(\varepsilon, \delta) \) with the same top set (or its opposite), by Lemma 10, can be estimated from above by \( \sqrt{2\frac{\pi(\varepsilon)}{(\rho(\varepsilon))^2} + 4d\rho(\varepsilon)^2} \).
The distance from a point in $\Theta(\omega(\varepsilon), \delta(\varepsilon))$ to the corresponding column of $G$ is estimated in (40), it is $\leq d \cdot \varphi_{d_0}(\varepsilon)$, so it is $\leq d \cdot \varphi_1(\varepsilon)$, and the distance from a column of $G$ to the corresponding column of $\tilde{G}$ is $\leq d \cdot \varphi_{d(d_0-1)+1}(\varepsilon) \leq d \cdot \varphi_1(\varepsilon)$. Since we have to normalize this vector, the total distance from a point of $(\tilde{\Omega}(\omega(\varepsilon), \delta(\varepsilon)))_{C_3(d)}(d\omega(\varepsilon))$ to the nearest normalized column of $\tilde{G}$ can be estimated from above by

$$C_3(d)\omega(\varepsilon) + \sqrt{2 \frac{\nu(\varepsilon)}{(\rho(\varepsilon))^2} + 4d\rho(\varepsilon)^2 + 4d \cdot \varphi_1(\varepsilon)}$$

It is clear that this function, let us denote it by $\zeta(\varepsilon)$, tends to 0 as $\varepsilon \downarrow 0$, recall that $\rho(\varepsilon) = e^k$, $\nu(\varepsilon) = \varepsilon^{3k}$, $\omega(\varepsilon) = \varepsilon^{4k}$, $\varphi_1(\varepsilon) = \varepsilon^{\left(\frac{1}{2} + \frac{1}{d}\right) - 1}$. The obtained measure corresponds to a zonotope from $T_{d_0}$. Let us denote this zonotope by $T_{d_0}$.

Since the dual norms to the gauge functions of $Z$ and $T$ are their support functions, we get the estimate

$$d(T, Z) \leq \sup_{u \in \tilde{S}} \tilde{h}_Z(u) \cdot \sup_{u \in \tilde{S}} \tilde{h}_T(u).$$

So it is enough to show that

$$C_1(d, \varepsilon) \leq \frac{\tilde{h}_T(u)}{\tilde{h}_Z(u)} \leq C_2(d, \varepsilon),$$

where $\lim_{\varepsilon \downarrow 0} C_1(d, \varepsilon) = \lim_{\varepsilon \downarrow 0} C_2(d, \varepsilon) = 1$.

Observe that Lemma 5 implies that there exists a constant $0 < C_7(d) < \infty$ such that

$$C_7(d) \leq \tilde{h}_Z(u), \quad \forall u \in \tilde{S}.$$ (50)

We have

$$\tilde{h}_Z(u) = \int_{\tilde{S}} |\langle u, z \rangle| d\tilde{\mu}(z) \leq \int_{\tilde{S} \setminus \mathcal{B}} |\langle u, z \rangle| d\tilde{\mu}(z)$$

$$+ \int_{\tilde{S}} |\langle u, z \rangle| d\tilde{\mu}(z) + \sum_{z \in \text{supp } \tilde{\mu} \cap \mathcal{B}} (|\langle u, z \rangle - \langle u, \mathcal{A}(z) \rangle|) \tilde{\mu}(z)$$

$$\leq C_4(d) \frac{\delta(\varepsilon)}{\omega^{d-1}(\varepsilon)} + \tilde{h}_T(u) + \zeta(\varepsilon)\tilde{\mu}(\tilde{S}), \quad \forall u \in \tilde{S}.$$

In a similar way we get

$$\tilde{h}_T(u) = \int_{\tilde{S}} |\langle u, z \rangle| d\tilde{\mu}(z) \leq \int_{\mathcal{B}} |\langle u, z \rangle| d\tilde{\mu}(z)$$

$$+ \sum_{z \in \text{supp } \tilde{\mu} \cap \mathcal{B}} (|\langle u, z \rangle - \langle u, \mathcal{A}(z) \rangle|) \tilde{\mu}(z)$$

$$\leq \tilde{h}_Z(u) + \zeta(\varepsilon)\tilde{\mu}(\tilde{S}), \quad \forall u \in \tilde{S}.$$
Using (50) we get
\[
1 - \frac{C_4(d) \zeta(\varepsilon) \hat{\mu}(\hat{S})}{C_7(d)} \leq \frac{\hat{h}_T(u)}{h_Z(u)} \leq 1 + \frac{\zeta(\varepsilon) \hat{\mu}(\hat{S})}{C_7(d)}.
\]
It is an estimate of the form (49), Q.E.D. ■

It is clear that Lemma 15 completes our proof of Lemma 2. ■

4 Proof of Theorem 4

PROOF. We start by proving Theorem 4 for polyhedral $X$. In this case we can consider $X$ as a subspace of $\ell^m_\infty$ for some $m \in \mathbb{N}$. Since $X$ has an MVSE which is not a parallelepiped, there exists a linear projection $P : \ell^m_\infty \to X$ such that $P(B_m^\infty)$ has the minimal possible volume, but $P(B_m^\infty)$ is not a parallelepiped. Let $d = \dim X$, let $\{q_1, \ldots, q_{m-d}\}$ be an orthonormal basis in ker $P$ and let $\{\tilde{q}_1, \ldots, \tilde{q}_d\}$ be an orthonormal basis in the orthogonal complement of ker $P$. As it was shown in Lemma 4, $P(B_m^\infty)$ is linearly equivalent to the zonotope spanned by rows of $\tilde{Q} = [\tilde{q}_1, \ldots, \tilde{q}_d]$. By the assumption this zonotope is not a parallelepiped. It is easy to see that this assumption is equivalent to: there exists a minimal linearly dependent collection of rows of $\tilde{Q}$ containing $\geq 3$ rows. This condition implies that we can reorder the coordinates in $\ell^m_\infty$ and multiply the matrix $\tilde{Q}$ from the right by an invertible $d \times d$ matrix $C_1$ in such a way that $\tilde{Q} C_1$ has a submatrix of the form
\[
\begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 \\
a_1 & a_2 & \ldots & a_d
\end{pmatrix},
\]
where $a_1 \neq 0$ and $a_2 \neq 0$. Let $X$ be a matrix whose columns form a basis of $X$. The argument of [21] (see the conditions (1)–(3) on p. 96) implies that $X$ can be multiplied from the right by an invertible $d \times d$ matrix $C_2$ in such a way that $XC_2$ is of the form
\[
\begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 \\
sign a_1 & \sign a_2 & \ldots & * \\
\vdots & \vdots & \ddots & \vdots
\end{pmatrix},
\]
where at the top there is an $d \times d$ identity matrix, and all minors of the matrix $XC_2$ have absolute values $\leq 1$. 

31
Changing signs of the first two columns, if necessary, we get that the subspace \( X \subset \ell^m_\infty \) is spanned by columns of the matrix

\[
\begin{pmatrix}
\pm 1 & 0 & 0 & \ldots & 0 \\
0 & \pm 1 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1 \\
1 & 1 & * & \ldots & * \\
& b_1 & c_1 & * & \ldots & * \\
& b_2 & c_2 & * & \ldots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
& b_{m-1} & c_{m-1} & \ldots & * & * \\
\end{pmatrix}.
\]

The condition on the minors implies that \(|b_i| \leq 1|, |c_i| \leq 1|, and |b_i - c_i| \leq 1| for each \( i \). Therefore the subspace, spanned in \( \ell^m_\infty \) by the first two columns of the matrix (51) is isometric to \( \mathbb{R}^2 \) with the norm

\[||(\alpha, \beta)|| = \max(|\alpha|, |\beta|, |\alpha + \beta|).\]

It is easy to see that the unit ball of this space is linearly equivalent to a regular hexagon. Thus, Theorem 4 is proved in the case when \( X \) is polyhedral.

Proving the result for general, not necessarily polyhedral, space, we shall denote the space by \( Y \). We use Theorem 3. Actually we need only the following corollary of it: Each MVSE is a polyhedron. Therefore we can apply the following result to each MVSE.

**Lemma 16** [22, Lemma 1] Let \( Y \) be a finite dimensional space and let \( A \) be a polyhedral MVSE for \( Y \). Then there exists another norm on \( Y \) such that the obtained normed space \( X \) satisfies the conditions:

1. \( X \) is polyhedral;
2. \( B_X \supset B_Y \);
3. \( A \) is an MVSE for \( X \).

So we consider the space \( Y \) as being embedded into a polyhedral space \( \tilde{X} \) with the embedding satisfying the conditions of Lemma 16. By the first part of the proof the space \( X \) satisfies the conditions of Theorem 4 and we may assume that \( X \) is a subspace \( \ell^m_\infty \) in the way described in the first part of the proof. So \( X \) is spanned by columns - let us denote them by \( e_1, \ldots, e_d \) - of the matrix (51) in \( \ell^m_\infty \). It is easy to see that to finish the proof it is enough to show that the vectors \( e_1, e_2, e_1 - e_2 \) are in \( B_Y \).

It turns out each of these points is the center of a facet of a minimum-volume parallelepiped containing \( B_X \). In fact, let \( \{f_i\}_{i=1}^m \) be the unit vector basis of \( \ell^m_\infty \). Let \( P_1 \) and \( P_2 \) be the projections onto \( Y \) with the kernels \( \operatorname{lin}\{f_{d+1}, \ldots, f_m\} \) and \( \operatorname{lin}\{f_1, f_{d+2}, \ldots, f_m\} \), respectively (recall that \( Y \), as a linear space, coincides with \( X \)). The analysis from [20, pp. 318–319] shows that \( P_1(B^m_\infty) \) and \( P_2(B^m_\infty) \) have the minimal possible volume among
all linear projections of $B_m^\infty$ into $X$. It is easy to see that $P_1(B_m^\infty)$ and $P_2(B_m^\infty)$ are parallelepipeds.

We show that $e_1, e_2$ are centers of facets of $P_1(B_m^\infty)$, and that $e_1 - e_2$ is the center of a facet of $P_2(B_m^\infty)$. In fact, the centers of facets of $P_1(B_m^\infty)$ coincide with $P_1(f_1), \ldots, P_1(f_d)$, and it is easy to check that $P_1(f_i) = e_i$ for $i = 1, \ldots, d$. As for $P_2$, we observe that $e_1 - e_2 \in \text{lin}\{f_1, f_2, f_{d+2}, \ldots, f_m\}$, and the coefficient near $f_2$ in the expansion of $e_1 - e_2$ is $\pm 1$. Therefore $P_2(f_2) = \pm(e_1 - e_2)$.

Since the projections $P_1$ and $P_2$ satisfy the minimality condition from [21, Lemma 1] (see, also [20, pp. 318–319]), the parallelepipeds $P_1(B_m^\infty)$ and $P_2(B_m^\infty)$ are MVSE for $X$. Hence, by the conditions of Lemma 16, they are MVSE for $Y$ also. Hence, they are minimum-volume parallelepipeds containing $B_Y$. On the other hand, it is known, see [26, Lemma 3-1], that centers of facets of minimal-volume parallelepipeds containing $B_Y$ should belong to $B_Y$, we get $e_1, e_2, e_1 - e_2 \in B_Y$. The theorem follows.

I would like to thank Gideon Schechtman for turning my attention to the fact that the class $T_d$ was studied in works on lattice tiles.

References


33


