ON THE ENTROPY NUMBERS AND THE KOLMOGOROV WIDTHS

GUERGANA PETROVA AND PRZEMYSŁAW WOJTASZCZYK

ABSTRACT. Direct estimates between linear or nonlinear Kolmogorov widths and entropy numbers are presented. These estimates are derived using the recently introduced Lipschitz widths. Applications for m-term approximation are obtained.

1. INTRODUCTION

We consider a Banach space $(X, \|\cdot\|_X)$ (or a Hilbert space H) equipped with a norm $\|\cdot\|_X$ and a compact subset $\mathcal{K} \subset X$ of X. Typically, $\mathcal{K}$ is a finite ball in smoothness spaces like the Lipschitz, Sobolev, or Besov spaces.

A well known classical result, called the Carl's inequality, see [2] or [7], compares a certain characteristic of the set $\mathcal{K}$, called entropy numbers $e_k(\mathcal{K})_X$, with its approximability by linear spaces, measured by its Kolmogorov width $d_k(\mathcal{K})_X$. The Carl's inequality states that for each $r > 0$, there is a constant $C(r)$ such that for all $n \in \mathbb{N}$,

$$\max_{1\le k\le n} k^r e_k(\mathcal{K})_X \le C(r) \max_{1\le m\le n} m^r d_{m-1}(\mathcal{K})_X. \tag{1.1}$$

Inequality (1.1) has been generalized in [10], where the nonlinear Kolmogorov widths $d_n(\mathcal{K}, N)_X$ have been used instead of the linear Kolmogorov widths $d_k(\mathcal{K})_X$. More precisely, it has been shown there that for each $r > 0$, there is a constant $C(r, \lambda)$ such that for all $n \in \mathbb{N}$,

$$\max_{1\le k\le n} k^r e_k(\mathcal{K})_X \le C(r,\lambda) \max_{1\le m\le n} m^r d_{m-1}(\mathcal{K}, \lambda^m)_X, \tag{1.2}$$

with $\lambda > 1$ a fixed constant. In addition, it was also proven that for each $r > 0$, there is a constant $C(r, a)$ such that for all $n \in \mathbb{N}$,

$$\max_{1\le k\le n} k^r e_{(a+r)k\log k}(\mathcal{K})_X \le C(r,a) \max_{1\le m\le n} m^r d_{m-1}(\mathcal{K}, m^{am})_X, \tag{1.3}$$

where $a > 0$ is a fixed constant and $k \log k$ cannot be replaced by a slower growing function of k.

All these inequalities are primarily useful when the linear or nonlinear Kolmogorov widths decay as a power of m. In this paper, we give finer extensions of the (generalized) Carl's inequalities (1.1), (1.2) and (1.3), using

2010 *Mathematics Subject Classification.* 41A46, 41A65.

Key words and phrases. Kolmogorov widths, entropy numbers

Dedicated to Ron DeVore, with the utmost respect and admiration.

the recently introduced in [8] Lipschitz widths. We start with some definitions, presented in §2, and continue, see §3, with a comparison between the nonlinear Kolmogorov widths and the Lipschitz widths. Our main results are presented in §4, where we give a direct comparison between the entropy numbers of $\mathcal{K}$ and its linear and nonlinear Kolmogorov widths. Finally, in §5, we derive what these estimates mean for the m-term approximation in Hilbert spaces.

2. Preliminaries

We start this section with the definition of Kolmogorov widths. If we fix the value of $n \geq 0$, the *Kolmogorov n-width* $d_n(\mathcal{K})_X$ of $\mathcal{K}$ is defined as

$$d_0(\mathcal{K})_X := \sup_{f\in\mathcal{K}} \|f\|_X, \quad d_n(\mathcal{K})_X := \inf_{\dim(X_n)=n} \sup_{f\in\mathcal{K}} \operatorname{dist}(f, X_n)_X, \quad n \geq 1,$$

where the infimum is taken over all linear spaces $X_n \subset X$ of dimension n. These are the classical Kolmogorov widths introduced in [6], or consult [7] for their modern exposition. To distinguish them from the introduced later nonlinear Kolmogorov widths, we call them *linear Kolmogorov n-widths.* They describe the optimal performance possible for the approximation of the model class $\mathcal{K}$ using linear spaces of dimension n. However, they do not tell us how to select a (near) optimal space Y of dimension n for this purpose. Let us also note that in the definition of Kolmogorov width, we are not requiring that the mapping which sends $f \in \mathcal{K}$ into an approximation to f is a linear map.

A generalization of this concept was introduced in [10], where the so called *nonlinear Kolmogorov (n, N)-width* $d_n(\mathcal{K}, N)_X$ was defined for $N \geq 1$ as

$$d_0(\mathcal{K}, N)_X := \sup_{f\in K} \|f\|_X,$$

$$d_n(\mathcal{K}, N)_X := \inf_{\mathcal{L}_N} \sup_{f\in\mathcal{K}} \inf_{X_n\in\mathcal{L}_N} \operatorname{dist}(f, X_n)_X, \quad n \geq 1,$$

where the last infimum is over the sets $\mathcal{L}_N$ of at most N linear spaces $X_n \subset X$ of dimension n. Note that here the choice of the linear subspace $X_n \in \mathcal{L}_N$ from which we choose the best approximation to f depends on f. Clearly, $d_n(\mathcal{K}, 1)_X = d_n(\mathcal{K})_X$, and the bigger the N is, the more flexibility we have to approximate f. These nonlinear Kolmogorov widths are used in estimating from below the best m-term approximation, see e.g. [3, 10]. The cases considered in [10] are the cases when $N = \lambda^n$, and $N = n^{an}$, where $\lambda > 1$ and $a > 0$ are fixed constants, respectively. A useful observation that we are going to utilize is that both Kolmogorov widths are homogenous. Namely, if $\mathcal{K} \subset X$ and $t \in \mathbb{R}$, we have

$$d_n(t\mathcal{K}, N)_X = |t| d_n(\mathcal{K}, N)_X \quad \text{and} \quad d_n(t\mathcal{K})_X = |t| d_n(\mathcal{K})_X, \tag{2.1}$$

where $t\mathcal{K} := \{tf : f \in \mathcal{K}\}$.

In going further, we introduce first the *minimal ϵ-covering number* $N_\epsilon(\mathcal{K})$ of a compact set $\mathcal{K} \subset X$. A collection $\{g_1, \ldots, g_m\} \subset X$ of elements of X is called an ϵ-covering of $\mathcal{K}$ if

$$\mathcal{K} \subset \bigcup_{j=1}^{m} B(g_j, \epsilon), \quad \text{where} \quad B(g_j, \epsilon) := \{f \in X : \|f - g_j\|_X \le \epsilon\}.$$

An ϵ-covering of $\mathcal{K}$ whose cardinality is minimal is called *minimal ϵ-covering* of $\mathcal{K}$. We denote by $N_\epsilon(\mathcal{K})$ the cardinality of the minimal ϵ-covering of $\mathcal{K}$. *Minimal inner ϵ-covering number* $\tilde{N}_\epsilon(\mathcal{K})$ of a compact set $\mathcal{K} \subset X$ is defined exactly as $N_\epsilon(\mathcal{K})$ but we additionally require that the centers $\{g_1, \ldots, g_m\}$ of the covering are elements from $\mathcal{K}$.

Entropy numbers $e_n(\mathcal{K})_X$, $n \ge 0$, of the compact set $\mathcal{K} \subset X$ are defined as the infimum of all $\epsilon > 0$ for which 2^n balls with centers from X and radius ϵ cover $\mathcal{K}$. If we put the additional restriction that the centers of these balls are from $\mathcal{K}$, then we define the so called *inner entropy numbers* $\tilde{e}_n(\mathcal{K})_X$. Formally, we write

$$e_n(\mathcal{K})_X = \inf\{\epsilon > 0 \ : \ \mathcal{K} \subset \bigcup_{j=1}^{2^n} B(g_j, \epsilon), \ g_j \in X, \ j = 1, \ldots, 2^n\},$$

$$\tilde{e}_n(\mathcal{K})_X = \inf\{\epsilon > 0 \ : \ \mathcal{K} \subset \bigcup_{j=1}^{2^n} B(h_j, \epsilon), \ h_j \in \mathcal{K}, \ j = 1, \ldots, 2^n\}.$$

A collection $\{f_1, \ldots, f_\ell\} \subset \mathcal{K}$ of elements from $\mathcal{K}$ is called an ϵ-packing of $\mathcal{K}$ if

$$\min_{i \ne j} \|f_i - f_j\|_X > \epsilon.$$

An ϵ-packing of $\mathcal{K}$ whose size is maximal is called *maximal ϵ-packing* of $\mathcal{K}$. We denote by $\tilde{P}_\epsilon(\mathcal{K})$ the cardinality of the maximal ϵ-packing of $\mathcal{K}$. We have the following inequalities for every $\epsilon > 0$ and every compact set $\mathcal{K}$

$$\tilde{P}_\epsilon(\mathcal{K}) \ge \tilde{N}_\epsilon(\mathcal{K}) \ge \tilde{P}_{2\epsilon}(\mathcal{K}), \tag{2.2}$$

and

$$e_n(\mathcal{K})_X \le \tilde{e}_n(\mathcal{K})_X \le 2e_n(\mathcal{K})_X. \tag{2.3}$$

Finally, we introduce the *Lipschitz widths* $d_n^\gamma(\mathcal{K})_X$, $\gamma \ge 0$, $n \ge 1$, of the compact set $\mathcal{K} \subset X$, see [8]. We denote by $(\mathbb{R}^n, \|.\|_{Y_n})$, $n \ge 1$, the n-dimensional Banach space with a fixed norm $\|\cdot\|_{Y_n}$. For $\gamma \ge 0$, we first define the *fixed Lipschitz* width $d^\gamma(\mathcal{K}, Y_n)_X$,

$$d^\gamma(\mathcal{K}, Y_n)_X := \inf_{\Phi_n} \sup_{f \in \mathcal{K}} \inf_{y \in B_{Y_n}} \|f - \Phi_n(y)\|_X,$$

where the infimum is taken over all Lipschitz mappings

$$\Phi_n : (B_{Y_n}, \|\cdot\|_{Y_n}) \to X, \quad B_{Y_n} := \{y \in \mathbb{R}^n : \ \|y\|_{Y_n} \le 1\},$$

that satisfy the Lipschitz condition

$$\sup_{y,y'\in B_{Y_n}} \frac{\|\Phi_n(y)-\Phi_n(y')\|_X}{\|y-y'\|_{Y_n}} \leq \gamma,$$

with constant γ. We then define the *Lipschitz* width

$$d_n^\gamma(\mathcal{K})_X := \inf_{k\leq n}\inf_{\|\cdot\|_{Y_k}} d^\gamma(\mathcal{K},Y_k)_X,$$

where the infimum is taken over all norms $\|\cdot\|_{Y_k}$ in $\mathbb{R}^k$ and all $k\leq n$. We observe the following analog to (2.1)

$$|t|d_n^{\gamma|t|}(t\mathcal{K})_X = d_n^\gamma(\mathcal{K})_X, \text{ where } t\mathcal{K} := \{tf : f\in\mathcal{K}\}. \tag{2.4}$$

3. Comparison between nonlinear Kolmogorov widths and Lipschitz widths

In this section, we derive direct inequalities between the nonlinear Kolmogorov widths and the Lipschitz widths. We then use known relations between entropy numbers and Lipschitz widths to derive improvements of the (generalized) Carl's inequalities.

We first note the following comparison between the linear Kolmogorov widths and the Lipschitz widths, proven in [8], see Corollary 5.2.

Theorem 3.1. *For every $n\geq 1$ and every compact set $\mathcal{K}\subset X$ we have*

$$d_n^\gamma(\mathcal{K})_X \leq d_n(\mathcal{K})_X, \quad \textit{for every } \gamma \geq 2\sup_{f\in\mathcal{K}}\|f\|_X.$$

We next proceed with estimates between the nonlinear Kolmogorov width and the Lipschitz widths. Clearly, it follows from the definition that

$$d_n(\mathcal{K},N)_X \geq d_{nN}(\mathcal{K})_X \geq d_{nN}^\gamma(\mathcal{K})_X, \quad \gamma = 2\sup_{f\in\mathcal{K}}\|f\|,$$

where we have used in the last inequality the above theorem. Better estimates in the case of $\mathcal{K}$ being a subset of a Hilbert space H or a general Banach space X are described in the following lemmas.

Lemma 3.2. *For every $n\geq 1$, $N>1$, and every compact $\mathcal{K}$, subset of a Hilbert space H such that $\sup_{f\in\mathcal{K}}\|f\|_H=1$, we have*

$$d_{n+1}^{(N+1)}(\mathcal{K})_H \leq d_n(\mathcal{K},N)_H, \quad \textit{and} \quad d^3_{n+\lceil\log_2 N\rceil}(\mathcal{K})_H \leq d_n(\mathcal{K},N)_H. \tag{3.1}$$

Proof: Let us fix $n,N\geq 1$, and consider the n-dimensional linear spaces $X_1,\dots,X_N$, $X_i\subset H$, $i=1,\dots,N$. We define a norm $\|\cdot\|_{Y_{n+1}}$ on $\mathbb{R}^{n+1}$,

$$\|(x,x_{n+1})\|_{Y_{n+1}} := \max\left\{\|x\|_{\ell_2(\mathbb{R}^n)}, |x_{n+1}|\right\}, \quad x:=(x_1,\dots,x_n),$$

whose unit ball is

$$B_{Y_{n+1}} := \{(x,x_{n+1}) \ : \ \|x\|_{\ell_2(\mathbb{R}^n)}\leq 1 \text{ and } |x_{n+1}|\leq 1\}.$$

Clearly

$$B_{Y_{n+1}} = B_{\ell_2(\mathbb{R}^n)}\times[-1,1], \quad \text{where} \quad B_{\ell_2(\mathbb{R}^n)} := \{x\in\mathbb{R}^n \ : \ \|x\|_{\ell_2(\mathbb{R}^n)}\leq 1\}.$$

We want to construct a Lipschitz mapping from $(B_{Y_{n+1}}, \|\cdot\|_{Y_{n+1}})$ to H whose image approximates well $\mathcal{K}$. We divide the interval $[-1,1]$ into N subintervals I_j, $j=0,\dots,N-1$,

$$I_j := [a_j, a_{j+1}], \quad a_j := 2j/N - 1,$$

with centers c_j and consider the univariate continuous piecewise linear functions ψ_j, $\psi_j : ([-1,1], |\cdot|) \to [0,1]$, $j=0,\dots,N-1$, whose break points are $\{a_0,\dots,a_j,c_j,a_{j+1},\dots,a_{N-1}\}$, and

$$\psi_j(c_j) = 1, \quad \psi_j(a_k) = 0, \quad k = 0,\dots,N-1.$$

Let $(B_{X_j}, \|\cdot\|_H)$ be the unit ball of the space $X_j \subset H$. We fix an orthonormal basis $\{\varphi_1^j,\dots,\varphi_n^j\}$ in X_j and consider the isometry map $\bar{\psi}_j$ from $B_{\ell_2(\mathbb{R}^n)}$ onto B_{X_j},

$$\bar{\psi}_j : (B_{\ell_2(\mathbb{R}^n)}, \|\cdot\|_{\ell_2(\mathbb{R}^n)}) \to (B_{X_j}, \|\cdot\|_H),$$

defined as

$$\bar{\psi}_j(x) = \bar{\psi}_j(x_1,\dots,x_n) := \sum_{i=1}^{n} x_i \varphi_i^j. \tag{3.2}$$

We use these mappings to construct $\Phi_{n+1} : (B_{Y_{n+1}}, \|\cdot\|_{Y_{n+1}}) \to H$ as

$$\Phi_{n+1}(x, x_{n+1}) := \sum_{j=0}^{N-1} \psi_j(x_{n+1}) \cdot \bar{\psi}_j(x).$$

Let us fix $(x, x_{n+1}), (x', x'_{n+1}) \in B_{Y_{n+1}}$ and denote by

$$A := \|\Phi_{n+1}(x, x_{n+1}) - \Phi_{n+1}(x', x'_{n+1})\|_H.$$

We want to derive an upper bound for A. Note that $\psi_j(x_{n+1}) \neq 0$ if and only if $x_{n+1} \in I_j$.We consider the following two cases:

- if $x_{n+1}, x'_{n+1} \in I_j$ for some $j = 0,\dots,N-1$, then $\psi_j(x_{n+1}) \neq 0$, $\psi_j(x'_{n+1}) \neq 0$, $\psi_k(x_{n+1}) = \psi_k(x'_{n+1}) = 0$ for all $k \neq j$, and therefore

$$\begin{aligned} A &= \|\psi_j(x_{n+1})\bar{\psi}_j(x) - \psi_j(x'_{n+1})\bar{\psi}_j(x')\|_H \\ &\leq |\psi_j(x_{n+1})|\|\bar{\psi}_j(x) - \bar{\psi}_j(x')\|_H \\ &+ |\psi_j(x_{n+1}) - \psi_j(x'_{n+1})|\|\bar{\psi}_j(x')\|_H \\ &\leq \|x - x'\|_{\ell_2(\mathbb{R}^n)} + N|x_{n+1} - x'_{n+1}| \\ &\leq (N+1)\|(x, x_{n+1}) - (x', x'_{n+1})\|_{Y_{n+1}}. \end{aligned}$$

- if $x_{n+1} \in I_j, x'_{n+1} \in I_k$ for some $j, k = 0,\dots,N-1$, $k \neq j$, we obtain that

$$A = \|\psi_j(x_{n+1})\bar{\psi}_j(x) - \psi_k(x'_{n+1})\bar{\psi}_k(x')\|_H.$$

  We can assume without loss of generality that

$$x_{n+1} \leq a_{j+1} \leq a_k \leq x'_{n+1}.$$

Since $\psi_j(a_{j+1})=\psi_k(a_k)=0$, we have

$$\begin{aligned}
A &\le \|\psi_j(x_{n+1})\bar\psi_j(x)-\psi_j(a_{j+1})\bar\psi_j(x)\|_H\\
&+ \|\psi_k(a_k)\bar\psi_k(x)-\psi_k(x'_{n+1})\bar\psi_k(x')\|_H\\
&\le |\psi_j(x_{n+1})-\psi_j(a_{j+1})|\|\bar\psi_j(x)\|_H\\
&+ \|\psi_k(a_k)\bar\psi_k(x)-\psi_k(x'_{n+1})\bar\psi_k(x')\|_H\\
&\le N|a_{j+1}-x_{n+1}|+\|x-x'\|_{\ell_2(\mathbb{R}^n)}+N|x'_{n+1}-a_k|\\
&\le N|x'_{n+1}-x_{n+1}|+\|x-x'\|_{\ell_2(\mathbb{R}^n)}\\
&\le (N+1)\|(x,x_{n+1})-(x',x'_{n+1})\|_{Y_{n+1}},
\end{aligned}$$

where we have used arguments similar to the first case.

In both cases we have that

$$\|\Phi_{n+1}(x,x_{n+1})-\Phi_{n+1}(x',x'_{n+1})\|_H\le (N+1)\|(x,x_{n+1})-(x',x'_{n+1})\|_{Y_{n+1}},$$

and therefore Φ_{n+1} is an $(N+1)$-Lipschitz mapping.

Since $\sup_{f\in\mathcal{K}}\|f\|_H=1$, the approximant f_j to f from X_j will belong to B_{X_j} since f_j is the orthogonal projection of f onto X_j. Thus, it follows from the definition of $\bar\psi_j$ that there is $x^j\in B_{\ell_2(\mathbb{R}^n)}$, such that $\bar\psi_j(x^j)=f_j$, and therefore

$$\Phi_{n+1}(x^j,c_j)=f_j,\quad\text{and}\quad \|f-f_j\|_H=\mathrm{dist}(f,X_j)_H,$$

which gives

$$d_{n+1}^{(N+1)}(\mathcal{K})_H\le d_n(\mathcal{K},N)_H.$$

To show the second part of (3.1), we determine $\ell\in\mathbb{N}$ such that

$$2^{\ell-1}<N\le 2^\ell,$$

and define a norm $\|\cdot\|_{Y_{n+\ell}}$ on $\mathbb{R}^{n+\ell}$ by

$$\|(x,y)\|_{Y_{n+\ell}}:=\max\left\{\|x\|_{\ell_2(\mathbb{R}^n)},\|y\|_{\ell_\infty(\mathbb{R}^\ell)}\right\},$$

where

$$x:=(x_1,\dots,x_n),\quad y:=(y_1,\dots,y_\ell).$$

The unit ball with respect to this norm is

$$B_{Y_{n+\ell}}:=\{(x,y)\in\mathbb{R}^{n+\ell}\ :\ \|x\|_{\ell_2(\mathbb{R}^n)}\le 1 \text{ and } \|y\|_{\ell_\infty(\mathbb{R}^\ell)}\le 1\}.$$

Like before, we have $B_{Y_{n+\ell}}=B_{\ell_2(\mathbb{R}^n)}\times[-1,1]^\ell$. Next, we consider the disjoint cubes Q_j, $j=1,\dots,2^\ell$, of side length 1 such that

$$[-1,1]^\ell=\cup_{j=1}^{2^\ell}Q_j.$$

We denote by $\mathbf{c}_j:=(c_1^j,\dots,c_\ell^j)\in\mathbb{R}^\ell$ the center of Q_j, $j=1,\dots,2^\ell$, and define the functions $\phi_j:([-1,1]^\ell,\|\cdot\|_{\ell_\infty(\mathbb{R}^\ell)})\to[0,1]$ as

$$\phi_j(y):=2\left(\frac12-\|\mathbf{c}_j-y\|_{\ell_\infty(\mathbb{R}^\ell)}\right)_+,\quad j=1,\dots,2^\ell,$$

and $\Psi_{n+\ell} : (B_{Y_{n+\ell}}, \|\cdot\|_{Y_{n+\ell}}) \to H$ as

$$\Psi_{n+\ell}(x,y) := \sum_{j=1}^{2^\ell} \phi_j(y) \cdot \bar{\psi}_j(x),$$

where $\bar{\psi}_j$ are the mappings defined in (3.2).

Using the fact that for any two numbers a, b, we have $|a_+ - b_+| \le |a-b|$, we obtain that

$$|\phi_j(y) - \phi_j(y')| \le 2|\|\mathbf{c}_j - y\|_{\ell_\infty(\mathbb{R}^\ell)} - \|\mathbf{c}_j - y'\|_{\ell_\infty(\mathbb{R}^\ell)}| \le 2\|y - y'\|_{\ell_\infty(\mathbb{R}^\ell)}.$$

Moreover, the supports of the ϕ_j's are disjoint, with Q_j being the support of ϕ_j, and $|\phi_j(y)| \le 1$ for all j. Now, following similar arguments as the ones for Φ_{n+1}, and denoting

$$B := \|\Psi_{n+\ell}(x,y) - \Psi_{n+\ell}(x',y')\|_H,$$

we derive that:

- if $y, y' \in Q_j$ for some $j = 1, \dots, 2^\ell$,

$$B = \|\phi_j(y)\bar{\psi}_j(x) - \phi_j(y')\bar{\psi}_j(x')\|_H \le 3\|(x,y) - (x',y')\|_{Y_{n+\ell}}.$$

- if $y \in Q_j$ and $y' \in Q_k$, $k \ne j$, we consider the line segment

$$y + t(y' - y), \quad 0 \le t \le 1,$$

and fix

$$d_j := y + t_0(y' - y) \in \partial Q_j,$$

and

$$b_k := y + t_1(y' - y) \in \partial Q_k.$$

Clearly $t_0 \le t_1$, $\phi_j(d_j) = \phi_k(b_k) = 0$,

$$\|y - d_j\|_{\ell_\infty(\mathbb{R}^\ell)} + \|y' - b_k\|_{\ell_\infty(\mathbb{R}^\ell)} = (t_0 + 1 - t_1)\|y - y'\|_{\ell_\infty(\mathbb{R}^\ell)} \le \|y - y'\|_{\ell_\infty(\mathbb{R}^\ell)},$$

and similarly to the estimate for A, one obtains

$$\begin{aligned}
B &= \|\phi_j(y)\bar{\psi}_j(x) - \phi_k(y')\bar{\psi}_k(x')\|_H \\
&\le |\phi_j(y) - \phi_j(d_j)|\|\bar{\psi}_j(x)\|_H + \|\phi_k(b_k)\bar{\psi}_k(x) - \phi_k(y')\bar{\psi}_k(x')\|_H \\
&\le 2\|d_j - y\|_{\ell_\infty(\mathbb{R}^\ell)} + \|x - x'\|_{\ell_2(\mathbb{R}^n)} + 2\|y' - b_k\|_{\ell_\infty(\mathbb{R}^\ell)} \\
&\le 2\|y - y'\|_{\ell_\infty(\mathbb{R}^\ell)} + \|x - x'\|_{\ell_2(\mathbb{R}^n)} \\
&\le 3\|(x,y) - (x',y')\|_{Y_{n+\ell}}.
\end{aligned}$$

Therefore, $\Psi_{n+\ell}$ is a 3-Lipschitz mapping. As before, since $\sup_{f\in\mathcal{K}} \|f\|_H = 1$, we obtain

$$d^3_{n+\lceil \log_2 N\rceil}(\mathcal{K})_H \le d_n(\mathcal{K}, N)_H,$$

where we have used the fact that $\ell = \lceil \log_2 N\rceil$ and $\phi_j(\mathbf{c}_j) = 1$, $j = 0, \dots, N$. The proof is completed. $\square$

The case of arbitrary Banach space X is based on the following lemma.

Lemma 3.3. *Let Y be an n-dimensional subspace of a Banach space X and $(B_Y, \|\cdot\|_Y)$ be its unit ball. Let $(B_Z, \|\cdot\|_H)$ be the unit ball in an n-dimensional subspace Z of a Hilbert space H. Then, there exists a linear map*

$$\bar{\psi} : (B_Z, \|\cdot\|_H) \to Y,$$

with Lipschitz constant (i.e. norm) at most $\sqrt{n}$ such that $B_Y \subset \bar{\psi}(B_Z)$. In addition, if $X = L_p$, then the Lipschitz constant of $\bar{\psi}$ is at most $n^{|1/2-1/p|}$.

Proof: It follows from the Fritz John theorem, see Chapter 3 in [9] or [1], that there exists an invertible linear operator $\phi : (\mathbb{R}^n, \|\cdot\|_{\ell_2(\mathbb{R}^n)}) \to Y$ onto Y such that

$$\phi(B_{\ell_2(\mathbb{R}^n)}) \subset B_Y \subset \sqrt{n}\phi(B_{\ell_2(\mathbb{R}^n)}). \tag{3.3}$$

Let us fix an orthonormal basis $\varphi_1, \dots, \varphi_n$ for Z and consider the coordinate mapping $\kappa_Z : Z \to \mathbb{R}^n$ defined as

$$\kappa_Z(g) = (x_1, \dots, x_n) = x, \quad \text{where} \quad g = \sum_{j=1}^n x_j \varphi_j.$$

This mapping is isometry when $\mathbb{R}^n$ is equipped with the norm

$$\|x\|_{\ell_2(\mathbb{R}^n)} = \sqrt{\sum_{j=1}^n x_j^2} = \|g\|_Z.$$

We now define the linear mapping

$$\tilde{\psi} := \phi \circ \kappa_Z : (Z, \|\cdot\|_H) \to Y,$$

and notice that

$$\tilde{\psi}(B_Z) \subset B_Y \subset \sqrt{n}\tilde{\psi}(B_Z).$$

The first inclusion gives that $\tilde{\psi}$ has a norm (Lipschitz constant) ≤ 1, and thus $\bar{\psi} := \sqrt{n}\tilde{\psi}$ has a Lipschitz constant $\sqrt{n}$. The second inclusion shows that $B_Y \subset \bar{\psi}(B_Z)$, and therefore $\bar{\psi}$ is the desired mapping. It follows from [5, Cor. 5] that in the case of $X = L_p$, we can replace $\sqrt{n}$ in (3.3) by $n^{|1/2-1/p|}$. $\square$

Remark 3.4. Note that since $\bar{\psi}$ is linear, we have that $\bar{\psi}(0) = 0$, and for every $z \in B_Z$,

$$\|\bar{\psi}(z)\|_Y = \|\bar{\psi}(z) - \bar{\psi}(0)\|_Y \leq \sqrt{n}\|z\|_H \leq \sqrt{n}, \tag{3.4}$$

where we can replace $\sqrt{n}$ by $n^{|1/2-1/p|}$ in the case when $X = L_p$.

Lemma 3.5. *For every $n \geq 1$, $N > 1$, and every compact set $\mathcal{K}$ subset of a Banach space X with $\sup_{f \in \mathcal{K}} \|f\|_X = 1$, we have*

$$d_{n+1}^{2(N+1)\sqrt{n}}(\mathcal{K})_X \leq d_n(\mathcal{K}, N)_X, \quad \text{and} \quad d_{n+\lceil \log_2 N \rceil}^{6\sqrt{n}}(\mathcal{K})_X \leq d_n(\mathcal{K}, N)_X. \tag{3.5}$$

When $X = L_p$, we have

$$d_{n+1}^{2(N+1)n^{|1/2-1/p|}}(\mathcal{K})_{L_p} \le d_n(\mathcal{K},N)_{L_p}, \quad \text{and} \quad d_{n+\lceil \log_2 N\rceil}^{6n^{|1/2-1/p|}}(\mathcal{K})_{L_p} \le d_n(\mathcal{K},N)_{L_p}.$$

Proof: We fix n, $N > 1$, and consider the n dimensional linear spaces $X_1,\dots,X_N$, $X_j \subset X$, $j = 1,\dots,N$, with $(B_{X_j}, \|\cdot\|_X)$ being the unit ball of X_j. For a fixed $j = 1,\dots,N$, we apply Lemma 3.3 with $Y = X_j$ and $Z = \ell_2(\mathbb{R}^n)$ to find an M-Lipschitz mapping $\bar\Psi_j$, where $M = \sqrt{n}$ or $n^{|1/p-1/2|}$, depending on whether X is a general Banach space or L_p, such that

$$\bar\Psi_j : (B_{\ell_2(\mathbb{R}^n)}, \|\cdot\|_{\ell_2(\mathbb{R}^n)}) \to X_j, \quad \text{and} \quad B_{X_j} \subset \bar\Psi_j(B_{\ell_2(\mathbb{R}^n)}). \tag{3.6}$$

We show (3.5) by proceeding as in the proof of Lemma 3.2 and defining a mapping $\Theta_{n+1} : (B_{Y_{n+1}}, \|\cdot\|_{Y_{n+1}}) \to X$ as

$$\Theta_{n+1}(x, x_{n+1}) := 2\sum_{j=0}^{N-1} \psi_j(x_{n+1}) \cdot \bar\Psi_j(x),$$

where ψ_j and $(B_{Y_{n+1}}, \|\cdot\|_{Y_{n+1}})$ are as in Lemma 3.2. We fix (x, x_{n+1}), (x', x'_{n+1}), denote by

$$C := \|\Theta_{n+1}(x, x_{n+1}) - \Theta_{n+1}(x', x'_{n+1})\|_X,$$

and show in a similar way that

- if $x_{n+1}, x'_{n+1} \in I_j$ for some $j = 0,\dots,N-1$,

$$\begin{aligned}
\frac{C}{2} &= \|\psi_j(x_{n+1})\bar\Psi_j(x) - \psi_j(x'_{n+1})\bar\Psi_j(x')\|_X \\
&\le |\psi_j(x_{n+1})|\|\bar\Psi_j(x) - \bar\Psi_j(x')\|_X \\
&+ |\psi_j(x_{n+1}) - \psi_j(x'_{n+1})|\|\bar\Psi_j(x')\|_X \\
&\le M\|x - x'\|_{\ell_2(\mathbb{R}^n)} + NM|x_{n+1} - x'_{n+1}| \\
&\le M(N+1)\|(x, x_{n+1}) - (x', x'_{n+1})\|_{Y_{n+1}},
\end{aligned}$$

  where we have used (3.4).
- if $x_{n+1} \in I_j, x'_{n+1} \in I_k$ for some $j,k = 0,\dots,N-1$, $k \ne j$,

$$\begin{aligned}
\frac{C}{2} &\le \|\psi_j(x_{n+1})\bar\Psi_j(x) - \psi_j(a_{j+1})\bar\Psi_j(x)\|_X \\
&+ \|\psi_k(a_k)\bar\Psi_k(x) - \psi_k(x'_{n+1})\bar\Psi_k(x')\|_X \\
&\le |\psi_j(x_{n+1}) - \psi_j(a_{j+1})|\|\bar\Psi_j(x)\|_X \\
&+ \|\psi_k(a_k)\bar\Psi_k(x) - \psi_k(x'_{n+1})\bar\Psi_k(x')\|_X \\
&\le NM|a_{j+1} - x_{n+1}| + M\|x - x'\|_{\ell_2(\mathbb{R}^n)} + NM|x'_{n+1} - a_k| \\
&\le NM|x'_{n+1} - x_{n+1}| + M\|x - x'\|_{\ell_2(\mathbb{R}^n)} \\
&\le M(N+1)\|(x, x_{n+1}) - (x', x'_{n+1})\|_{Y_{n+1}}.
\end{aligned}$$

In conclusion,

$$\|\Theta_{n+1}(x, x_{n+1}) - \Theta_{n+1}(x', x'_{n+1})\|_H \le 2M(N+1)\|(x, x_{n+1}) - (x', x'_{n+1})\|_{Y_{n+1}},$$

and therefore Θ_{n+1} is a $2M(N+1)$-Lipschitz mapping.

Note that if f_j is the approximant to f from X_j, then

$$\|f-f_j\|_X \le \|f\|_X \quad \Rightarrow \quad \|f_j\|_X \le \|f-f_j\|_X + \|f\|_X \le 2\|f\|_X \le 2, \tag{3.7}$$

where we have used that $\sup_{f\in\mathcal{K}}\|f\|_X = 1$. Thus $f_j \in 2B_{X_j}$. It follows from Lemma 3.3 that since $B_{X_j} \subset \bar{\Psi}_j(B_{\ell_2(\mathbb{R}^n)})$, there is $x^j \in B_{\ell_2(\mathbb{R}^n)}$, such that $\bar{\Psi}_j(x^j) = \frac{1}{2}f_j$. Therefore

$$\Theta_{n+1}(x^j, c_j) = f_j, \quad \text{and} \quad \|f-f_j\|_X = \operatorname{dist}(f, X_j)_X,$$

which gives

$$d_{n+1}^{2M(N+1)}(\mathcal{K})_X \le d_n(\mathcal{K}, N)_X.$$

To show the second part of (3.5), we define $\Xi_{n+\ell} : (B_{Y_{n+\ell}}, \|\cdot\|_{Y_{n+\ell}}) \to X$ as

$$\Xi_{n+\ell}(x,y) := 2\sum_{j=1}^{2^\ell} \phi_j(y)\cdot \bar{\Psi}_j(x),$$

where ϕ_j and $(B_{Y_{n+\ell}}, \|\cdot\|_{Y_{n+\ell}})$ are the same as in Lemma 3.2 and $\bar{\Psi}_j$ is defined in (3.6). For fixed $(x,y),(x',y') \in B_{Y_{n+\ell}}$, we denote by

$$D := \|\Xi_{n+\ell}(x,y) - \Xi_{n+\ell}(x',y')\|_X$$

and consider the following cases

- if $y, y' \in Q_j$ for some $j = 1, \dots, 2^\ell$, we have
$$\frac{D}{2} \le 3M\|(x,y)-(x',y')\|_{Y_{n+\ell}}.$$
- if $y \in Q_j$ and $y' \in Q_k$, $k \ne j$, similarly to the estimate for C, we obtain

$$\begin{aligned}
\frac{D}{2} &= \|\phi_j(y)\bar{\Psi}_j(x) - \phi_k(y')\bar{\Psi}_k(x')\|_X \\
&\le |\phi_j(y) - \phi_j(d_j)|\|\bar{\Psi}_j(x)\|_X + \|\phi_k(b_k)\bar{\psi}_k(x) - \phi_k(y')\bar{\Psi}_k(x')\|_X \\
&\le 2M\|d_j - y\|_{\ell_\infty(\mathbb{R}^\ell)} + M\|x-x'\|_{\ell_2(\mathbb{R}^n)} + 2M\|y'-b_k\|_{\ell_\infty(\mathbb{R}^\ell)} \\
&\le 2M\|y-y'\|_{\ell_\infty(\mathbb{R}^\ell)} + M\|x-x'\|_{\ell_2(\mathbb{R}^n)} \\
&\le 3M\|(x,y)-(x',y')\|_{Y_{n+\ell}}.
\end{aligned}$$

The latter estimate implies that $\Xi_{n+\ell}$ is a $6M$-Lipschitz mapping, and since $\sup_{f\in\mathcal{K}}\|f\|_X = 1$, we obtain

$$d^{6M}_{n+\lceil \log_2 N\rceil}(\mathcal{K})_X \le d_n(\mathcal{K}, N)_X.$$

The proof is completed. □

Remark 3.6. Note that Lemma 3.5 with $X = L_2$ can be used instead of Lemma 3.2. However, we have decided to present both lemmas since better Lipschitz constants are obtained when working directly with a Hilbert space H.

Remark 3.7. It follows from (2.1) and (2.4) that lemmas similar to Lemma 3.2 and Lemma 3.5 can be stated in the case when $\sup_{f\in\mathcal{K}}\|f\|_H \neq 1$, or $\sup_{f\in\mathcal{K}}\|f\|_X \neq 1$, respectively.

4. Main results

In this section, we provide estimates from above and below that connect the behavior of the linear and nonlinear Kolmogorov widths of $\mathcal{K}$ with its entropy numbers. In what follows we assume that $\sup_{f\in\mathcal{K}}\|f\|_H = 1$ in the case of Hilbert space, or $\sup_{f\in\mathcal{K}}\|f\|_X = 1$ in the case of a general Banach space. Similar results hold if this supremum is not 1.

Our approach of deriving estimates from below utilizes some known results for Lipschitz widths stated below, see Theorem 4.7 in [8].

Theorem 4.1. *Let $\mathcal{K} \subset X$ be a compact subset of a Banach space X, $n \in N$, and $d_n^\gamma(\mathcal{K})_X$ be the Lipschitz width for $\mathcal{K}$ with Lipschitz constant $\gamma \geq 2\mathrm{rad}(\mathcal{K})$. Then the following holds:*

(1) *If for $\alpha > 0$, $\beta \in R$ and a constant $C > 0$, we have*

$$e_n(\mathcal{K})_X \geq C\frac{[\log_2 n]^\beta}{n^\alpha}, \quad n = 1, 2, \dots, \qquad \textit{then} \quad d_n^\gamma(\mathcal{K})_X \geq C'\frac{[\log_2 n]^\beta}{n^\alpha[\log_2 n]^\alpha},$$

for $n = 1, 2, \dots$, where $C' > 0$ is a fixed constant.

(2) *If for $\alpha > 0$ and $C > 0$, we have*

$$e_n(\mathcal{K})_X \geq C\frac{1}{[\log_2 n]^\alpha}, \quad n = 1, 2, \dots, \qquad \textit{then} \quad d_n^\gamma(\mathcal{K})_X \geq \frac{C'}{[\log_2 n]^\alpha},$$

for $n = 1, 2, \dots$, where $C' > 0$ is a fixed constant.

(3) *If for $0 < \alpha < 1$ and $C, c > 0$, we have*

$$e_n(\mathcal{K})_X \geq C2^{-cn^\alpha}, \quad n = 1, 2, \dots, \qquad \textit{then} \quad d_n^\gamma(\mathcal{K})_X \geq C'2^{-c'n^{\alpha/(1-\alpha)}},$$

for $n = 1, 2, \dots$, where $C', c' > 0$, are fixed constants.

4.1. Estimates from below for the linear Kolmogorov width.

The above theorem, combined with Theorem 3.1, gives the following relations between linear Kolmogorov widths and entropy numbers.

Theorem 4.2. *Let $\mathcal{K} \subset X$ be a compact subset of a Banach space X, $n \in \mathbb{N}$, and $d_n(\mathcal{K})_X$ be the n-th linear Kolmogorov width for $\mathcal{K}$. Then the following holds:*

(1) *If for $\alpha > 0$, $\beta \in \mathbb{R}$, $C > 0$, we have*

$$e_n(\mathcal{K})_X \geq C\frac{[\log_2 n]^\beta}{n^\alpha}, \quad n = 1, 2, \dots, \qquad \textit{then} \quad d_n(\mathcal{K})_X \geq C'\frac{[\log_2 n]^\beta}{n^\alpha[\log_2 n]^\alpha},$$

for $n = 1, 2, \dots$, where $C' > 0$ is a fixed constant.

(2) *If for $\alpha > 0$, $C > 0$, we have*

$$e_n(\mathcal{K})_X \geq \frac{C}{[\log_2 n]^\alpha}, \quad n = 1, 2, \dots, \qquad \textit{then} \quad d_n(\mathcal{K})_X \geq C'\frac{1}{[\log_2 n]^\alpha},$$

for $n = 1, 2, \ldots$, where $C' > 0$ is a fixed constant.

(3) *If for $0 < \alpha < 1$, $C, c > 0$ we have*

$$e_n(\mathcal{K})_X \geq C 2^{-cn^\alpha}, \quad n = 1, 2, \ldots, \qquad \textit{then} \quad d_n(\mathcal{K})_X \geq C' 2^{-c' n^{\alpha/(1-\alpha)}},$$

for $n = 1, 2, \ldots$, where $C', c' > 0$ are fixed constants.

Proof: The statement follows from Theorem 3.1, Theorem 4.1 and the inequality $\sup_{f \in K} \|f\|_X \geq \mathrm{rad}(\mathcal{K})$. $\square$

4.2. **Estimates from below for the nonlinear Kolmogorov width, the Hilbert space case.** Using Lemma 3.2 and Theorem 4.1, we obtain similar estimates for $d_{n-1}(\mathcal{K}, N)_H$.

Theorem 4.3. *Let $\mathcal{K} \subset H$ be a compact subset of a Hilbert space H and $d_n(\mathcal{K}, N)_H$, $n \in \mathbb{N}$, $N > 1$, be the nonlinear Kolmogorov width for $\mathcal{K}$. Then the following holds:*

- *If for $\alpha > 0$, $\beta \in \mathbb{R}$, and $C > 0$ the entropy numbers satisfy $e_n(\mathcal{K})_H \geq C \frac{[\log_2 n]^\beta}{n^\alpha}$, $n = 1, 2, \ldots$, then there is a constant $C'' > 0$ such that for every $N > 1$we have*

$$d_{n-1}(\mathcal{K}, N)_H \geq C'' \frac{[\log_2(n + \lceil \log_2 N \rceil)]^{\beta - \alpha}}{[n + \lceil \log_2 N \rceil]^\alpha}, \quad n = 1, 2, \ldots. \tag{4.1}$$

- *If for $\alpha > 0$ and $C > 0$, the entropy numbers satisfy the inequality $e_n(\mathcal{K})_H \geq \frac{C}{[\log_2 n]^\alpha}$, $n = 1, 2, \ldots$, then there is a constant $C'' > 0$ such that for every $N > 1$ we have*

$$d_{n-1}(\mathcal{K}, N)_H \geq C'' \frac{1}{[\log_2(n + \lceil \log_2 N \rceil)]^\alpha}, \quad n = 1, 2, \ldots. \tag{4.2}$$

- *If for $0 < \alpha < 1$ and $C, c > 0$, the entropy numbers satisfy the inequality $e_n(\mathcal{K})_H \geq C 2^{-cn^\alpha}$, $n = 1, 2, \ldots$, then there are constants $C'', c'' > 0$ such that for every $N > 1$*

$$d_{n-1}(\mathcal{K}, N)_H \geq C'' 2^{-c''(n + \lceil \log_2 N \rceil)^{\alpha/(1-\alpha)}}, \quad n = 1, 2, \ldots. \tag{4.3}$$

Proof: To show (4.1), we apply Lemma 3.2, Theorem 4.1 with a value $\gamma = \max\{2\mathrm{rad}(\mathcal{K}), 3\}$, and use the monotonicity of the Lipschitz width as a function of γ to derive that

$$\begin{aligned} d_{n-1}(\mathcal{K}, N)_H &\geq d^3_{n + \lceil \log_2 N \rceil}(\mathcal{K})_H \\ &\geq d^\gamma_{n + \lceil \log_2 N \rceil}(\mathcal{K})_H \geq C \frac{[\log_2(n + \lceil \log_2 N \rceil)]^{\beta - \alpha}}{[n + \lceil \log_2 N \rceil]^\alpha}. \end{aligned}$$

We omit the proof of the rest of the theorem since it is similar to the case already discussed. $\square$

Note that the above theorem holds for any value of N. In the cases when $N = \lambda^n$, with $\lambda > 1$, or $N = n^{an}$, with $a > 0$, we obtain two corollaries.

Corollary 4.4. *Let $\mathcal{K} \subset H$ be a compact subset of a Hilbert space H. Then the following holds:*

- *If* $e_n(\mathcal{K})_H \geq C\frac{[\log_2 n]^\beta}{n^\alpha}$, $\quad n = 1, 2, \ldots,$ *then*
$$d_{n-1}(\mathcal{K}, \lambda^n)_H \geq C'' \frac{[\log_2 n]^{\beta-\alpha}}{n^\alpha}, \quad n = 2, 3, \ldots.$$
- *If* $e_n(\mathcal{K})_H \geq C\frac{1}{[\log_2 n]^\alpha}$, $\quad n = 1, 2, \ldots,$ *then*
$$d_{n-1}(\mathcal{K}, \lambda^n)_H \geq C'' \frac{1}{[\log_2 n]^\alpha}, \quad n = 2, 3, \ldots.$$
- *If* $e_n(\mathcal{K})_H \geq C2^{-cn^\alpha}$, $\quad n = 1, 2, \ldots,$ *then*
$$d_{n-1}(\mathcal{K}, \lambda^n)_H \geq C''2^{-c''n^{\alpha/(1-\alpha)}}, \quad n = 2, 3, \ldots.$$

Corollary 4.5. *Let $\mathcal{K} \subset H$ be a compact subset of a Hilbert space H. Then the following holds:*

- *If* $e_n(\mathcal{K})_H \geq C\frac{[\log_2 n]^\beta}{n^\alpha}$, $\quad n = 1, 2, \ldots$ *then*
$$d_{n-1}(\mathcal{K}, n^{an})_H \geq C'' \frac{[\log_2 n]^{\beta-2\alpha}}{n^\alpha}, \quad n = 2, 3, \ldots.$$
- *If* $e_n(\mathcal{K})_H \geq C\frac{1}{[\log_2 n]^\alpha}$, $\quad n = 1, 2, \ldots,$ *then*
$$d_{n-1}(\mathcal{K}, n^{an})_H \geq C'' \frac{1}{[\log_2 n]^\alpha}, \quad n = 2, 3, \ldots.$$
- *If* $e_n(\mathcal{K})_H \geq C2^{-cn^\alpha}$, $\quad n = 1, 2, \ldots,$ *then*
$$d_{n-1}(\mathcal{K}, n^{an})_H \geq C''2^{-c''[n\log_2 n]^{\alpha/(1-\alpha)}}, \quad n = 2, 3, \ldots.$$

Proof: We outline the proof of only the first statement. It follows from (4.1) with $N = n^{an}$ that
$$\begin{aligned} d_{n-1}(\mathcal{K}, n^{an})_H &\geq C'' \frac{[\log_2(n + an\log_2 n)]^{\beta-\alpha}}{[n + an\log_2 n]^\alpha} \geq C_1 \frac{[\log_2(n + an\log_2 n)]^{\beta-\alpha}}{[n\log_2 n]^\alpha} \\ &\geq C_2 \frac{[\log_2 n]^{\beta-\alpha}}{[n\log_2 n]^\alpha}, \end{aligned}$$
where we have used that for n big enough
$$\log_2 n \leq \log_2(n + an\log_2 n) \leq 2\log_2 n.$$
□

4.2.1. *Examples.* Here, we provide an example which shows that some of the estimates in Corollary 4.4 are sharp. We consider the Hilbert space $\ell_2 := \{x = (x_1, x_2, \ldots) : \sum_{j=1}^\infty |x_j|^2 < \infty\}$ with a standard basis $\{e_j)_{j=1}^\infty$ and the strictly decreasing sequence $\sigma = \{\sigma_j\}_{j=1}^\infty$ of positive numbers σ_j which converge to 0 with $\sigma_1 = 1$. We then define the compact set
$$\mathcal{K}_\sigma := \{\sigma_j e_j\}_{j=1}^\infty \cup \{0\} \subset \ell_2$$
and prove the following lemma.

Lemma 4.6. *Every set $\mathcal{K}_\sigma \subset \ell_2$ has inner entropy numbers*

$$\tilde{e}_n(\mathcal{K}_\sigma)_{\ell_2} = \sqrt{\sigma_{2^n}^2 + \sigma_{2^n+1}^2}, \quad n = 1, 2, \dots,$$

and nonlinear Kolmogorov width

$$d_n(\mathcal{K}_\sigma, N)_{\ell_2} \leq \sigma_{nN+1}, \quad N > 1, \quad n = 1, 2, \dots.$$

Proof: Since

$$\|\sigma_j e_j - \sigma_{j'} e_{j'}\|_{\ell_2} = \sqrt{\sigma_j^2 + \sigma_{j'}^2} \leq \sqrt{\sigma_j^2 + \sigma_{j+1}^2}, \quad \text{for all} \quad j' \geq j+1,$$

and

$$\|\sigma_j e_j - 0\|_{\ell_2} = \sigma_j < \sqrt{\sigma_j^2 + \sigma_{j+1}^2},$$

we have that the ball with center $\sigma_j e_j$ and radius $r_j := \sqrt{\sigma_j^2 + \sigma_{j+1}^2}$ contains 0 and all points $\sigma_{j'} e_{j'}$ with $j' > j$, but none of the points $\sigma_{j'} e_{j'}$ with $j' < j$. Thus, if we look for 2^n balls with centers in $\mathcal{K}_\sigma$, covering $\mathcal{K}_\sigma$, and with smallest radius, these are the balls $B(\sigma_j e_j, r_{2^n})$, $j = 1, 2, \dots, 2^n$, with centers $\sigma_j e_j$ and radius r_{2^n}. The j-th ball does not contain the first $(j-1)$ points $\sigma_{j'} e_{j'}$, $1 \leq j' \leq j-1$, from $\mathcal{K}_\sigma$, but contains the rest of the points $\{\sigma_i e_i\}_{i=j}^\infty \cup \{0\}$. Therefore, we have that

$$\tilde{e}_n(\mathcal{K}_\sigma)_{\ell_2} = r_{2^n}.$$

To prove the second statement, we define the n-dimensional spaces

$$X_s := \operatorname{span}\{e_j\}_{(s-1)n+1}^{sn}, \quad s = 1, 2, \dots, N.$$

Clearly $0, \sigma_j e_j \in \bigcup_{s=1}^N X_s$ for $j = 1, \dots, nN$, and for $j > nN$ we have

$$\operatorname{dist}(\sigma_j e_j, \bigcup_{s=1}^N X_s)_{\ell_2} = \sigma_j.$$

Thus, $d_n(\mathcal{K}_\sigma, N)_{\ell_2} \leq \sigma_{nN+1}$, and the proof is completed. $\square$

For our particular example we fix $\alpha > 0$, select the sequence $\{\sigma_j\}_{j=1}^\infty$ to be

$$\sigma_j = \frac{1}{[\log_2 \log_2(j+3)]^\alpha}, \quad j = 1, 2, \dots, \tag{4.4}$$

and show in the following lemma that the estimate in Corollary 4.4 cannot be improved.

Lemma 4.7. *The set $\mathcal{K} := \mathcal{K}_\sigma$ defined by the sequence* (4.4) *has the following properties:*

$$e_n(\mathcal{K})_{\ell_2} \asymp (\log_2 n)^{-\alpha}, \quad \textit{and} \quad d_{n-1}(\mathcal{K}, \lambda^n)_{\ell_2} \asymp (\log_2 n)^{-\alpha}, \quad n = 2, 3, \dots.$$

Proof: It follows from (2.3) and Lemma 4.6 that

$$e_n(\mathcal{K})_{\ell_2} \asymp \sigma_{2^n} \asymp (\log_2 n)^{-\alpha},$$

and that

$$\begin{aligned} d_{n-1}(\mathcal{K},\lambda^n)_{\ell_2} &\leq \sigma_{(n-1)\lambda^n+1} = \frac{1}{[\log_2\log_2((n-1)\lambda^n+1)]^\alpha} \\ &\leq \frac{C}{(\log_2\log_2\lambda^n)^\alpha} = \frac{C}{(\log_2 n+\log_2\log_2\lambda)^\alpha} \\ &\leq \frac{C'}{(\log_2 n)^\alpha}. \end{aligned}$$

The estimate from below follows from Corollary 4.4. □

4.3. **Estimates from below for the nonlinear Kolmogorov width, the Banach space case.** To prove an estimate from below in the Banach space case, we use the following statement from [8], see Theorem 7.3 in [8].

Theorem 4.8. *Let $\mathcal{K}\subset X$ be a compact subset of a Banach space X. Consider the Lipschitz width $d_n^{\gamma_n}(\mathcal{K})_X$ with $\gamma_n=cn^\delta\lambda^n$, $\delta\in\mathbb{R}$, $\lambda>1$, and $c>0$. If for some constants $c_1>0,\alpha>0$ we have $e_n(\mathcal{K})_X>c_1(\log_2 n)^{-\alpha}$, $n=1,2,\dots$, then there exists a constant $C>0$ such that*

$$d_n^{\gamma_n}(\mathcal{K})_X\geq C(\log_2 n)^{-\alpha},\quad n=1,2,\dots.$$

We now use Lemma 3.5 and the above statement to prove the following theorem.

Theorem 4.9. *Let $\mathcal{K}\subset X$ be a compact subset of a Banach space X and $d_n(\mathcal{K},N)_X$, $n\in\mathbb{N}$, $N>1$, be the nonlinear Kolmogorov width for $\mathcal{K}$. If there is $\alpha>0$ and $C>0$ such that the entropy numbers $e_n(\mathcal{K})_X\geq C\frac{1}{[\log_2 n]^\alpha}$, $n=1,2,\dots$, then there is an absolute constant $C''>0$ such that*

$$d_{n-1}(\mathcal{K},\lambda^n)_X\geq C''\frac{1}{[\log_2 n]^\alpha},\quad n=2,\dots.$$

Proof: We apply Lemma 3.5, Theorem 4.8 with $\gamma=2(\lambda^n+1)\sqrt{n}$ and use the monotonicity of the Lipschitz width as a function of γ to derive that

$$d_{n-1}(\mathcal{K},\lambda^n)_X\geq d_n^{2(\lambda^n+1)\sqrt{n}}(\mathcal{K})_X\geq d_n^{c\sqrt{n}\lambda^n}(\mathcal{K})_X\geq C'\frac{1}{[\log_2 n]^\alpha}.$$

□

4.4. **Estimates from above for the entropy numbers.** The next proposition provides us with a tool to derive estimates for the entropy numbers of $\mathcal{K}$ if we have a knowledge about the behavior of the nonlinear Kolmogorov widths $d_n(\mathcal{K},N)_X$.

Proposition 4.10. *Let $\mathcal{K}\subset X$ with $\mathrm{rad}(\mathcal{K})<1$ be a compact subset of a Banach space X and $d_n(\mathcal{K},N)_X$, $N>1$, $n\in\mathbb{N}$, be the nonlinear Kolmogorov width for $\mathcal{K}$. If for some $1>\epsilon>0$ we have $d_n(\mathcal{K},N)_X<\epsilon$, then there exists an absolute constant $c>0$ such that $\tilde{P}_{3\epsilon}(\mathcal{K})\leq N(c/\epsilon)^n$ and*

$$e_{\lceil\log_2\mu\rceil}(\mathcal{K})_X\leq 3cN^{1/n}\mu^{-1/n},\quad \textit{with}\quad \mu=\tilde{P}_{3\epsilon}(\mathcal{K}).$$

Proof: Since $d_n(\mathcal{K}, N)_X < \epsilon$, it follows from the definition of the nonlinear Kolmogorov width that there exist n-dimensional subspaces $X_j \subset X$, $j = 1, \dots, N$, for which

$$\sup_{f\in\mathcal{K}} \inf_{X_j,\, j=1,\dots,N} \mathrm{dist}(f, X_j)_X < \epsilon.$$

Let $\{k_j\}_{j=1}^{\mu} \subset \mathcal{K}$ be a maximal 3ϵ-packing in $\mathcal{K}$, i.e. $\mu = \tilde{P}_{3\epsilon}(\mathcal{K})$. Then, for each k_i there exists $x_i \in X_{j(i)} \subset \bigcup_{j=1}^{N} X_j$ such that $\|k_i - x_i\|_X < \epsilon$, $i = 1, \dots, \mu$, and we can estimate the difference $\|x_i - x_{i'}\|_X$, $i \neq i'$,

$$\|x_i - x_{i'}\|_X \geq \|k_i - k_{i'}\|_X - \|x_{i'} - k_{i'}\|_X - \|x_i - k_i\|_X > \epsilon. \tag{4.5}$$

The condition $\mathrm{rad}(\mathcal{K}) < 1$ implies that there exists $y \in X$ such that

$$\mathcal{K} \subset B(y, 1)_X,$$

and therefore

$$\|x_i - y\|_X \leq \|x_i - k_i\|_X + \|k_i - y\|_X < 1 + \epsilon.$$

Let y_j be the closest to y element from $X_{j(i)} \subset \bigcup_{j=1}^{N} X_j$. Then for any $x_i \in \{x_1, \dots, x_N\}$ ($x_1, \dots, x_N$ may not be necessary all different) we have

$$\|y_j - x_i\|_X \leq \|y_j - y\|_X + \|y - x_i\|_X \leq 2\|y - x_i\|_X < 2(1 + \epsilon),$$

which leads to

$$\{x_i\}_{i=1}^{\mu} \subset \bigcup_{j=1}^{N} B_{X_{j(i)}}(y_j, 2(1+\epsilon)) \subset \bigcup_{j=1}^{N} B_{X_{j(i)}}(y_j, 4), \quad \text{where} \quad y_j \in X_{j(i)}.$$

It follows from (4.5) that the set $\{x_i\}_{i=1}^{\mu}$ is an ϵ-packing for $\bigcup_{j=1}^{N} B_{X_{j(i)}}(y_j, 4)$. We next use (2.2) to derive

$$\mu \leq \widetilde{\mathcal{N}}_{\epsilon/2}\left(\bigcup_{j=1}^{N} B_{X_{j(i)}}(y_j, 4)\right) \leq N(c/\epsilon)^n,$$

where $c > 0$ is an absolute constant. Note that we have applied the inequality

$$\widetilde{\mathcal{N}}_{\epsilon/2}(B_{X_{j(i)}}(y_j, 4)) \leq (c/\epsilon)^n, \quad j = 1, \dots, N,$$

from [7, Chp. 15 Prop.1.3]. In terms of entropy numbers we can write

$$e_{\lceil \log_2 \mu \rceil}(\mathcal{K})_X \leq 3\epsilon \leq 3cN^{1/n}\mu^{-1/n}.$$

$\square$

We use Proposition 4.10 to obtain estimates from above for the entropy numbers $e_m(\mathcal{K})_X$ of $\mathcal{K}$. A similar estimate but for a different range of m and some specific values of N has been recently presented in [11].

Lemma 4.11. *Let $\mathcal{K} \subset X$ be a compact subset of a Banach space X with* $\mathrm{rad}(\mathcal{K}) < 1$. *If for $\alpha > 0$, $\beta \in \mathbb{R}$, $\lambda > 1$, and $c_0 > 0$ we have that*

$$d_n(\mathcal{K}, \lambda^n)_X \leq c_0 \frac{[\log_2 n]^\beta}{n^\alpha},$$

for some $n > n_0(c_0, \alpha, \beta, \lambda)$, then

$$e_m(\mathcal{K})_X < C \frac{[\log_2 m]^{\alpha+\beta}}{m^\alpha}, \quad \textit{with} \quad m = 2\alpha n \log_2 n,$$

where C is a fixed constant depending only on $\lambda, \alpha, \beta, c_0$.

Proof: It follows from Proposition 4.10 with $\epsilon = c_0 \frac{[\log_2 n]^\beta}{n^\alpha}$ that

$$\log_2 \mu \leq n[\log_2(\lambda c) + \alpha \log_2 n - \log_2 c_0 - \beta \log_2(\log_2 n)] \leq 2\alpha n \log_2 n,$$

for $n > n_0$ where n_0 depends only c_0, λ, α and β. For such n's we have

$$e_{2\alpha n \log_2 n}(\mathcal{K})_X \leq \frac{3c_0 [\log_2 n]^\beta}{n^\alpha}.$$

Setting $m = 2\alpha n \log_2 n$ gives

$$e_m(\mathcal{K})_X \leq C m^{-\alpha} [\log_2 n]^{\beta+\alpha}.$$

Since for n sufficiently large, $2^{-1} \log_2 n < \log_2 m < 3 \log_2 n$, the proof is completed. $\square$

Remark 4.12. Similar statement as Lemma 4.11 holds if

$$d_n(\mathcal{K}, n^{an})_X \leq c_0 \frac{[\log_2 n]^\beta}{n^\alpha},$$

where $a > 0$ is a positive constant.

The results in Lemma 4.11 and Remark 4.12 hold also for sequences. Namely, using the monotonicity of the quantities involved, the following is true.

Remark 4.13. Let $\mathcal{K} \subset X$ be a compact subset of a Banach space X with $\mathrm{rad}(\mathcal{K}) < 1$. If there are constants $\alpha > 0$, $\beta \in \mathbb{R}$, $\lambda > 1$, and $c_0 > 0$ such that we have

$$d_n(\mathcal{K}, \lambda^n)_X \leq c_0 \frac{[\log_2 n]^\beta}{n^\alpha}, \quad n = 1, 2, \ldots,$$

or for $a > 0$

$$d_n(\mathcal{K}, n^{an})_X \leq c_0 \frac{[\log_2 n]^\beta}{n^\alpha}, \quad n = 1, 2, \ldots,$$

then

$$e_n(\mathcal{K})_X < C \frac{[\log_2 n]^{\alpha+\beta}}{n^\alpha}, \quad n = 1, 2, \ldots,$$

where C is a fixed constant depending only on α, β, c_0 and λ or a.

5. Applications

In this section, we describe how some of the above results can translate to estimates about m-term approximation. We follow the framework outlined in Theorem 4.1 from [10].

We assume that we have a system $\mathcal{D} = \{g_j\}_{j=1}^{\infty} \subset X$ and de la Vallee-Poussin linear operators V_k associated with the sequences n_k, $\{(V_k, n_k)\}_{k=1}^{\infty}$, satisfying the conditions:

(1) There is a constant $A_2 > 1$ such that

$$V_k(g_j) = \begin{cases} g_j, & j = 1, \dots, n_k, \\ 0, & j > A_2 n_k, \\ \alpha_{k,j} g_j, & \text{otherwise}, \quad \text{where} \quad \alpha_{k,j} \in \mathbb{R}. \end{cases}$$

(2) The norms of V_k as operators from X to X are uniformly bounded, i.e. there is a constant $A_3 > 0$ such that $\|V_k\|_{X \to X} \le A_3$, $k = 1, 2, \dots$.

We denote by $S_{n_k}(f)$ the best approximation to $f \in \mathcal{K}$ by elements from $\mathrm{span}\{g_1, \dots, g_{n_k}\}$,

$$E_{n_k}(f, \mathcal{D})_X := \inf_{c_1, \dots, c_{n_k}} \|f - \sum_{j=1}^{n_k} c_j g_j\|_X = \|f - S_{n_k}(f)\|_X,$$

and by

$$\sigma_m(f, \mathcal{D}')_X := \inf_{\{c_j\},\, \Lambda:\, |\Lambda| = m} \|f - \sum_{j \in \Lambda \cap \mathcal{D}'} c_j g_j\|_X$$

the best m-term approximation of f by a linear combination of m elements from $\mathcal{D}'$, where $\mathcal{D}'$ could be a subset of $\mathcal{D}$ or $\mathcal{D}$ itself. We also define

$$E_{n_k}(\mathcal{K}, \mathcal{D})_X := \sup_{f \in \mathcal{K}} E_{n_k}(f, \mathcal{D})_X, \quad \sigma_m(\mathcal{K}, \mathcal{D}')_X := \sup_{f \in \mathcal{K}} \sigma_m(f, \mathcal{D}')_X.$$

Then the following lemma holds.

Lemma 5.1. *If the Banach space X admits de la Vallee-Poussin linear operators V_k that satisfy* (1)-(2), *with constants $A_2 > 1$, $A_3 > 0$, then we have for $1 < m < A_2 n_k$,*

$$d_m \left(\mathcal{K}, \left(\frac{A_2 b n_k}{m} \right)^m \right)_X \le (1 + 2A_3) \max\{E_{n_k}(\mathcal{K}, \mathcal{D})_X, \sigma_m(\mathcal{K}, \mathcal{D})_X\}, \tag{5.1}$$

where $b > 1$ is an absolute constant.

Proof: Clearly, we have the inequality

$$\begin{aligned} \|f - V_k(f)\|_X &\le \|f - S_{n_k}(f)\|_X + \|S_{n_k}(f) - V_k(f)\|_X \\ &= E_{n_k}(f, \mathcal{D})_X + \|V_k(S_{n_k}(f) - f)\|_X \le (1 + A_3) E_{n_k}(f, \mathcal{D})_X. \end{aligned} \tag{5.2}$$

If we denote by $\mathcal{D}_{A_2 n_k} := \{g_1, \dots, g_{A_2 n_k}\}$, then it follows from the properties of V_k that for any index set Λ with $|\Lambda| = m$ and any coefficients $\{c_j\}_{j=1}^m$,

$$\sigma_m(V_k(f), \mathcal{D}_{A_2 n_k})_X \le \|V_k(f) - V_k(\sum_{j\in\Lambda} c_j g_j)\|_X \le A_3 \|f - \sum_{j\in\Lambda} c_j g_j\|_X,$$

and therefore

$$\sigma_m(V_k(f), \mathcal{D}_{A_2 n_k})_X \le A_3 \sigma_m(f, \mathcal{D})_X. \tag{5.3}$$

Since

$$\sigma_m(f, \mathcal{D}_{A_2 n_k})_X \le \|f - V_k(f)\|_X + \sigma_m(V_k(f), \mathcal{D}_{A_2 n_k})_X,$$

it follows from (5.2) and (5.3) that

$$\begin{aligned} \sigma_m(f, \mathcal{D}_{A_2 n_k})_X &\le (1 + A_3) E_{n_k}(f, \mathcal{D})_X + A_3 \sigma_m(f, \mathcal{D})_X \\ &\le (1 + 2A_3) \max\{E_{n_k}(\mathcal{K}, \mathcal{D})_X, \sigma_m(\mathcal{K}, \mathcal{D})_X\}. \end{aligned}$$

Taking a supremum over $f \in \mathcal{K}$ in the latter inequality gives

$$\sigma_m(\mathcal{K}, \mathcal{D}_{A_2 n_k})_X \le (1 + 2A_3) \max\{E_{n_k}(\mathcal{K}, \mathcal{D})_X, \sigma_m(\mathcal{K}, \mathcal{D})_X\}. \tag{5.4}$$

Note that the total number of m-dimensional subspaces, $1 < m < A_2 n_k$, of the linear space $\operatorname{span}\{g_1, \dots, g_{A_2 n_k}\}$ is $\binom{A_2 n_k}{m}$. Using the Stirling formula, one can show that there is an absolute constant $b > 1$ such that

$$\binom{A_2 n_k}{m} \le \left(\frac{A_2 b n_k}{m}\right)^m.$$

Then the definition of nonlinear Kolmogorov width and its monotonicity with respect to N gives

$$d_m\left(\mathcal{K}, \left(\frac{A_2 b n_k}{m}\right)^m\right)_X \le d_m\left(\mathcal{K}, \binom{A_2 n_k}{m}\right)_X \le \sigma_m(\mathcal{K}, \mathcal{D}_{A_2 n_k})_X.$$

The latter inequality combined with (5.4) leads to

$$d_m\left(\mathcal{K}, \left(\frac{A_2 b n_k}{m}\right)^m\right)_X \le (1 + 2A_3) \max\{E_{n_k}(\mathcal{K}, \mathcal{D})_X, \sigma_m(\mathcal{K}, \mathcal{D})_X\},$$

where $1 < m < A_2 n_k$, and the proof is completed. $\Box$.

We next state a theorem that follows from Lemma 5.1 and our inequalities for nonlinear Kolmogorov widths in Hilbert spaces. Note that our theorem does not require the additional assumptions on the error $E_n(\mathcal{K}, \mathcal{D})_H$ that are needed in Theorem 4.1 from [10] and describes the behavior of the errors in cases not covered by this theorem.

Theorem 5.2. *If the Hilbert space H admits de la Vallee-Poussin linear operators V_k that satisfy* (1)-(2)*, then the following holds:*

- *If $e_{n_k}(\mathcal{K})_H \ge C \frac{[\log_2 n_k]^\beta}{n_k^\alpha}$, $k = 1, 2, \dots$, then there is an absolute constant $C'' > 0$ such that*

$$\max\{E_{n_k}(\mathcal{K}, \mathcal{D})_H, \sigma_m(\mathcal{K}, \mathcal{D})_H\} \ge C'' \frac{[\log_2 m(1 + \log_2(A_2 n_k/m))]^{\beta-\alpha}}{m^\alpha [1 + \log_2(A_2 b n_k/m)]^\alpha},$$

for $1 < m < n_k$, $k = 1, 2, 3, \dots$.

- *If* $e_{n_k}(\mathcal{K})_H \geq C\frac{1}{[\log_2 n_k]^\alpha}$, $k=1,2,\dots$, *then there is an absolute constant* $C''>0$ *such that*

$$\max\{E_{n_k}(\mathcal{K},\mathcal{D})_H, \sigma_m(\mathcal{K},\mathcal{D})_H\} \geq C''\frac{1}{[\log_2 m(1+\log_2(A_2bn_k/m))]^\alpha},$$

for $1<m<n_k$, $k=1,2,3,\dots$.

- *If* $e_{n_k}(\mathcal{K})_H \geq C2^{-cn_k^\alpha}$, $k=1,2,\dots$, *then there are absolute constants* $C''>0$ *and* $c''>0$ *such that*

$$\max\{E_{n_k}(\mathcal{K},\mathcal{D})_H, \sigma_m(\mathcal{K},\mathcal{D})_H\} \geq C''2^{-c''[m(1+\log_2(A_2bn_k/m))]^{\alpha/(1-\alpha)}},$$

for $1<m<n_k$, $k=1,2,3,\dots$.

Proof: We use Theorem 4.3 in the case $N=\left(\frac{A_2bn_k}{m}\right)^m$, Lemma 5.1 and the fact that

$$d_m\left(\mathcal{K},\left(\frac{A_2bn_k}{m}\right)^m\right)_X \geq d_{n_k-1}\left(\mathcal{K},\left(\frac{A_2bn_k}{m}\right)^m\right)_X, \quad 1<m<n_k.$$

Note that we have utilized the fact that the constants in Theorem 4.3 do not depend on N. □

We can derive several corollaries from the above theorem, one of which we state below. If we take $m=n_k/2$ in Theorem 5.2, we obtain the following statement.

Corollary 5.3. *If the Hilbert space* H *admits de la Vallee-Poussin linear operators* V_k *that satisfy* (1)-(2)*, then the following holds:*

- *If* $e_{n_k}(\mathcal{K})_H \geq C\frac{[\log_2 n_k]^\beta}{n_k^\alpha}$, $k=1,2,\dots$, *then there is an absolute constant* $C''>0$ *such that*

$$\max\{E_{n_k}(\mathcal{K},\mathcal{D})_H, \sigma_{n_k/2}(\mathcal{K},\mathcal{D})_H\} \geq C''\frac{[\log_2 n_k]^{\beta-\alpha}}{n_k^\alpha}, \quad k=1,2,\dots.$$

- *If* $e_{n_k}(\mathcal{K})_H \geq C\frac{1}{[\log_2 n_k]^\alpha}$, $k=1,2,\dots$, *then there is an absolute constant* $C''>0$ *such that*

$$\max\{E_{n_k}(\mathcal{K},\mathcal{D})_H, \sigma_{n_k/2}(\mathcal{K},\mathcal{D})_H\} \geq \frac{C''}{[\log_2 n_k]^\alpha}, \quad k=1,2,\dots.$$

- *If* $e_{n_k}(\mathcal{K})_H \geq C2^{-cn_k^\alpha}$, $k=1,2,\dots$, *then there are absolute constants* $C''>0$, $c''>0$ *such that*

$$\max\{E_{n_k}(\mathcal{K},\mathcal{D})_H, \sigma_{n_k/2}(\mathcal{K},\mathcal{D})_H\} \geq C''2^{-c''n_k^{\alpha/(1-\alpha)}}, \quad k=1,2,\dots.$$

Note that since $A_2>1$, we can take $m=n_k$ in Lemma 5.1, use the fact that

$$E_{n_k}(\mathcal{K},\mathcal{D})_H \geq \sigma_{n_k}(\mathcal{K},\mathcal{D})_H,$$

and obtain from this lemma that if the Banach space X admits de la Vallee-Poussin linear operators satisfying (1)-(2), then

$$d_{n_k}(\mathcal{K},(A_2b)^{n_k})_H \leq (1+2A_3)E_{n_k}(\mathcal{K},\mathcal{D})_H.$$

We can now use Corollary 4.4 and the monotonicity of the nonlinear Kolmogorov width with respect to N to conclude that

$$d_{n_k}\left(\mathcal{K},(A_2b)^{n_k+1}\right)_H \leq d_{n_k}\left(\mathcal{K},(A_2b)^{n_k}\right)_H,$$

and derive the following statement.

Corollary 5.4. *If the Hilbert space H admits de la Vallee-Poussin linear operators V_k that satisfy* (1)-(2)*, then the following holds:*

- *If $e_{n_k+1}(\mathcal{K})_H \geq C\frac{[\log_2 n_k]^\beta}{n_k^\alpha}$, $\quad k=1,2,\dots$, then there is an absolute constant $C''>0$ such that*

$$E_{n_k}(\mathcal{K},\mathcal{D})_H \geq C''\frac{[\log_2 n_k]^{\beta-\alpha}}{n_k^\alpha}, \quad k=1,2,\dots.$$

- *If $e_{n_k+1}(\mathcal{K})_H \geq C\frac{1}{[\log_2 n_k]^\alpha}$, $\quad k=1,2,\dots$, then there is an absolute constant $C''>0$ such that*

$$E_{n_k}(\mathcal{K},\mathcal{D})_H \geq \frac{C''}{[\log_2 n_k]^\alpha}, \quad k=1,2,\dots.$$

- *If $e_{n_k+1}(\mathcal{K})_H \geq C2^{-cn_k^\alpha}$, $\quad k=1,2,\dots$, then there are absolute constants $C''>0$, $c''>0$, such that*

$$E_{n_k}(\mathcal{K},\mathcal{D})_H \geq C''2^{-c''n_k^{\alpha/(1-\alpha)}}, \quad k=1,2,\dots.$$

Acknowledgment: This work was supported by the NSF Grant DMS 2134077, Tripods Grant CCF-1934904, and the ONR Contract N00014-20-1-278.

References

[1] Ball, Keith, *An elementary introduction to modern convex geometry. Flavors of geometry,* 1–58, Math. Sci. Res. Inst. Publ., 31, Cambridge Univ. Press, Cambridge, 1997.

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

[3] Hongwei Huang, Heping Wang, *Greedy algorithm with regard to the needlet system on the sphere*, J. Math. Anal. Appl. 454 (2017), 557–570.

[4] T. Kühn, *Entropy Numbers of General Diagonal Operators*, Rev. Mat. Complut. 2005 18.2 479–491.

[5] D.R. Lewis, *Finite dimensional subspaces of L_p*, Studia Mathematica LXIII (1978), 207–212.

[6] A. Kolmogoroff (A.N. Kolmogorov), *Über die beste annäherung von funktionen einer gegeben funktionenklasse*, Annals of Mathematics, 37.1 (1936), 107–110.

[7] G.G. Lorentz, M.v. Golitschek, Y. Makovoz, *Constructive Approximation, Advanced Problems*, Springer Verlag, 1996.

[8] G. Petrova, P. Wojtaszczyk, *Lipschitz widths*, submitted, arXiv:2111.01341.

[9] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge Univ. Press, Cambridge, 1989.

[10] V. Temlyakov, *Nonlinear Kolmogorov widths* , Mathematical Notes, **63**, (1998), 785–795.

[11] V. Temlyakov, *An inequality for the entropy numbers and its application*, arxiv 1301.7624v1.

(G. Petrova) Department of Mathematics, Texas A&M University, College Station, TX 77843, USA
Email address: gpetrova@math.tamu.edu

(P. Wojtaszczyk) Institut of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland
Email address: wojtaszczyk@impan.pl