Entropy numbers of finite-dimensional embeddings

Marta Kossaczká*, Jan Vybíral†

February 1, 2018

Abstract

Entropy numbers and covering numbers of sets and operators are well known geometric notions, which found many applications in various fields of mathematics, statistics, and computer science. Their values for finite-dimensional embeddings $id : \ell_p^n \to \ell_q^n$, $0 < p, q \leq \infty$, are known (up to multiplicative constants) since the pioneering work of Schütt in 1984, with later improvements by Edmunds and Triebel, Kühn and Guédon and Litvak. The aim of this survey is to give a self-contained presentation of the result and an overview of the different techniques used in its proof.

Keywords: Entropy number, covering number, finite-dimensional vector space, volume argument, ε-net

1 Introduction

The concept of covering numbers can be traced back to the work of Kolmogorov [23, 24] and became since then an important tool used in several areas of theoretical and applied mathematics. Furthermore, Pietsch introduced in his book [30] a formal definition of an inverse function of covering numbers under the name of entropy numbers. Later on, Carl and Triebel [6, 9] investigated the relation between entropy numbers and other geometric and approximation quantities related to sets and operators, most importantly to eigenvalues of compact operators.

Due to the natural definition of covering and entropy numbers, and due to their relations to other geometric notions, these concepts found applications in many areas of pure and applied mathematics, including geometry of Banach spaces [3, 4, 7, 16, 31], information theory [20, 32, 36, 37, 38], and random processes [27, 28]. They also appeared in the theory of compressed sensing [5, 11] in the study of optimality of recovery of sparse vectors and in the study of eigenvalue problems in Banach spaces [13, 25]. One of the most important classes of operators, whose entropy numbers are well understood and often applied, are the identities between finite-dimensional vector spaces. The main aim of this note is to present a self-contained overview of this area. To state the main result in detail, we need to recall some notation.

The couple $(X, \|\cdot\|_X)$ is called a quasi-Banach space, if X is a real or complex vector space and the mapping $\|\cdot\|_X : X \to [0, \infty)$ satisfies

*Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00, Prague 8, Czech Republic

†Dept. of Mathematics FNSPE, Czech Technical University in Prague, Trojanova 13, 12000 Prague, Czech Republic; jan.vybiral@fjfi.cvut.cz

(i) $\|x\|_X = 0$ if, and only if, $x = 0$,

(ii) $\|\alpha x\|_X = |\alpha| \cdot \|x\|_X$ for all $\alpha \in \mathbb{R}$ (or in $\mathbb{C}$) and all $x \in X$,

(iii) there exists a constant $C \geq 1$ such that $\|x + y\|_X \leq C(\|x\|_X + \|y\|_X)$ for all $x, y \in X$,

(iv) X is complete with respect to $\|\cdot\|_X$.

If the constant C in (iii) can be chosen to be equal to one, X is actually a Banach space. If $\|\cdot\|_X$ satisfies the axioms above with (iii) replaced by

$$\|x + y\|_X^p \leq \|x\|_X^p + \|y\|_X^p, \quad x, y \in X$$

for some $0 < p \leq 1$, then $(X, \|\cdot\|_X)$ is called a p-Banach space and $\|\cdot\|_X$ is a p-norm. It follows that a Banach space X is also a p-Banach space for $p = 1$. It is easy to see that every p-norm is a quasi-norm with $C = 2^{1/p-1}$. On the other hand, by the Aoki-Rolewicz theorem [1, 33], every quasi-norm is equivalent to some p-norm for a suitably chosen p. We refer to [22] for a survey on quasi-Banach spaces.

If X is a vector space equipped with some (quasi-)norm or p-norm $\|\cdot\|_X$, we denote by B_X its unit ball, i.e. the set $B_X = \{x \in X : \|x\|_X \leq 1\}$. The symbol $\mathcal{L}(X, Y)$ stands for the set of all bounded linear operators from X to Y. For $0 < p \leq \infty$, we define $\ell_p^n(\mathbb{R})$ (or $\ell_p^n(\mathbb{C})$) to be the Euclidean space $\mathbb{R}^n$ (or $\mathbb{C}^n$) equipped with the (quasi-)norm

$$\|x\|_p = \|(x_i)_{i=1}^n\|_p = \begin{cases} \Big(\sum\limits_{i=1}^n |x_i|^p\Big)^{1/p}, & 0 < p < \infty\,; \\ \max\limits_{1 \leq i \leq n} |x_i|\,, & p = \infty. \end{cases}$$

The unit ball of $\ell_p^n(\mathbb{R})$ will be denoted by B_p^n. It is easy to see, that $\ell_p^n(\mathbb{R})$ and $\ell_p^n(\mathbb{C})$ are Banach spaces if $p \geq 1$ and p-Banach spaces if $0 < p \leq 1$. Therefore, we will denote $\bar{p} = \min(1, p)$ and use the triangle inequality in $\ell_p^n(\mathbb{R})$ and $\ell_p^n(\mathbb{C})$ in the form

$$\|x + y\|_p^{\bar{p}} \leq \|x\|_p^{\bar{p}} + \|y\|_p^{\bar{p}}.$$

We define now the concept of entropy numbers of a bounded linear operator T between two (quasi-)Banach spaces X and Y. Essentially, we are allowed to use 2^{k-1} balls of radius r in Y to cover the image of the unit ball of X by T and $e_k(T)$ denotes the smallest r, for which this is still possible.

Definition 1. Let X and Y be Banach spaces, p-Banach spaces, or quasi-Banach spaces. Let $T : X \to Y$ be a bounded linear operator and let $k \geq 1$ be a positive integer. The kth (dyadic) entropy number of T is defined as

$$e_k(T) := \inf\Big\{r > 0 : \exists y_1, \ldots, y_{2^{k-1}} \in Y \text{ with } T(B_X) \subset \bigcup_{j=1}^{2^{k-1}} (y_j + rB_Y)\Big\}. \tag{1}$$

The relation of the entropy numbers to the covering numbers of Kolmogorov is quite straightforward. If $K \subset Y$ and $r > 0$, then the covering number $N(K, Y, r)$ is

the smallest number N such that there exist points $y_1, \dots, y_N$ with $K \subset \bigcup_{j=1}^N (y_j + rB_Y)$. The entropy numbers $e_k(T)$ can then be equivalently defined as

$$e_n(T) = \inf\{r > 0 : N(T(B_X), Y, r) \le 2^{k-1}\}.$$

Although easy to define, the entropy numbers of some specific operator T are usually rather difficult to calculate, or estimate. One class of operators, where the upper and lower bounds on entropy numbers are known, are the identities between finite-dimensional vector spaces. The main aim of this note is to present a self-contained proof of the following result.

Theorem 2. *Let $0 < p, q \le \infty$ and let $n \in \mathbb{N}$.*

a) If $0 < p \le q \le \infty$ then for all $k \in \mathbb{N}$ it holds

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \sim \begin{cases} 1 & \text{if} \quad 1 \le k \le \log_2 n, \\ \left(\dfrac{\log_2(1+n/k)}{k}\right)^{\frac{1}{p}-\frac{1}{q}} & \text{if} \quad \log_2 n \le k \le n, \\ 2^{-\frac{k-1}{n}} n^{\frac{1}{q}-\frac{1}{p}} & \text{if} \quad n \le k. \end{cases} \tag{2}$$

b) If $0 < q \le p \le \infty$ then for all $k \in \mathbb{N}$ it holds

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \sim 2^{-\frac{k-1}{n}} n^{\frac{1}{q}-\frac{1}{p}}. \tag{3}$$

The constants of equivalence in both (2) *and* (3) *may depend on p and q, but are independent of k and n.*

Theorem 2 was first proved for $1 \le p < q \le \infty$ by Schütt [34] with partial results given before by Höllig [21] and Pietsch [30, Section 12.2]. The case of quasi-Banach spaces (i.e. when p and/or q is smaller than one) was studied by Edmunds and Triebel in [13, Section 3.2.2], who provided the estimates from above in (2) including the intermediate k's with $\log_2 n \le k \le n$. Finally, the corresponding lower bound was supplied by Kühn [26] and, independently, by Guédon and Litvak in [18]. Although all the estimates in Theorem 2 may be found in the literature since nearly two decades, there seems to be no place, were all the parts would appear together - the reader has to combine several sources, sometimes using different notations.

Interestingly enough, the proof of Theorem 2 requires a combination of several different techniques, which are of independent interest. The most natural are the so-called volume arguments. Quite intuitively, if we want to cover a body $K \subset \mathbb{R}^n$ by a number of other bodies $L_1, \dots, L_N \subset \mathbb{R}^n$, then their volume combined must be larger than the volume of K. Here, volume can be any positive measure on $\mathbb{R}^n$. As L_j's are now different translates of a fixed dilation of one fixed body L, it is most convenient to work with a shift-invariant measure, which behaves well with respect to dilations. It is therefore most natural to work with the usual Lebesgue measure on $\mathbb{R}^n$ and $\mathrm{vol}(K)$ will be the usual Lebesgue volume of a measurable set $K \subset \mathbb{R}^n$. If $K = B_X$ and $L = B_Y$ are unit balls of some quasi-Banach spaces X and Y, and if the number N is fixed to be $N = 2^{k-1}$ for a positive integer k, this can be immediately translated into lower bounds of the entropy numbers $e_k(id : X \to Y)$.

With a bit of additional work, volume arguments can be also used to give upper bounds on entropy numbers. Indeed, if $r > 0$ is fixed, we take a maximal set $y_1, \dots, y_N \in B_X \subset Y$ with $\|y_i - y_j\|_Y \geq r$. Then, on one hand, the sets $y_j + c_1 r B_Y$ are disjoint and B_X is covered by the union of $y_j + c_2 r B_Y$ for suitably chosen constants $c_1, c_2 > 0$. On the other hand, the sets $y_j + c_1 r B_Y$ are included in a certain multiple of B_X, which must therefore have a larger volume than all the disjoint sets $y_j + c_1 r B_Y$ combined. This gives an upper bound on their number N, leading to an upper bound on the entropy numbers.

Although the volume arguments represent a powerful technique, which can in principle be applied to all the parameter settings in Theorem 2, the obtained bounds are not always optimal. Actually, it turns out that the volume arguments provide optimal bounds (up to a multiplicative constant) only when k is large or, equivalently, when r is small. For smaller k's, direct combinatorial estimates are needed to provide both the lower and the upper bounds. Geometrically it means, that any good cover of $T(B_X)$ with a small number of sets $y_j + r B_Y$ needs to have big overlap and/or to cover also some large neighborhood of $T(B_X)$.

The structure of the paper is as follows. Section 2 gives basic properties of entropy numbers and Gamma function, presents the calculation of the volume of B_p^n, and provides a couple of lemmas used later. The proof of Theorem 2 comes in Section 3. Finally, Section 4 collects few additional remarks and topics, including the extension of Theorem 2 to the complex setting.

2 Preparations

We start by recalling few well-known basic facts about entropy numbers. Although the reader may find the proof, for instance, in [8] or [13], we include it for the sake of completeness.

Theorem 3. *Let X, Y, Z be p-Banach spaces for some $0 < p \leq 1$. Let $S, T \in \mathcal{L}(X, Y)$ and $R \in \mathcal{L}(Y, Z)$. Then, for all $k, l \in \mathbb{N}$, it holds*

(i) $\|T\| \geq e_1(T) \geq e_2(T) \geq \dots \geq 0$ *(monotonicity);*

(ii) $e_{k+l-1}^p(S+T) \leq e_k^p(S) + e_l^p(T)$ *(subadditivity);*

(iii) $e_{k+l-1}(R \circ T) \leq e_k(R) \cdot e_l(T)$ *(submultiplicativity).*

Proof. Monotonicity of entropy numbers follows directly from (1). Similarly, $e_1(T) \leq \|T\|$ is implied by $T(B_X) \subset \|T\| \cdot B_Y$.

For the proof of subadditivity, let k, l be positive integers and let $\varepsilon > 0$. Then there are $y_1, \dots, y_{2^{k-1}} \in Y$ and $z_1, \dots, z_{2^{l-1}} \in Y$ with

$$S(B_X) \subset \bigcup_{i=1}^{2^{k-1}} \Big(y_i + (e_k(S) + \varepsilon) B_Y \Big) \quad \text{and} \quad T(B_X) \subset \bigcup_{j=1}^{2^{l-1}} \Big(z_j + (e_l(T) + \varepsilon) B_Y \Big).$$

Then

$$\begin{aligned}(S+T)(B_X) &\subset S(B_X)+T(B_X)\\ &\subset \Big[\bigcup_{i=1}^{2^{k-1}}\Big(y_i+(e_k(S)+\varepsilon)B_Y\Big)\Big]+\Big[\bigcup_{j=1}^{2^{l-1}}\Big(z_j+(e_l(T)+\varepsilon)B_Y\Big)\Big]\\ &=\bigcup_{i,j}\Big(y_i+z_j+(e_k(S)+\varepsilon)B_Y+(e_l(T)+\varepsilon)B_Y\Big)\\ &\subset \bigcup_{i,j}\Big(y_i+z_j+[(e_k(S)+\varepsilon)^p+(e_l(T)+\varepsilon)^p]^{1/p}B_Y\Big)\end{aligned}$$

and taking the infimum over $\varepsilon>0$ gives the result.

The proof of submultiplicativity follows in a similar way. Indeed, if k,l are positive integers and $\varepsilon>0$, we find $y_1,\dots,y_{2^{l-1}}\in Y$ and $z_1,\dots,z_{2^{k-1}}\in Z$ with

$$T(B_X)\subset\bigcup_{i=1}^{2^{l-1}}\Big(y_i+(e_l(T)+\varepsilon)B_Y\Big)\quad\text{and}\quad R(B_Y)\subset\bigcup_{j=1}^{2^{k-1}}\Big(z_j+(e_k(R)+\varepsilon)B_Z\Big).$$

The proof then follows from

$$\begin{aligned}(R\circ T)(B_X)=R(T(B_X))&\subset R\Big(\bigcup_{i=1}^{2^{l-1}}\Big(y_i+(e_l(T)+\varepsilon)B_Y\Big)\Big)\\ &=\bigcup_{i=1}^{2^{l-1}}\Big(Ry_i+(e_l(T)+\varepsilon)R(B_Y)\Big)\\ &\subset\bigcup_{i,j}\Big(Ry_i+(e_l(T)+\varepsilon)z_j+(e_l(T)+\varepsilon)(e_k(R)+\varepsilon)B_Z\Big).\qquad\square\end{aligned}$$

We will also need few basic facts about the Gamma function, which is defined by $\Gamma(t)=\int_0^\infty x^{t-1}e^{-x}dx$ for every positive real number $t>0$. By partial integration, we get

$$\Gamma(1+t)=t\,\Gamma(t)\quad\text{for every}\quad t>0 \tag{4}$$

and $\Gamma(n)=(n-1)!$ for every positive integer n. Furthermore, by standard calculus it follows that Γ is a continuous function on $(0,\infty)$.

Lemma 4. *Let $0<p<\infty$. Then*

$$\Gamma(1+x/p)^{1/x}\sim x^{1/p},\quad x\geq 1, \tag{5}$$

where the constants of equivalence may depend on p.

Proof. The result is a corollary of Stirling's formula [35, Chapter A.2, Theorem 2.3], but we give also a simple proof for reader's convenience.

If $x=n$ is a positive integer and $p=1$, the result follows from

$$n\log(n)-n\leq\log(n!)=\log(\Gamma(1+n))\leq n\log(n), \tag{6}$$

where the right-hand side of (6) comes from $n!\leq n^n$ and the left-hand side can be obtained by taking the Riemann sum of $\int_1^n\log(x)dx$ at right endpoints.

We modify this strategy also for $x \geq 1$ and $p > 0$. Let $k \in \mathbb{N}_0$ be the unique integer with

$$\frac{1}{p} + k \leq \frac{x}{p} < \frac{1}{p} + k + 1.$$

Iterative application of (4) leads to

$$\Gamma(1 + x/p) = \Gamma(x/p - k + 1) \prod_{j=0}^{k-1} (x/p - j) \leq C_p (x/p)^k, \tag{7}$$

where $C_p = \sup_{1/p+1 \leq t \leq 1/p+2} \Gamma(t)$. We conclude, that

$$\Gamma(1 + x/p)^{1/x} x^{-1/p} \leq C_p^{1/x} p^{-k/x} x^{k/x - 1/p} \leq C_p^{1/x} \max(1, p^{-1/p}) x^{-1/px},$$

which is bounded (by a constant depending on p) for $x \in [1, \infty)$.

On the other hand, using the first identity in (7) and Riemann sums, we obtain

$$\begin{aligned}
\log\left(\frac{\Gamma(1 + x/p)^{1/x}}{x^{1/p}}\right) &= \frac{1}{x} \log \Gamma(1 + x/p) - \frac{1}{p} \log(x) \\
&\geq \frac{1}{x} \log(c_p) + \frac{1}{x} \sum_{j=0}^{k-1} \log(x/p - j) - \frac{1}{p} \log(x) \\
&\geq \frac{1}{x} \log(c_p) + \frac{1}{x} \int_{x/p-k}^{x/p} \log(t) dt - \frac{1}{p} \log(x) \\
&= \frac{1}{x} \log(c_p) + \frac{1}{p} \log(1/p) - \frac{1}{p} - \frac{1}{x} f(x/p - k),
\end{aligned}$$

where $c_p = \inf_{1/p+1 \leq t \leq 1/p+2} \Gamma(t) > 0$ and $f(t) = t \log(t) - t$. The last expression is bounded (by a constant depending on p) for $x \in [1, \infty)$, as $1/p \leq x/p - k < 1/p + 1$. □

To apply the volume arguments, it is of course necessary to calculate (or at least to estimate) the volume of the unit ball B_p^n in $\ell_p^n(\mathbb{R})$. The exact value has been known (at least) since the work of Dirichlet, cf. [10]. We give a more modern proof, cf. [31].

Theorem 5. *Let $0 < p \leq \infty$ and let $n \in \mathbb{N}$. Then it holds*

$$\mathrm{vol}(B_p^n) = \frac{2^n \cdot \Gamma(1 + \frac{1}{p})^n}{\Gamma(1 + \frac{n}{p})}. \tag{8}$$

Proof. Let f be a smooth non-increasing positive function on $[0, \infty)$ quickly decaying

to zero at infinity. Then, by Fubini's theorem and partial integration,

$$\int_{\mathbb{R}^n} f(\|x\|_p)dx = -\int_{\mathbb{R}^n}\int_{\|x\|_p}^{\infty} f'(t)dtdx = -\int_0^{\infty}\Big(\int_{x:\|x\|_p\le t} 1\,dx\Big)f'(t)dt$$
$$= -\int_0^{\infty}\mathrm{vol}(tB_p^n)f'(t)dt = -\mathrm{vol}(B_p^n)\int_0^{\infty} t^n f'(t)dt$$
$$= \mathrm{vol}(B_p^n)\cdot\int_0^{\infty} nt^{n-1}f(t)dt.$$

For $f(t)=e^{-t^p}$, we get

$$\int_{\mathbb{R}^n} e^{-\|x\|_p^p}dx = \mathrm{vol}(B_p^n)\cdot\int_0^{\infty} nt^{n-1}e^{-t^p}dt = \mathrm{vol}(B_p^n)\cdot\frac{n}{p}\cdot\int_0^{\infty} s^{n/p-1}e^{-s}ds$$
$$= \frac{n\,\mathrm{vol}(B_p^n)\Gamma(n/p)}{p} = \mathrm{vol}(B_p^n)\Gamma(1+n/p). \tag{9}$$

The proof is then finished by Fubini's theorem

$$\mathrm{vol}(B_p^n)\Gamma(1+n/p) = \int_{\mathbb{R}^n} e^{-\|x\|_p^p}dx = \int_{\mathbb{R}^n} e^{-|x_1|^p-\dots-|x_n|^p}dx = \Big(2\int_0^{\infty} e^{-t^p}dt\Big)^n$$
$$= 2^n\Big(\frac{1}{p}\int_0^{\infty} s^{1/p-1}e^{-s}ds\Big)^n = 2^n\Big(\frac{\Gamma(1/p)}{p}\Big)^n = 2^n\Gamma(1+1/p)^n.$$

□

Theorem 5 combined with Lemma 4 gives

$$\mathrm{vol}(B_p^n)^{1/n} = \frac{2\Gamma(1+\frac{1}{p})}{\Gamma(1+\frac{n}{p})^{1/n}} \sim n^{-1/p}, \quad n\ge 1, \tag{10}$$

where the constants of equivalence depend again on p.

Next, we collect two simple facts about the ℓ_p^n-(quasi-)norms. The easy proof is left to the reader.

Lemma 6. *(i) Let $0<p,q\le\infty$ and $n\in\mathbb{N}$. Then*

$$\|id:\ell_p^n(\mathbb{R})\to\ell_q^n(\mathbb{R})\| = \max(1, n^{1/q-1/p}).$$

(ii) Let $0<p<q<\infty$. Then $\|x\|_q\le\|x\|_p^{p/q}\cdot\|x\|_\infty^{1-p/q}$.

The following lemma is the analogue of [30, Proposition 12.1.13] for quasi-Banach spaces and appeared already in [19]. Although it uses the volume arguments, no calculation of $\mathrm{vol}(B_X)$ is necessary, because two such terms cancel each other out. The estimate obtained is optimal up to the constant $4^{1/p}$, cf. Section 4.1 for details.

Lemma 7. *Let $0 < p \le 1$, $n \in \mathbb{N}$, and let X be a real n-dimensional p-Banach space. Then, for all $k \in \mathbb{N}$,*

$$e_k(id : X \to X) \le 4^{1/p} \cdot 2^{-\frac{k-1}{n}}. \tag{11}$$

Proof. If $k-1 \le 2n/p$, the result is trivial as the right-hand side of (11) is larger than or equal to 1. We will therefore assume that $k-1 > 2n/p$ or, equivalently, $2^{(k-1)p/n} > 4$.

Let $x_1, \dots, x_N$ be a maximal family of elements of B_X with $\|x_i - x_j\|_X > \tau$ for $i \ne j$, where $\tau > 0$ is given by $\frac{1+\tau^p/2}{\tau^p/2} = 2^{p(k-1)/n}$. Then, by triangle inequality, the sets $x_i + 2^{-1/p}\tau B_X$ are mutually disjoint, they are all included in $(1+\tau^p/2)^{1/p} B_X$ and B_X is covered by the union of $x_i + \tau B_X$ over $i = 1, \dots, N$. By volume comparison, we get

$$N \operatorname{vol}(2^{-1/p}\tau B_X) \le \operatorname{vol}[(1+\tau^p/2)^{1/p} B_X]$$

and, therefore,

$$N \le \frac{(1+\tau^p/2)^{n/p}}{2^{-n/p}\tau^n} = \Big(\frac{1+\tau^p/2}{\tau^p/2}\Big)^{n/p} = 2^{k-1}.$$

We conclude that

$$e_k(id : X \to X) \le \tau = \left[\frac{2}{2^{(k-1)p/n} - 1}\right]^{1/p} \le \left[\frac{4}{2^{(k-1)p/n}}\right]^{1/p} = 4^{1/p} 2^{-\frac{k-1}{n}}. \quad \square$$

The behavior of entropy numbers with respect to interpolation of Banach spaces was studied intensively. It is rather easy to show, that they behave well if one of the endpoints is fixed and we refer to [13] for details. Surprisingly, it was shown only recently in [12], that a general interpolation formula for entropy numbers with both endpoints interpolated is out of reach. We give only a simplified version in a form, which shall be needed later on.

Lemma 8. *Let $0 < p \le q < \infty$ and let k, l, n be positive integers. Then*

$$e_{k+l-1}(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \le 2^{1/\bar{p}} e_k^{p/q}(id : \ell_p^n(\mathbb{R}) \to \ell_p^n(\mathbb{R})) \cdot e_l^{1-p/q}(id : \ell_p^n(\mathbb{R}) \to \ell_\infty^n(\mathbb{R})),$$

where $\bar{p} = \min(1, p)$.

Proof. Let $\varepsilon > 0$ be arbitrary. We put

$$e^0 := (1+\varepsilon) e_k(id : \ell_p^n(\mathbb{R}) \to \ell_p^n(\mathbb{R})) \quad \text{and} \quad e^1 := (1+\varepsilon) e_l(id : \ell_p^n(\mathbb{R}) \to \ell_\infty^n(\mathbb{R})).$$

Then B_p^n can be covered by 2^{k-1} balls in $\ell_p^n(\mathbb{R})$ with radius e^0 and by 2^{l-1} balls in $\ell_\infty^n(\mathbb{R})$ with radius e^1. We can therefore decompose

$$B_p^n = \bigcup_i A_i,$$

where this union contains at most 2^{k-1} components A_i, and each A_i lies in some ball in $\ell_p^n(\mathbb{R})$ with radius e^0. Similarly we can write

$$B_p^n = \bigcup_j B_j,$$

with at most 2^{l-1} components B_j, each of them lying in some ball in $\ell_\infty^n(\mathbb{R})$ with radius e^1.

Finally, we denote $C_{i,j} = A_i \cap B_j$ and choose $z_{i,j} \in C_{i,j}$ arbitrarily any time $C_{i,j}$ is non-empty. Let now $x \in C_{i,j}$. Then both x and $z_{i,j}$ are in $C_{i,j} \subset A_i$. Therefore $\|x - z_{i,j}\|_p \leq 2^{1/\bar{p}} e^0$. Using $C_{i,j} \subset B_j$, we get in the same way also $\|x - z_{i,j}\|_\infty \leq 2e^1 \leq 2^{1/\bar{p}} e^1$. Hence, by Lemma 6,

$$\|x - z_{i,j}\|_q \leq 2^{1/\bar{p}} (e^0)^{p/q} (e^1)^{1-p/q},$$

and balls with centers in $z_{i,j}$'s with this radius in $\ell_q^n(\mathbb{R})$ cover B_p^n. Finally, there is at most 2^{k+l-2} such points. □

Some of the arguments used in the proof of Theorem 2 are of combinatorial nature. As a preparation, we present the following lemma from [26], another variant is discussed in the last section.

Lemma 9. *Let $m, n \in \mathbb{N}$ with $m \leq \frac{n}{4}$ and define*

$$H_m = \{x \in \{-1, 0, 1\}^n : \|x\|_1 = 2m\}.$$

Furthermore, for $x, y \in H_m$, define their Hamming distance *as $h(x, y) = \#\{i : x_i \neq y_i\}$. Then there is a set $A_m \subset H_m$ with $\#A_m \geq [n/(2m)]^m$, such that any two distinct $x, y \in A_m$ satisfy $h(x, y) > m$.*

Proof. First note, that

$$\#H_m = \binom{n}{2m} 2^{2m}. \tag{12}$$

Second, if $x \in H_m$ is arbitrary, then the set $\{y \in H_m : h(x, y) \leq m\}$ has cardinality at most

$$\binom{n}{m} 3^m. \tag{13}$$

Indeed, there is $\binom{n}{m}$ ways to choose m coordinates, where x and y may differ, and three possible values for each of the coordinates for y.

The set A_m can be constructed by the following greedy algorithm. First, choose $x^1 \in H_m$ arbitrarily. If $x^1, \ldots, x^l$ were already selected and if there is some $y \in H_m$, which has the Hamming distance from all these points at least equal to $m + 1$, put $x^{l+1} := y$. Otherwise, the algorithm stops and A_m is the collection of all $x^1, \ldots, x^N$, which were selected so far. This ensures, that $h(x, y) \geq m + 1$ for any two distinct $x, y \in A_m$.

Finally, from (12) and (13) it follows that

$$\#A_m = N \geq \frac{\binom{n}{2m} 2^{2m}}{\binom{n}{m} 3^m} = \frac{4^m}{3^m} \cdot \frac{(n-m)\ldots(n-2m+1)}{(2m)\ldots(m+1)} \geq \frac{4^m}{3^m} \Big(\frac{n-m}{2m}\Big)^m \geq \Big(\frac{n}{2m}\Big)^m.$$

We have used that $t \to \frac{n-2m+t}{m+t}$ is decreasing on $(0, \infty)$ and $n \geq 4m$. □

3 Proof of Theorem 2

This section is devoted to the proof of the main result, Theorem 2. For reader's convenience, we repeat its statement and then prove step-by-step all the upper and lower bounds of $e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R}))$ for all possible values of p, q, k and n.

Theorem 2. *Let $0 < p, q \leq \infty$ and let $n \in \mathbb{N}$.*

a) If $0 < p \leq q \leq \infty$ then for all $k \in \mathbb{N}$ it holds

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \sim \begin{cases} 1 & \text{if} \quad 1 \leq k \leq \log_2 n, \\ \left(\dfrac{\log_2(1+n/k)}{k}\right)^{\frac{1}{p}-\frac{1}{q}} & \text{if} \quad \log_2 n \leq k \leq n, \\ 2^{-\frac{k-1}{n}} n^{\frac{1}{q}-\frac{1}{p}} & \text{if} \quad n \leq k. \end{cases} \tag{2}$$

b) If $0 < q \leq p \leq \infty$ then for all $k \in \mathbb{N}$ it holds

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \sim 2^{-\frac{k-1}{n}} n^{\frac{1}{q}-\frac{1}{p}}. \tag{3}$$

The constants of equivalence in both (2) and (3) may depend on p and q, but are independent of k and n.

Proof. We split the proof of the different estimates in Theorem 2 by the used technique. Furthermore, we denote throughout the proof $\bar{p} = \min(1, p)$ and $\bar{q} = \min(1, q)$.

<u>*Step 1: Elementary estimates*</u>

(i) If $0 < p \leq q \leq \infty$, we have by Theorem 3 and Lemma 6

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \leq \|id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})\| = 1$$

or, equivalently, we can always cover B_p^n with B_q^n. This gives the optimal upper bound for $1 \leq k \leq \log_2 n$.

(ii) On the other side, the canonical vectors $e^1, \ldots, e^n$ defined by

$$(e^i)_l = \delta_{il}, \quad i, l = 1, \ldots, n,$$

satisfy $e^i \in B_p^n$ and $\|e^i - e^j\|_q = 2^{1/q}$ for $i, j \in \{1, \ldots, n\}$ and $i \neq j$. Therefore, if we cover B_p^n with $2^{k-1} < n$ balls in $\ell_q^n(\mathbb{R})$ of radius $r > 0$, at least one of these balls must contain two different e^i, e^j with $i \neq j$, simply by the pigeonhole principle. We denote the center of such a ball by $z \in \mathbb{R}^n$ and obtain by triangle inequality

$$2^{\bar{q}/q} = \|e^i - e^j\|_q^{\bar{q}} \leq \|e^i - z\|_q^{\bar{q}} + \|z - e^j\|_q^{\bar{q}} \leq 2r^{\bar{q}}.$$

Therefore,

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \geq 2^{1/q - 1/\bar{q}}$$

for $1 \leq k \leq \log_2 n$. Together with (i), this finishes the proof of the equivalence in (2) for this range of k's.

(iii) Lemma 7 together with Lemma 6 can be used to prove the upper bound in (3). Let $0 < q \le p \le \infty$ and let k and n be positive integers. Then

$$\begin{aligned} e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) &\le e_k(id : \ell_p^n(\mathbb{R}) \to \ell_p^n(\mathbb{R})) \cdot \|id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})\| \\ &\le 4^{1/p} \cdot 2^{-\frac{k-1}{n}} \cdot n^{1/q-1/p}. \end{aligned}$$

Step 2: Volume arguments, estimates from below

If B_p^n is covered by 2^{k-1} balls in $\ell_q^n(\mathbb{R})$ of radius $r > 0$, then their volume must be larger than the volume of B_p^n. This simple observation can be turned into the estimate

$$\mathrm{vol}(B_p^n) \le 2^{k-1}\,\mathrm{vol}(rB_q^n) = 2^{k-1} r^n\,\mathrm{vol}(B_q^n)$$

and, by (10),

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \ge 2^{-\frac{k-1}{n}} \left(\frac{\mathrm{vol}(B_p^n)}{\mathrm{vol}(B_q^n)}\right)^{1/n} \sim 2^{-\frac{k-1}{n}} n^{1/q-1/p}.$$

This proves the lower bound in (3) for all k's and in (2) for $k \ge n$. Together with Step 1, the proof of (3) is therefore finished.

Step 3: Volume arguments, estimates from above

Let $0 < p \le q \le \infty$ and let $\tau > 0$. Similarly to the proof of Lemma 7, we let $\{y_1, \dots, y_N\} \subset B_p^n$ be any maximal τ-distant set in the ℓ_q^n-(quasi)-norm. In detail, this means that we assume that $\|y_i - y_j\|_q > \tau$ for $i, j \in \{1, \dots, N\}$ with $i \ne j$ and that for every $y \in B_p^n$, there is $i \in \{1, \dots, N\}$ with $\|y - y_i\|_q \le \tau$.

Then we observe a couple of simple facts

(i) If $z \in (y_i + 2^{-1/\bar{q}}\tau B_q^n) \cap (y_j + 2^{-1/\bar{q}}\tau B_q^n)$ for $i \ne j$, then

$$\tau^{\bar{q}} < \|y_i - y_j\|_q^{\bar{q}} \le \|y_i - z\|_q^{\bar{q}} + \|z - y_j\|_q^{\bar{q}} \le 2(2^{-1/\bar{q}}\tau)^{\bar{q}},$$

gives a contradiction. Hence, the sets $y_j + 2^{-1/\bar{q}}\tau B_q^n$, $j = 1, \dots, N$, are disjoint.

(ii) If $z \in y_j + 2^{-1/\bar{q}}\tau B_q^n$, we obtain for $\nu = \|id : \ell_q^n(\mathbb{R}) \to \ell_p^n(\mathbb{R})\| = n^{1/p-1/q}$

$$\|z\|_p^{\bar{p}} \le \|y_j\|_p^{\bar{p}} + \|z - y_j\|_p^{\bar{p}} \le 1 + \nu^{\bar{p}}\|z - y_j\|_q^{\bar{p}} \le 1 + (2^{-1/\bar{q}}\tau\nu)^{\bar{p}}.$$

It follows that all the sets $y_j + 2^{-1/\bar{q}}\tau B_q^n$, $j = 1, \dots, N$ are included in $(1 + (2^{-1/\bar{q}}\tau\nu)^{\bar{p}})^{1/\bar{p}} B_p^n$.

By volume comparison, we conclude

$$N\,\mathrm{vol}(2^{-1/\bar{q}}\tau B_q^n) \le \mathrm{vol}[(1 + (2^{-1/\bar{q}}\tau\nu)^{\bar{p}})^{1/\bar{p}} B_p^n]$$

and

$$N \le \frac{(1 + (2^{-1/\bar{q}}\tau\nu)^{\bar{p}})^{n/\bar{p}}\,\mathrm{vol}(B_p^n)}{2^{-n/\bar{q}}\tau^n\,\mathrm{vol}(B_q^n)} = \left[\frac{1 + (2^{-1/\bar{q}}\tau\nu)^{\bar{p}}}{2^{-\bar{p}/\bar{q}}\tau^{\bar{p}}}\right]^{n/\bar{p}} V_n(p,q), \tag{14}$$

where we denoted $V_n(p,q) = \frac{\mathrm{vol}(B_p^n)}{\mathrm{vol}(B_q^n)}$. For a positive integer k, we define τ by putting the right-hand side of (14) equal to 2^{k-1}. We obtain

$$\left[\frac{1 + (2^{-1/\bar{q}}\tau\nu)^{\bar{p}}}{2^{-\bar{p}/\bar{q}}\tau^{\bar{p}}}\right]^{n/\bar{p}} V_n(p,q) = 2^{k-1} \quad \text{and} \quad \tau = \frac{2^{1/\bar{q}}}{[2^{(k-1)\bar{p}/n} V_n(p,q)^{-\bar{p}/n} - \nu^{\bar{p}}]^{1/\bar{p}}}.$$

Observe, that this is always possible if the denominator is positive. This is the case if

$$2^{\frac{k-1}{n}} > 2\nu V_n(p,q)^{1/n} = 2n^{1/p-1/q}\Big(\frac{\mathrm{vol}(B_p^n)}{\mathrm{vol}(B_q^n)}\Big)^{1/n}. \tag{15}$$

By (10), the right-hand side of (15) is equivalent to a constant, and (15) is satisfied if $k \geq \gamma_{p,q} n$ for some $\gamma_{p,q} > 0$ depending only on p and q.

We conclude that, for $k \geq \gamma_{p,q} n$,

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \leq \tau \leq \frac{C_{p,q}}{[2^{(k-1)\bar{p}/n} V_n(p,q)^{-\bar{p}/n}]^{1/\bar{p}}} \leq C'_{p,q} 2^{-\frac{k-1}{n}} n^{1/q-1/p},$$

where we used (15) in the last but one inequality with $C_{p,q} = 2^{1+1/\bar{q}}/(2^{\bar{p}}-1)^{1/\bar{p}}$.

Step 4: Combinatorial part - estimate from below

Let $0 < p \leq q \leq \infty$. Let $m, n \in \mathbb{N}$ with $m \leq n/4$ and let A_m be the set from Lemma 9. Then $\tilde{A}_m := (2m)^{-1/p} A_m \subset B_p^n$ and for two distinct $x, y \in \tilde{A}_m$ it holds $\|x-y\|_q \geq 2^{-1/p} m^{1/q-1/p}$. As a consequence, two different points from $\tilde{A}_m$ cannot be covered by an ℓ_q^n-ball, unless its radius is at least $2^{-1/p-1/\bar{q}} m^{1/q-1/p}$.

We summarize, that if for some $k \in \mathbb{N}$, there is $m \in \mathbb{N}$ with

$$m \leq n/4 \quad \text{and} \quad 2^{k-1} < \Big(\frac{n}{2m}\Big)^m, \tag{16}$$

then $e_k(id : \ell_p^n \to \ell_q^n) \geq 2^{-1/p-1/\bar{q}} m^{1/q-1/p}$.

Let us therefore fix $k, n \in \mathbb{N}$ with $n \geq 64$ and $\log_2 n \leq k \leq n/8$. Then $k \geq \log_2 n \geq \log_2(n/k+1)$ and we may choose m to be any integer with

$$\frac{k}{\log_2(n/k+1)} \leq m \leq \frac{2k}{\log_2(n/k+1)}.$$

Then $m \leq 2k/\log_2(8) = 2k/3 \leq n/12$. The function $f : t \to t\log_2(n/(2t))$ is increasing on $(0, \frac{n}{2e})$ and therefore

$$\begin{aligned} m\log_2\Big(\frac{n}{2m}\Big) &= f(m) \geq f\Big(\frac{k}{\log_2(n/k+1)}\Big) \\ &= \frac{k}{\log_2(n/k+1)} \log_2\Big(\frac{n\log_2(n/k+1)}{2k}\Big) \geq k, \end{aligned}$$

where in the last inequality we used that

$$\frac{n}{2k}\log_2(n/k+1) \geq \frac{3n}{2k} \geq \frac{n}{k} + 1.$$

We therefore get also

$$2^k \leq \Big(\frac{n}{2m}\Big)^m.$$

This finishes the proof of (16). It follows that

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \geq 2^{-1/p-1/\bar{q}} m^{1/q-1/p} \geq c_{p,q}\Big(\frac{\log_2(n/k+1)}{k}\Big)^{1/p-1/q}$$

for $n \geq 64$ and $\log_2 n \leq k \leq n/8$. Furthermore, by Step 2, we know that $e_n(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \gtrsim n^{1/q-1/p}$. This allows to use the monotonicity of entropy numbers and to obtain the lower bound in (2) for all $n \geq 64$ and all positive integers k. The (finitely many) remaining values of n can then also be incorporated at the cost of possibly larger constants.

Step 5: Combinatorial part - estimate from above

We first prove the result for $0 < p < q = \infty$, the general case then follows by interpolation, i.e. by Lemma 8. Let $1 \leq m \leq n$ be two natural numbers. Then every $x \in B_p^n$ can be approximated in the ℓ_∞^n-norm by a vector $\tilde{x}$ with at most m non-zero components and $\|x - \tilde{x}\|_\infty \leq r^0 := m^{-1/p}$. Indeed, we can take $\tilde{x}$ equal to x on the indices of the m largest (in the absolute value) components of x, and zero elsewhere.

Let now $l \geq \gamma_{p,\infty} m$ be an integer, where $\gamma_{p,\infty} > 0$ is the constant defined at the end of Step 3. Then we know that $e_l(id : \ell_p^m(\mathbb{R}) \to \ell_\infty^m(\mathbb{R})) \leq r^1 := C_{p,\infty} 2^{-\frac{l-1}{m}} m^{-1/p}$. Therefore, for any $\varepsilon > 0$, there exist points $x_1, \ldots, x_{2^{l-1}} \in \mathbb{R}^m$ with

$$B_p^m \subset \bigcup_{j=1}^{2^{l-1}} (x_j + (1+\varepsilon) r^1 B_\infty^m).$$

we collect the points with support of the size at most m and equal to some of x_j's on its support. In this way, we obtain at most $\binom{n}{m} 2^{l-1}$ centers of ℓ_∞^n-balls of radius $r^0 + (1+\varepsilon) r^1$, which cover B_p^n.

We conclude, that if $1 \leq m \leq n$ and $l \geq \gamma_{p,\infty} m$, then

$$2^{k-1} \geq \binom{n}{m} 2^{l-1} \quad \Longrightarrow \quad e_k(id : \ell_p^n(\mathbb{R}) \to \ell_\infty^n(\mathbb{R})) \leq m^{-1/p} \Big(1 + C_{p,\infty} 2^{-\frac{l-1}{m}}\Big). \tag{17}$$

For $1 \leq m \leq n$, we choose any integer l with $\gamma_{p,\infty} m \leq l \leq (\gamma_{p,\infty} + 1) m$. Using the elementary estimate

$$\binom{n}{m} \leq \frac{n^m}{m!} \leq n^m \Big(\frac{e}{m}\Big)^m \leq \Big(\frac{3n}{m}\Big)^m \leq \Big(\frac{n}{m} + 1\Big)^{3m}$$

we obtain

$$\log_2 \Big[\binom{n}{m} 2^{l-1}\Big] = l - 1 + \log_2 \binom{n}{m} \leq (\gamma_{p,\infty} + 1) m - 1 + 3m \log_2 \Big(\frac{n}{m} + 1\Big).$$

Together with (17), we arrive at

$$k \geq \gamma_p m \log_2(n/m + 1) \quad \Longrightarrow \quad e_k(id : \ell_p^n(\mathbb{R}) \to \ell_\infty^n(\mathbb{R})) \leq C_p m^{-1/p},$$

where $\gamma_p = \gamma_{p,\infty} + 4$ and $C_p = 1 + C_{p,\infty}$ depend only on p.

Next, we define $\alpha_p > 2$ to be a real number large enough to ensure

$$\gamma_p \alpha_p + 1 \leq 2^{\alpha_p - 2}. \tag{18}$$

Moreover, we assume that k and n are positive integers with $n \geq k \geq 2\alpha_p \gamma_p \log_2 n$. This allows us to choose m to be any integer with

$$\frac{k}{2\alpha_p \gamma_p \log_2(n/k + 1)} \leq m \leq \frac{k}{\alpha_p \gamma_p \log_2(n/k + 1)}. \tag{19}$$

From $\gamma_p > 4$ and $\alpha_p > 2$, it follows that $m \le n$ and, by monotonicity of $f : t \to t\log_2(n/t+1)$ on $(0,\infty)$, we get

$$\begin{aligned}
\gamma_p f(m) &= \gamma_p m \log_2(n/m+1) \le \gamma_p f\Big(\frac{k}{\gamma_p\alpha_p\log_2(n/k+1)}\Big)\\
&= \gamma_p \cdot \frac{k}{\gamma_p\alpha_p\log_2(n/k+1)}\log_2\Big(\frac{n}{k}\gamma_p\alpha_p\log_2\Big(\frac{n}{k}+1\Big)+1\Big)\\
&\le \frac{k}{\alpha_p\log_2(n/k+1)}\Big[\log_2\Big(\frac{n}{k}+1\Big)+\log_2(\gamma_p\alpha_p+1)+\log_2\Big(\log_2\Big(\frac{n}{k}+1\Big)+1\Big)\Big]\\
&\le \frac{2k}{\alpha_p}+\frac{k\log_2(\gamma_p\alpha_p+1)}{\alpha_p\log_2(n/k+1)} \le \frac{k}{\alpha_p}\Big[2+\log_2(\gamma_p\alpha_p+1)\Big] \le k,
\end{aligned}$$

where we used (18) in the last step.

This finishes the proof of the upper bound in (2) for $q=\infty$ and $2\alpha_p\gamma_p\log_2 n \le k \le n$. The k's between $\log_2 n$ and $2\alpha_p\gamma_p\log_2 n$ can be incorporated by monotonicity at the cost of larger constants.

Finally, the case $0 < p < q < \infty$ follows by interpolation. Indeed, Lemma 8 implies that

$$e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \le 2^{1/\bar{p}} e_k^{1-\frac{p}{q}}(id : \ell_p^n(\mathbb{R}) \to \ell_\infty^n(\mathbb{R})),$$

and the result follows from the bound just proven for $e_k(id : \ell_p^n(\mathbb{R}) \to \ell_\infty^n(\mathbb{R}))$. □

4 Extensions

We collect further facts about the entropy numbers of identities between finite-dimensional spaces, which seem to be very well known in the community. Several parts of Theorem 2 can be proved in different ways and we present some of these alternative approaches. Finally, we extend Theorem 2 also to the complex setting.

4.1 Alternative use of volume arguments

Without much work, one can show that Lemma 7 is optimal up to the constant $4^{1/\bar{p}}$. Indeed, if X is a n-dimensional real vector spaces equipped with a (quasi)-norm and B_X is covered by 2^{k-1} translations of rB_X, then

$$\mathrm{vol}(B_X) \le 2^{k-1} r^n \,\mathrm{vol}(B_X).$$

This implies that $e_k(id : X \to X) \ge 2^{-\frac{k-1}{n}}$.

If $0 < p \le q \le \infty$, we can write

$$2^{-\frac{k-1}{n}} \le e_k(id : \ell_p^n(\mathbb{R}) \to \ell_p^n(\mathbb{R})) \le e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R})) \cdot \|id : \ell_q^n(\mathbb{R}) \to \ell_p^n(\mathbb{R})\|.$$

By Lemma 6, the last norm is equal to $n^{1/p-1/q}$ and we obtain the lower bound in (2) for $k \ge n$.

The upper bound on $e_k(id : \ell_p^n(\mathbb{R}) \to \ell_q^n(\mathbb{R}))$ for $0 < p < q \le \infty$ and $\log_2 n \le k \le n$ was proven in Section 3 first for $q=\infty$ and, afterwards, for $p < q < \infty$ by interpolation. Actually, one can adapt the original proof for $q=\infty$ also to the general setting of $q \le \infty$. Unfortunately, this comes at a price of further technical difficulties and we, together with [30], preferred to go through interpolation.

4.2 Alternative use of combinatorial arguments

Next comment considers the lower bound in (2) for $\log_2 n \le k \le n$. Lemma 9 constructs a large subset (namely $(2m)^{-1/p}A_m$) of B_p^n with the elements having large mutual distances in ℓ_q^n. There is, however, more ways to achieve a similar result. We present a statement, which is very well-known in coding theory, cf. [17, 29], and appeared also in [14, 15] and [2] in connection with Gelfand widths and optimality results of sparse recovery. There the reader can also find a short proof, which is very much in the spirit of the proof of Lemma 9.

Lemma 10. *Let $k \le n$ be two positive integers. Then there are N subsets $T_1, \dots, T_N$ of $\{1, \dots, n\}$, such that*

(i) $N \ge \left(\frac{n}{4k}\right)^{k/2}$,

(ii) $|T_i| = k$ *for all* $i = 1, \dots, N$ *and*

(iii) $|T_i \cap T_j| < k/2$ *for all* $i \ne j$.

Similarly to Lemma 9, Lemma 10 can be used to produce a large set of points in B_p^n with large distance in ℓ_q^n. It is enough to take $x_j = k^{-1/p}\chi_{T_j}$, $j = 1, \dots, N$, where T_j's are the sets from Lemma 10 and χ_{T_j} is the indicator vector of T_j.

4.3 Complex spaces

It is surprisingly easy to extend the estimates of Theorem 2 to the setting of complex spaces $\ell_p^n(\mathbb{C})$. The main result then reads as follows.

Theorem 11. *Let $0 < p, q \le \infty$ and let $n \in \mathbb{N}$.*

a) If $0 < p \le q \le \infty$ then for all $k \in \mathbb{N}$ it holds

$$e_k(id : \ell_p^n(\mathbb{C}) \to \ell_q^n(\mathbb{C})) \sim \begin{cases} 1 & \text{if} \quad 1 \le k \le \log_2(2n), \\ \left(\dfrac{\log_2(1+2n/k)}{k}\right)^{\frac{1}{p}-\frac{1}{q}} & \text{if} \quad \log_2(2n) \le k \le 2n, \\ 2^{-\frac{k-1}{2n}} n^{\frac{1}{q}-\frac{1}{p}} & \text{if} \quad 2n \le k. \end{cases} \tag{20}$$

b) If $0 < q \le p \le \infty$ then for all $k \in \mathbb{N}$ it holds

$$e_k(id : \ell_p^n(\mathbb{C}) \to \ell_q^n(\mathbb{C})) \sim 2^{-\frac{k-1}{2n}} n^{\frac{1}{q}-\frac{1}{p}}. \tag{21}$$

The constants of equivalence in both (20) *and* (21) *may depend on p and q, but are independent of k and n.*

Proof. We observe, that the mapping

$$\mathcal{J}(z) = \mathcal{J}(z_1, \dots, z_n) = (\mathrm{Re}(z_1), \mathrm{Im}(z_1), \dots, \mathrm{Re}(z_n), \mathrm{Im}(z_n))$$

is bounded from $\ell_p^n(\mathbb{C})$ to $\ell_p^{2n}(\mathbb{R})$ with the norm bounded by a quantity independent of n. The same is true about

$$\mathcal{J}'(x) = \mathcal{J}'(x_1, \dots, x_{2n}) = (x_1 + ix_2, \dots, x_{2n-1} + ix_{2n})$$

as a mapping from $\ell_p^{2n}(\mathbb{R})$ to $\ell_p^n(\mathbb{C})$.

Using the submultiplicativity of entropy numbers, we get

$$e_k(id : \ell_p^n(\mathbb{C}) \to \ell_q^n(\mathbb{C})) \leq \|\mathcal{J} : \ell_p^n(\mathbb{C}) \to \ell_p^{2n}(\mathbb{R})\| \\ \cdot e_k(id : \ell_p^{2n}(\mathbb{R}) \to \ell_q^{2n}(\mathbb{R})) \cdot \|\mathcal{J}' : \ell_q^{2n}(\mathbb{R}) \to \ell_q^n(\mathbb{C})\|$$

and

$$e_k(id : \ell_p^{2n}(\mathbb{R}) \to \ell_q^{2n}(\mathbb{R})) \leq \|\mathcal{J}' : \ell_p^{2n}(\mathbb{R}) \to \ell_p^n(\mathbb{C})\| \\ \cdot e_k(id : \ell_p^n(\mathbb{C}) \to \ell_q^n(\mathbb{C})) \cdot \|\mathcal{J} : \ell_q^n(\mathbb{C}) \to \ell_q^{2n}(\mathbb{R})\|.$$

This can be summarized into

$$e_k(id : \ell_p^n(\mathbb{C}) \to \ell_q^n(\mathbb{C})) \sim e_k(id : \ell_p^{2n}(\mathbb{R}) \to \ell_q^{2n}(\mathbb{R}))$$

with constants of equivalence independent on k and n. The result then follows from Theorem 2. □

Acknowledgement

The second author was supported by the grant P201/18/00580S of the Grant Agency of the Czech Republic and by the Neuron Fund for Support of Science.

References

[1] T. Aoki, *Locally bounded linear topological spaces*, Proc. Imp. Acad. Tokyo 18 (10) (1942), 588–594

[2] H. Boche, R. Calderbank, G. Kutyniok, and J. Vybíral, *A survey of compressed sensing*, in: Compressed Sensing and its Applications, Birkhäuser, Boston, 2015

[3] J. Bourgain, J. Lindenstrauss, and V. Milman, *Approximation of zonoids by zonotopes*, Acta Math. 162 (1) (1989), 73–141

[4] J. Bourgain, A. Pajor, S.J. Szarek, and N. Tomczak-Jaegermann, *On the duality problem for entropy numbers of operators*, in: Geometric aspects of functional analysis, Springer, Berlin, Heidelberg, 1989

[5] E.J. Candès, J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory 52 (2) (2006), 489–509

[6] B. Carl, *Entropy numbers, s-numbers, and eigenvalue problems*, J. Funct. Anal. 41 (3) (1981), 290–306

[7] B. Carl, *Entropy numbers of diagonal operators with an application to eigenvalue problems*, J. Approx. Theory 32 (1981), 135–150

[8] B. Carl and I. Stephani, Entropy, compactness and the approximation of operators, Cambridge University Press, Cambridge, 1990

[9] B. Carl and H. Triebel, *Inequalities between eigenvalues, entropy numbers, and related quantities of compact operators in Banach spaces*, Math. Ann. 251 (2) (1980), 129–133

[10] P.G.L. Dirichlet, *Sur une nouvelle méthode pour la détermination des intégrales multiples*, J. Math. Pures Appl. 4 (1839), 164–168

[11] D.L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (4) (2006), 1289–1306

[12] D.E. Edmunds and Yu. Netrusov, *Entropy numbers and interpolation*, Math. Ann. 351 (4) (2011), 963–977

[13] D.E. Edmunds and H. Triebel, Function spaces, entropy numbers and differential operators, Cambridge University Press, Cambridge, 1996

[14] S. Foucart, A. Pajor, H. Rauhut, and T. Ullrich, *The Gelfand widths of ℓ_p-balls for* $0 < p \leq 1$, J. Compl. 26 (6) (2010), 629–640

[15] S. Foucart and H. Rauhut, A mathematical introduction to compressive sensing, Birkhäuser, Boston, 2013

[16] Y. Gordon, H. König, and C. Schütt, *Geometric and probabilistic estimates for entropy and approximation numbers of operators*, J. Approx. Theory 49 (3) (1987), 219–239

[17] R. Graham and N. Sloane, *Lower bounds for constant weight codes*, IEEE Trans. Inform. Theor. 26 (1) (1980), 37–43

[18] O. Guédon and A.E. Litvak, *Euclidean projections of a p-convex body*, in: Geometric aspects of functional analysis, Springer, Berlin, Heidelberg, 2000

[19] A. Hinrichs, A. Kolleck, and J. Vybíral, *Carl's inequality for quasi-Banach spaces*, J. Funct. Anal. 271 (8) (2016), 2293–2307

[20] Y. Guo, P.L. Bartlett, J. Shawe-Taylor, and R.C. Williamson, *Covering numbers for support vector machines*, in: Proceedings of the twelfth annual conference on Computational learning theory (pp. 267-277), ACM (1999)

[21] K. Höllig, *Diameters of classes of smooth functions*, in: Quantitative Approximation (pp. 163–175), Academic Press, New York, London, 1980

[22] N. Kalton, *Quasi-Banach spaces*, in: Handbook of the geometry of Banach spaces, Vol. 2 (pp. 1099–1130), North-Holland, Amsterdam, 2003

[23] A. Kolmogorov, *On certain asymptotic characteristics of some completely bounded metric spaces*, Dokl. Akad. Nauk. SSSR 108 (1956), 385–389

[24] A. Kolmogorov and B. Tikhomirov, *ε-entropy and ε-capacity of sets in functional space*, Uspekhi Mat. Nauk 14 (2) (86) (1959), 3–86

[25] H. König, Eigenvalue distribution of compact operators, Birkhäuser, Basel, 1986

[26] T. Kühn, *A lower estimate for entropy numbers*, J. Approx. Theory 110 (1) (2001), 120–124

[27] M. Ledoux and M. Talagrand, Probability in Banach spaces: isoperimetry and processes, Springer, Berlin, 1991

[28] M.A. Lifshits and W. Linde, *Approximation and entropy numbers of Volterra operators with application to Brownian motion*, Memoirs Amer. Math. Soc. 157 (2002)

[29] N. Noam and W. Avi, *Hardness vs randomness*, J. Comput. Syst. Sci. 49 (2) (1994), 149–167

[30] A. Pietsch, Operator ideals, North-Holland, Amsterdam, 1980

[31] G. Pisier, The volume of convex bodies and Banach space geometry, Cambridge University Press, Cambridge, 1999

[32] G. Raskutti, M.J. Wainwright, and B. Yu, *Minimax rates of estimation for high-dimensional linear regression over ℓ_q-balls.* IEEE Trans. Inform. Theory 57 (10) (2011), 6976–6994

[33] S. Rolewicz, *On a certain class of linear metric spaces*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astrono. Phys. 5 (1957), 471–473

[34] C. Schütt, *Entropy numbers of diagonal operators between symmetric Banach spaces*, J. Approx. Theory 40 (2) (1984), 121–128

[35] E.M. Stein and R. Shakarchi, Complex analysis, Princeton Lectures in Analysis II, Princeton University Press, Princeton, NJ, 2003

[36] R. Williamson, A.J. Smola, and B. Schölkopf, *Entropy numbers, operators and support vector kernels*, in: Computational Learning Theory, Springer, Berlin, Heidelberg, 1999

[37] R.C. Williamson, A.J. Smola, and B. Schölkopf, *Generalization performance of regularization networks and support vector machines via entropy numbers of compact operators*, IEEE Trans. Inform. Theory 47 (6) (2001), 2516–2532

[38] D.-X. Zhou, *The covering number in learning theory*, J. Compl. 18 (3) (2002), 739–767