Greedy Algorithms and Kolmogorov Widths in Banach Spaces

Van Kien Nguyen

April 12, 2018

Abstract

Let $X$ be a Banach space and $K$ be a compact subset in $X$. We consider a greedy algorithm for finding $n$-dimensional subspace $V_n \subset X$ which can be used to approximate the elements of $K$. We are interested in how well the space $V_n$ approximates the elements of $K$. For this purpose we compare the greedy algorithm with the Kolmogorov width which is the best possible error one can approximate $K$ by $n$-dimensional subspaces. Various results in this direction have been given, e.g., in Binev et al. (SIAM J. Math. Anal. (2011)), DeVore et al. (Constr. Approx. (2013)) and Wojtaszczyk (J. Math. Anal. Appl. (2015)). The purpose of the present paper is to continue this line. We shall show that under some additional assumptions the results in the above-mentioned papers can be improved.

1 Introduction

Recently, a new greedy algorithm for obtaining a good subspace $V_n$ of $n$-dimension to approximate elements of a compact set $K$ in a Banach space $X$ has been given. This greedy algorithm was studied initially when $X$ is a Hilbert space in the context of reduced basis methods for solving families of PDEs, see [6, 7]. Later, it was studied extensively not only in the setting of Hilbert spaces, let us mention, for instance, Binev et al. [1], Buffa et al. [2], DeVore et al. [3], and Wojtaszczyk [13]. The greedy algorithm for generating the subspace $V_n$ to approximate elements of $K$ is implemented as follows. We first select $f_0$ such that

$$
\|f_0\|_X = \max_{f \in K} \|f\|_X .
$$

Since $K$ is compact, such a $f_0$ always exists. At the general step, assuming that $\{f_0, \ldots, f_{n-1}\}$ and $V_n = \text{span}\{f_0, \ldots, f_{n-1}\}$ have been chosen, then we take $f_n$ such that

$$
\text{dist}(f_n, V_n)_X = \max_{f \in K} \text{dist}(f, V_n)_X .
$$

The error in approximating the elements of $K$ by $V_n$ is defined as

$$
\sigma_0(K)_X := \|f_0\|_X , \quad \sigma_n(K)_X := \text{dist}(f_n, V_n)_X = \max_{f \in K} \text{dist}(f, V_n)_X
$$

for $n \geq 1$. The sequence $\sigma_n(K)_X$ is non-increasing. It is important to note that the sequence $\{f_n\}_{n \geq 0}$ and also $\sigma_n(K)_X$ are not unique.

Let us mention that the best possible error one can achieve to approximate the elements of $K$ by $n$-dimensional subspaces is the Kolmogorov width $d_n(K)_X$, which is given by

$$
d_n(K)_X := \inf_{L} \sup_{f \in K} \text{dist}(f, L)_X , \quad n \geq 1,
$$

*E-mail: vnguyen@ins.uni-bonn.de, kiennv@utc.edu.vn
where the infimum is taken over all \( n \)-dimensional subspaces \( L \) of \( X \). We also put
\[
d_0(\mathcal{K})_X = \max_{f \in \mathcal{K}} \| f \|_X.
\]
We would like to emphasize that in practice, finding subspaces which give this performance is out of reach.

We are interested in how well the subspaces created by the greedy algorithm approximate the elements of \( \mathcal{K} \). For this purpose it is natural to compare \( \sigma_n(\mathcal{K})_X \) with the Kolmogorov width \( d_n(\mathcal{K})_X \). Various comparisons between \( \sigma_n(\mathcal{K})_X \) and \( d_n(\mathcal{K})_X \) have been made. The first attempt in this direction was given in [2] and improved in [1], where the authors considered the case when \( X \) is a Hilbert space \( H \). Under this assumption, it has been shown that
\[
\sigma_n(\mathcal{K})_H \leq C 2^n d_n(\mathcal{K})_H
\]
for an absolute constant \( C \). Observe that this result is useful only when \( d_n(\mathcal{K})_H \) decays faster than \( 2^{-n} \). A significant improvement of the above result was given in [3] where the authors prove that if the Kolmogorov width has polynomial decay with rate \( n^{-s} \), then the greedy algorithm also yields the same rate, i.e., \( \sigma_n(\mathcal{K})_H \leq C n^{-s} \). In the same paper, the estimate of this type for Banach spaces \( X \) was also considered, but there is an additional factor \( n^{-s} \epsilon \) (for any \( \epsilon > 0 \)), that is,
\[
\sigma_n(\mathcal{K})_X \leq C n^{-s} + \frac{1}{2} n^{\epsilon}
\]
where \( C \) depends on \( s \) and \( \epsilon \).

For a recent result in this direction we refer to [13]. Let \( \tilde{\gamma}_n(X) \) be the supremum of Banach-Mazur distance \( d(V_n, \ell_2^n) \) where \( V_n \) is the \( n \)-dimensional space in a quotient space of \( X \). If \( d_n(\mathcal{K})_X \leq C_0 n^{-s} \) and \( \tilde{\gamma}_n(X) \leq C_1 n^{\mu} \), then Wojtaszczyk [13] shows that there is a constant \( C \) such that
\[
\sigma_n(\mathcal{K})_X \leq C \left( \frac{\log(n+2)}{n} \right)^s n^\mu.
\]
Observe that the estimate given in (1.2) improves the result (1.1) since \( \tilde{\gamma}_n(X) \leq \sqrt{n} \). It has been shown in [13] that the above estimate is optimal in \( L_p \) up to a logarithmic factor. However, for a given Banach space \( X \), the factor \( \tilde{\gamma}_n(X) \) is not easy to compute. Hence, this raises the question whether we can replace the condition on \( \tilde{\gamma}(X) \) by \( \gamma_n(X) = \sup_{V_n} d(V_n, \ell_2^n) \) where the supremum is taken over \( n \)-dimensional subspaces \( V_n \) in \( X \), see Section 2 for the definition.

In the present paper we will give a new analysis of the performance of the greedy algorithm in which we show that the assumption on \( \tilde{\gamma}_n(X) \) can be relaxed to \( \gamma_n(X) \). In addition the rate of the logarithm in (1.2) can also be improved when \( s > 1/2 \). More precisely we shall prove that there is a constant \( C > 0 \) such that
\[
\sigma_n(\mathcal{K})_X \leq C \sqrt{\log(2n)} n^{-s + \mu}
\]
if \( d_n(\mathcal{K})_X \leq C_0 n^{-s} \) and \( \gamma_n(X) \leq C_1 n^{\mu} \).

Often, the compact set of interest \( \mathcal{K} \) is the image (or subset) of the closed unit ball \( B_E \) of a Banach space \( E \) under a compact operator \( T \in L(E, X) \). For this reason, we shall compare \( \sigma_n(\mathcal{K})_X \) with the Kolmogorov widths \( d_n(T(B_E))_X \). In this study, we obtain the estimate
\[
\sigma_{3n-1}(\mathcal{K})_X \leq 3e^2 \Gamma(E) \Gamma_n(X) \left( \prod_{k=0}^{n-1} d_k(T(B_E))_X \right)^{1/n}, \quad n \geq 1,
\]
here \( \Gamma_n(X) \) is the \( n \)-Grothendieck number of \( X \), which is closely related to \( \gamma_n(X) \), see Section 2. Note that if \( E \) is a Hilbert space then \( \Gamma(E) = 1 \). In Section 3 we will give an example showing that the estimate (1.3) is sharp in some situations.

The rest of our paper is organized as follows. In the next Section 2 we will collect some required tools. The main results are stated and proved in Section 3.


2 Some preparations

In this section we collect some tools needed to formulate our results in the next section. The Banach - Mazur distance of two isomorphic Banach spaces $X$ and $Y$ is defined by

$$d(X, Y) = \inf \{ \|T\| \cdot \|T^{-1}\| : T : X \to Y \text{ is an isomorphism}\}.$$ 

For a Banach space $X$ we introduce a sequence of numbers

$$\gamma_n(X) = \sup \{ d(V, \ell^2_n) : V \text{ is an } n\text{-dimensional subspace in } X \}.$$ 

The sequence $\gamma_n(X)$ is non-decreasing and $\gamma_1(X) = 1$. It is obvious that if $X$ is a Hilbert space then we have $\gamma_n(X) = 1, n = 1, 2, 3, \ldots$. In the case of an arbitrary Banach space $X$, it is known that $\gamma_n(X) \leq n^{1/2}$ and $\gamma_n(L_p) \leq n^{\frac{1}{2} - \frac{1}{p}}$ for $1 \leq p \leq \infty$.

Let $X$ and $Y$ be Banach spaces of finite dimension. Then there exists an operator $T : X \to Y$ such that $d(X, Y) = \|T\| \cdot \|T^{-1}\|$. We can additionally assume that $\|T^{-1}\| = 1$. Hence a new norm on $X$ defined by $\|x\|_e := \|Tx\|_Y$ satisfies

$$\|x\|_X \leq \|x\|_e \leq d(X, Y) \|x\|_X.$$ (2.1)

Moreover $T$ is an isometry between $(X, \| \cdot \|_e)$ and $Y$.

The local injective distance $\gamma_n(X)$ is closely related to the so-called Grothendieck number. Let $T \in \mathcal{L}(X, Y)$ be a linear bounded operator. The $n$-th Grothendieck number of $T$ is defined as

$$\Gamma_n(T) := \sup \left\{ \left| \det \left( \langle T x_i, b_j \rangle \right) \right|^{1/n} : x_1, \ldots, x_n \in B_X, b_1, \ldots, b_n \in B_Y \right\}.$$ 

If $T$ is the identity map of $X$ then we write $\Gamma_n(X)$. Let $0 \leq \delta \leq 1/2$. A Banach space $X$ is said to be of weak Hilbert type $\delta$ if there exists a constant $C \geq 1$ such that $\Gamma_n(X) \leq C n^\delta$ for $n \geq 1$. We denote the class of these spaces by $\Gamma_\delta$. Note that $\Gamma_{1/2}$ is the set of all Banach spaces, i.e.,

$$\Gamma_n(X) \leq cn^{1/2}, \quad \text{for all } X.$$ 

In particular we have $\Gamma_n(L_p) \leq n^{1/p - 1/2}$ for $1 \leq p \leq \infty$. The relation between $\gamma_n(X)$ and Grothendieck numbers is represented in the following lemma, see, e.g., [10].

**Lemma 2.1.** Let $X$ be a Banach space. Then

$$\gamma_n(X) \leq C_1 n^\delta \quad \text{if and only if} \quad \Gamma_n(X) \leq C_2 n^\delta$$

for some $C_1, C_2 \geq 1$.

For later use, let us introduce the notion of Kolmogorov and Gelfand widths of linear continuous operators. In the following we use the definition given in [12, Chapter 2], but see also [8, Chapter 11]. Note that there is a shift of 1 between definitions by Pinkus [12, Chapter 2] and Pietsch [8, Chapter 11]. Let $B_X$ be the closed unit ball of $X$. The Kolmogorov $n$-width of the operator $T \in \mathcal{L}(X, Y)$ is defined as

$$d_n(T) := d_n(T(B_X))_Y = \inf_{L_n} \sup_{\|x\| \leq 1} \inf_{y \in L_n} \|Tx - y\|_Y,$$

where the infimum is taken over all subspace $L_n$ of dimension $n$ in $Y$. The Gelfand $n$-th width of $T \in \mathcal{L}(X, Y)$ is given by

$$d^n(T) := d^n(T(B_X))_Y := \inf_{L^n} \sup_{\|x\| \leq 1, x \in L^n} \|Tx\|_Y,$$
where the infimum is taken over subspaces $L^n$ of $X$ of co-dimension at most $n$. We also put $d_0(T) = d^n(T) = \|T\|$. Note that Kolmogorov and Gelfand widths are closely related, i.e., $d^n(T) = d_n(T')$ for every $T \in \mathcal{L}(X, Y)$ and $d_n(T) = d^n(T)$ if $T$ is compact or $Y$ is a reflexive Banach space, see [12, Chapter 2]. Here recall that $T'$ is the dual operator of $T$. For basic properties of these quantities we refer to monographs [12, Chapters 2] and [8, Chapter 11].

The relation between Grothendieck number and Kolmogorov, Gelfand widths is given in the following lemma. For a proof we refer to [10].

**Lemma 2.2.** Let $X$ and $Y$ be Banach spaces and $T \in \mathcal{L}(X, Y)$. Then it holds

$$
\left( \prod_{k=0}^{n-1} d_k(T) \right)^{1/n} \leq \Gamma_n(T) \quad \text{and} \quad \left( \prod_{k=0}^{n-1} d^k(T) \right)^{1/n} \leq \Gamma_n(T)
$$

for all $n \geq 1$.

An operator $T \in \mathcal{L}(X, Y)$ is called absolutely 2-summing if there exists a constant $C$ such that

$$
\left( \sum_{i=1}^{n} \|Tx_i\|^2 \right)^{1/2} \leq C \sup \left\{ \left( \sum_{i=1}^{n} |\langle x_i, b \rangle|^2 \right)^{1/2} : b \in X', \|b\|_{X'} \leq 1 \right\}. \tag{2.2}
$$

The set of these operators is denoted by $\mathcal{B}_2(X, Y)$ and the norm $\|T|\mathcal{B}_2\|$ is given by the infimum of all $C > 0$ satisfying (2.2). The following assertion can be found in [10].

**Lemma 2.3.** Let $X$ and $Y$ be Banach spaces. Let $T \in \mathcal{B}_2(X, Y)$. Then we have

$$
\Gamma_n(T) \leq e n^{-1/2} \|T|\mathcal{B}_2\| \Gamma_n(X), \quad n \geq 1.
$$

## 3 Main results

Our first result can be formulated as follows.

**Theorem 3.1.** Let $X$ be a Banach space and $K$ a compact subset of $X$. Assume that

$$
d_n(K)_X \leq C_0 \max(1, n)^{-s}, \quad (n \geq 0) \quad \text{and} \quad \gamma_n(X) \leq C_1 n^\mu, \quad (n \geq 1)
$$

for $0 \leq \mu \leq \frac{1}{2}$ and $s > \mu$. Then we have

$$
\sigma_n(K)_X \leq C_0 C_1 2^{\mu 16^s \sqrt{\log(2n)}} n^{-s+\mu} \quad \text{for } n \geq 2. \tag{3.1}
$$

**Proof.** The idea of the proof follows from the proof of Proposition 2.2 in [13].

**Step 1.** Let $\varepsilon > 0$. From the assumption $d_n(K)_X \leq C_0 n^{-s}$, $n \geq 1$, we infer the existence of a sequence of subspaces $(T_k)_{k \geq 0}$ in $X$ and $\dim(T_k) = 2^k$ such that

$$
\max_{x \in K} \min_{g \in T_k} \|x - g\|_X \leq (C_0 + \varepsilon) 2^{-sk}.
$$

For $n \in \mathbb{N}$ fixed we put $V_k = T_0 + T_1 + \ldots + T_{k-1}$ for $k = 1, \ldots, n$. Then we have $V_k \subset V_{k+1}$ and $\dim(V_k) < 2^k$. Observe that

$$
\max_{x \in K} \min_{g \in V_k} \|x - g\|_X \leq \max_{x \in K} \min_{g \in T_{k-1}} \|x - g\|_X \leq (C_0 + \varepsilon) 2^{-s(k-1)}. \tag{3.2}
$$

We denote $N = 2^n$. Implementing the greedy algorithm for the set $K$ we get the sequence $\{f_0, \ldots, f_{N-1}\}$. Then it follows from (3.2) that

$$
\|f_\ell - g^k\|_X \leq (C_0 + \varepsilon) 2^{-s(k-1)}, \quad \ell = 0, \ldots, N - 1; \quad k = 1, \ldots, n \tag{3.3}
$$

4
for some $g_k^i \in V_k$. Let $X = \text{span}\{f_0, \ldots, f_{N-1}\}$ and $Y = \text{span}\{V_n, X\}$. It is obvious that $2^n \leq \dim(Y) < 2^{n+1}$. From (2.1) we infer the existence of a Euclidean norm $\| \cdot \|_e$ on $Y$ satisfying
\[
\|y\|_X \leq \|y\|_e \leq d(Y, \ell_2^{\dim(Y)}) \|y\|_X \leq \gamma_{\dim(Y)}(X) \|y\|_X \leq A \|y\|_X, \tag{3.4}
\]
where we put $A = \gamma_{2^{n+1}}(X)$.

Let $Q$ be the orthogonal projection from $Y$ onto $X$ in the Euclidean norm $\| \cdot \|_e$. We denote $\dim(Q(V_k)) = h_k$ for $k = 1, \ldots, n$. It is clear that $h_k \leq \dim(V_k) < 2^k$ and $Q(V_{k-1}) \subset Q(V_k)$.

From (3.3) and (3.4) we get
\[
\text{dist}(f_\ell, Q(V_k)) \| \cdot \|_e \leq \|f_\ell - Q(g_k^i)\|_e = \|Q(f_\ell - g_k^i)\|_e \leq \|f_\ell - g_k^i\|_e \leq (C_0 + \varepsilon) A 2^{-s(k-1)}. \tag{3.5}
\]

By $\{\phi_j\}_{j=0}^{N-1}$ we denote the orthonormal system obtained from $f_0, \ldots, f_{N-1}$ by Gram-Schmidt orthogonalization in the norm $\| \cdot \|_e$. It follows that the matrix $[\phi_j(f_\ell)]_{j,\ell=0}^{N-1}$ has a triangular form. In particular, on the diagonal we have
\[
\text{dist} \left( f_\ell, \text{span}\{f_0, \ldots, f_{\ell-1}\} \right) \| \cdot \|_e \geq \text{dist} \left( f_\ell, \text{span}\{f_0, \ldots, f_{\ell-1}\} \right)_X = \sigma_\ell(\mathcal{K})_X. \tag{3.6}
\]

**Step 2.** We consider the case
\[
0 < h_{m_1} = \ldots = h_{m_2 - 1} < h_{m_2} = \ldots = h_{m_3 - 1} < \ldots < h_{m_L} = \ldots = h_n
\]
where $m_1 = 1$, $m_{L+1} = n + 1$. We denote $\{x_j\}_{j=0}^{N-1}$ another orthonormal basis in $X$, such that
\[
Q(V_{m_i-1}) = \ldots = Q(V_{m_{i-1}}) = \text{span}\{x_0, \ldots, x_{h_{m_i-1}-1}\}
\]
for $i = 2, \ldots, L$. Considering the vector $[x_j(f_\ell)]_{j=0}^{N-1}$ we observe that
\[
\sum_{j=h_{m_L}}^{N-1} |x_j(f_\ell)|^2 = \text{dist}(f_\ell; Q(V_n))^2 \| \cdot \|_e \tag{3.7}
\]
and
\[
|x_0(f_\ell)|^2 \leq \|f_\ell\|_e^2, \quad \sum_{j=h_{m_{i-1}}+1}^{h_{m_i}-1} |x_j(f_\ell)|^2 \leq \text{dist}(f_\ell; Q(V_{m_{i-1}}))^2 \| \cdot \|_e, \tag{3.8}
\]
for $i = 2, \ldots, L$. Note that
\[
\prod_{j=0}^{N-1} \sigma_j(\mathcal{K})_X \leq \prod_{j=0}^{N-1} |\phi_j(f_\ell)| = |\det[\phi_j(f_\ell)]| = |\det[x_j(f_\ell)]|,
\]
see (3.6). By $k_j$ we denote the $j$-th column of the matrix $[x_j(f_\ell)]_{j,\ell=0}^{N-1}$. Applying Hadamard’s
inequality and then arithmetic-geometric mean inequality we obtain

\[
\left( \prod_{j=0}^{N-1} \sigma_j(K) \right)^2 \leq \left( \det[x_j(f_\ell)] \right)^2 \\
\leq \left( \prod_{j=0}^{h_1-1} \|k_j\|_e^2 \right) \left( \prod_{j=h_{mL}}^{N-1} \|k_j\|_e^2 \right) \left( \prod_{i=2}^{L} \prod_{j=h_{m_{i-1}}}^{h_{m_{i-1}}-1} \|k_j\|_e^2 \right) \\
\leq \left( \frac{1}{h_1} \sum_{j=0}^{h_1-1} \|k_j\|_e^2 \right) \left( \frac{1}{N-h_{mL}} \sum_{j=h_{mL}}^{N-1} \|k_j\|_e^2 \right) \\
\times \prod_{i=2}^{L} \left( \frac{1}{h_{m_i} - h_{m_{i-1}}} \sum_{j=h_{m_{i-1}}}^{h_{m_{i-1}}-1} \|k_j\|_e^2 \right) h_{m_i}^{-h_{m_{i-1}}}.
\] (3.9)

From (3.5), (3.7), and (3.8) we have

\[
\sum_{j=0}^{h_{m_{i-1}}} \|k_j\|_e^2 = \sum_{j=h_{m_{i-1}}}^{N-1} \|x_j(f_\ell)\|^2 \leq \sum_{\ell=0}^{N-1} \|x_\ell\|^2 \leq NA^2d_0(K)X \leq NA^2(C_0 + \varepsilon)^2
\]

(since \(d_0(K)X \leq C_0\), by our assumption), and for \(i = 2, \ldots, L\),

\[
\sum_{j=h_{m_{i-1}}}^{h_{m_{i-1}}-1} \|k_j\|_e^2 = \sum_{j=h_{m_{i-1}}}^{h_{m_{i-1}}-1} \sum_{\ell=0}^{N-1} |x_j(f_\ell)|^2 \\
\leq \sum_{\ell=0}^{N-1} \text{dist}(f_\ell; Q(V_{m_{i-1}})) \leq N(C_0 + \varepsilon)^2 A^2 2^{-2s(m_i-2)}.
\]

Similarly, we have

\[
\sum_{j=h_{mL}}^{N-1} \|k_j\|_e^2 \geq \sum_{\ell=0}^{N-1} \text{dist}(f_\ell; Q(V_n)) \leq N(C_0 + \varepsilon)^2 A^2 2^{-2s(n-1)}.
\]

Inserting this into (3.9) we find

\[
\left( \prod_{j=0}^{N-1} \sigma_j(K) \right)^2 \leq \left( \frac{NA^2(C_0 + \varepsilon)^2}{h_1} \right) \left( \frac{N(C_0 + \varepsilon)^2 A^2 2^{-2s(m_i-2)}}{N-h_{mL}} \right) \\
\times \prod_{i=2}^{L} \left( \frac{N(C_0 + \varepsilon)^2 A^2 2^{-2s(n-1)}}{h_{m_i} - h_{m_{i-1}}} \right) h_{m_i}^{-h_{m_{i-1}}} \\
= MA^2N(C_0 + \varepsilon)^{2N} \left( 2^{-2s(n-1)(N-h_{mL})} \prod_{i=2}^{L} 2^{-2s(m_i-2)(h_{m_i} - h_{m_{i-1}})} \right),
\] (3.10)

where we put

\[
M = \left( \frac{N}{h_1} \right)^{h_1} \left( \frac{N}{N-h_{mL}} \right)^{N-h_{mL}} \prod_{i=2}^{L} \left( \frac{N}{h_{m_i} - h_{m_{i-1}}} \right)^{h_{m_i} - h_{m_{i-1}}}.
\]
For nonnegative number \( a_1, \ldots, a_n \) and positive numbers \( p_1, \ldots, p_n \) we have
\[
a_1^{p_1} \cdots a_n^{p_n} \leq \left( \frac{a_1 p_1 + \cdots + a_n p_n}{p_1 + \cdots + p_n} \right)^{p_1 + \cdots + p_n},
\]
see, e.g., [4, Page 17]. Applying the above inequality for \( M \) we get
\[
M \leq (L + 1)^N \leq (n + 1)^N.
\]
(3.11)

Now we deal with the term
\[
U := \left( 2^{-2s(n - 1)(N - h_mL)} \prod_{i=2}^L 2^{-2s(m_i - 2)(h_mL - h_{m_i} - 1)} \right)
\]
\[
= 2^{2s[-(n-1)2^n+(n-1)h_mL-(mL-2)(h_mL-h_{mL-1})-\cdots-(mL-2)(h_mL-h_{m_L})]}
\]
\[
= 2^{2s[-(n-1)2^n+h_mL(n+1-mL)+h_mL-1(mL-mL-1)+\cdots+h_mL(mL-mL-1)+h_mL(mL-mL-1)]}.
\]

Using \( h_{m_i} < 2^{m_i} \) for \( i = 1, \ldots, L \) we can estimate
\[
U \leq 2^{2s[-(n-1)2^n+2^n+\cdots+2^n]} \leq 2^{2s[-(n-3)2^n]}.
\]
(3.12)

Plugging (3.11) and (3.12) into (3.10) we obtain
\[
\left( \prod_{j=0}^{N-1} \sigma_j(K)_X \right)^2 = (n + 1)^N A^{2N} (C_0 + \varepsilon)^{2N} 2^{(-n+3)2^n} 2^s.
\]

Finally from the assumption \( A \leq C_1 2^\mu(n+1) \) we find
\[
\sigma_{2^n-1}(K)_X \leq (C_0 + \varepsilon) C_1 \sqrt{n+1} \cdot 2(n+1)^n \mu 2^{(-n+3)s}
\]
\[
= (C_0 + \varepsilon) C_1 8^s \sqrt{n+1} \cdot 2(n+1)^n \mu 2^{-ns}
\]

and hence
\[
\sigma_j(K)_X \leq (C_0 + \varepsilon) C_1 2^\mu 16^s \sqrt{\log(2j)} \cdot j^{(\mu - s)}
\]
(3.13)

for \( 2^n \leq j < 2^{n+1} \). Since \( \varepsilon > 0 \) arbitrary we get (3.1).

\textit{Step 3.} We comment on the case
\[
0 = h_m = \ldots = h_{m_j-1} < h_{m_j} = \ldots = h_{m_{j-1}} < \cdots < h_{m_L} = \ldots = h_n.
\]

In this situation we proceed as in Step 2, but there is no first term in the product on the right-hand side of (3.9). Note that in case \( h_1 = \ldots = h_n = 0 \), there is no logarithmic factor on the right-hand side of (3.13). The proof is complete. \( \Box \)

\textbf{Remark 3.2.} Comparing with Theorem 2.3 in [13] we found that in the case \( s > 1/2 \) the estimate given in Theorem 3.1 improves the rate of the logarithmic term. Moreover the factor \( \gamma_n(X) \) is replaced by \( \tilde{\gamma}_n(X) \), which is somewhat better, since in general \( \gamma_n(X) \leq \tilde{\gamma}_n(X) \).

In the case of Lebesgue spaces we have the following.

\textbf{Corollary 3.3.} Let \( 1 \leq p \leq \infty \), \( s > \frac{1}{2} - \frac{1}{p} \), and \( K \) be a compact set in \( L_p \). Assume that
\[
d_n(K)_{L_p} \leq C_0 \max(1, n)^{-s}, \quad n \geq 0.
\]
Then we have
\[
\sigma_n(K)_{L_p} \leq C_0 16^s 2^{\frac{1}{2} - \frac{1}{p}} \sqrt{\log(2n)} n^{-s + \frac{1}{2} - \frac{1}{p}} \quad \text{for } n \geq 2.
\]

7
Let $E$ and $X$ be Banach spaces and $B_E$ be the closed unit ball of $E$. As a supplement we study the case $\mathcal{K} \subset T(B_E)$ where $T \in \mathcal{L}(E, X)$ is a compact operator. We shall compare the rate of convergence of $\sigma_n(\mathcal{K})$ with the Kolmogorov widths of $T(B_E)$. In this situation we have the following.

**Theorem 3.4.** Let $X$ be a Banach space and $\mathcal{K}$ be a compact set in $X$. Assume that there exists a compact operator $T \in \mathcal{L}(E, X)$ where $E$ is a reflexive Banach space such that $\mathcal{K} \subset T(B_E)$. Then we have

$$\left( \prod_{k=0}^{3n-1} \sigma_k(\mathcal{K})_X \right)^{1/3n} \leq 3e^2 \Gamma_n(E) \Gamma_n(X) \left( \prod_{k=0}^{n-1} d_k(T) \right)^{1/n}, \quad n \geq 1.$$  

**Proof.** First, note that $T(B_E)$ is a closed set in $X$ since $T$ is a compact operator and $E$ is reflexive. For $n \in \mathbb{N}$ fixed, running the greedy algorithm for $\mathcal{K}$ we get $\{f_0, \ldots, f_{3n-1}\}$ and $V_k = \text{span}\{f_0, \ldots, f_k\}$. We select $e_k \in B_E$ such that $Te_k = f_k$ for $k = 0, \ldots, 3n - 1$. For each $k \in \mathbb{N}$, as a consequence of the Hahn-Banach Theorem, see [5, Corollary 14.13], we can choose $b_k \in X'$ such that $\|b_k\|_{X'} = 1$,

$$\langle V_k, b_k \rangle = 0, \quad \text{and} \quad \langle f_k, b_k \rangle = \langle Te_k, b_k \rangle = \text{dist}(f_k, V_k)_X = \sigma_k(\mathcal{K})_X. \quad (3.14)$$

We define the operators $A \in \mathcal{L}(\ell^3_{2n}, E)$ and $B \in \mathcal{L}(X, \ell^3_{2n})$ by

$$A := \sum_{k=0}^{3n-1} u_k \otimes e_k \quad \text{and} \quad B := \sum_{k=0}^{3n-1} b_k \otimes u_k, \quad (3.15)$$

where $\{u_k\}_{k=0}^{3n-1}$ is the canonical basis of $\ell^3_{2n}$. We calculate the norm $\|B|B_2\|$, see the definition (2.2). Let $x_1, \ldots, x_N \in X$. We have

$$\left( \sum_{i=1}^{N} \|Bx_i\|_{\ell^3_{2n}}^2 \right)^{1/2} = \left( \sum_{i=1}^{N} \left( \sum_{k=0}^{3n-1} \langle x_i, b_k \rangle u_k \right)^2 \right)^{1/2} = \left( \sum_{i=1}^{N} \sum_{k=0}^{3n-1} |\langle x_i, b_k \rangle|^2 \right)^{1/2}$$

which implies

$$\left( \sum_{i=1}^{N} \|Bx_i\|_{\ell^3_{2n}}^2 \right)^{1/2} \leq \sqrt{3n} \sup_{k=0, \ldots, 3n-1} \left( \sum_{i=1}^{N} |\langle x_i, b_k \rangle|^2 \right)^{1/2} \leq \sqrt{3n} \sup_{b \in B_{X'}} \left( \sum_{i=1}^{N} |\langle x_i, b \rangle|^2 \right)^{1/2}.$$

Hence $\|B|B_2\| \leq \sqrt{3n}$. We consider the matrix $(Te_k, b_j) = (\langle BTAu_k, u_j \rangle)$ which has the lower triangular form. It follows from (3.14) that

$$\left( \prod_{k=0}^{3n-1} \sigma_k(\mathcal{K})_X \right)^{1/3n} = \left( \prod_{k=0}^{3n-1} |\langle Te_k, b_k \rangle| \right)^{1/3n} = |\text{det} \left( \langle BTAu_i, u_j \rangle \right)|^{1/3n}.$$

Note that for any operator $S \in \mathcal{L}(\ell^3_{2})$ we have $|\det S| \leq \prod_{k=0}^{n-1} d_k(S)$ since Kolmogorov widths equal to singular values of $S$, see [9]. Consequently we obtain

$$\left( \prod_{k=0}^{3n-1} \sigma_k(\mathcal{K})_X \right)^{1/3n} \leq \left( \prod_{k=0}^{n-1} d_k(BTA) \right)^{1/3n} = \left( \prod_{k=0}^{n-1} d_{3k}(BTA) \prod_{k=0}^{n-1} d_{3k+1}(BTA) \prod_{k=0}^{n-1} d_{3k+2}(BTA) \right)^{1/3n}. \quad (3.16)$$
From the property
\[ d_{m+n+k}(BTA) \leq d_m(B)d_n(T)d_k(A), \]
see [12, Page 32] or [8, Theorem 11.9.2], and the monotonicity \( d_{k+1} \leq d_k \) of Kolmogorov widths we conclude that
\[
\left( \prod_{k=0}^{3n-1} \sigma_k(K)_X \right)^{1/3n} \leq \left( \prod_{k=0}^{n-1} d_k(A) \right)^{1/n} \left( \prod_{k=0}^{n-1} d_k(T) \right)^{1/n} \left( \prod_{k=0}^{n-1} d_k(B) \right)^{1/n}. \tag{3.17}
\]

Lemmas 2.2 and 2.3 yield the estimate
\[
\left( \prod_{k=0}^{n-1} d_k(B) \right)^{1/n} \leq \Gamma_n(B) \leq en^{-1/2} \|B|B_2\| \Gamma_n(X) \leq en^{-1/2} (3n)^{1/2} \Gamma_n(X) = e\sqrt{3} \Gamma_n(X). \tag{3.18}
\]

Now we deal with the first product on the right-hand side of (3.17). We have
\[
\left( \prod_{k=0}^{n-1} d_k(A) \right)^{1/n} = \left( \prod_{k=0}^{n-1} d^k(A') \right)^{1/n} \leq \Gamma_n(A'),
\]
see Section 2. Here \( A' \in \mathcal{L}(E', \ell_2^n) \) is the dual operator of \( A \) which is of the form \( A' = \sum_{k=0}^{3n-1} e_k \otimes u_k \). Similar argument as for the operator \( B \) we also get \( \|A'|B_2\| \leq \sqrt{3}n \). Hence we found
\[
\Gamma_n(A') \leq en^{-1/2} \|A'|B_2\| \Gamma_n(E') \leq en^{-1/2} (3n)^{1/2} \Gamma_n(E') = e\sqrt{3} \Gamma_n(E')
\]
which leads to
\[
\left( \prod_{k=0}^{n-1} d_k(A) \right)^{1/n} \leq e\sqrt{3} \Gamma_n(E') = e\sqrt{3} \Gamma_n(E).
\]

Putting this and (3.18) into (3.17) we arrive at
\[
\left( \prod_{k=0}^{3n-1} \sigma_k(K)_X \right)^{1/3n} \leq 3e^2 \Gamma_n(E) \Gamma_n(X) \left( \prod_{k=0}^{n-1} d_k(T) \right)^{1/n}.
\]

The proof is complete. \( \square \)

**Remark 3.5.** Note that Theorem 3.4 still holds true if one replaces Kolmogorov widths by Gelfand widths, i.e.,
\[
\left( \prod_{k=0}^{3n-1} \sigma_k(K)_X \right)^{1/3n} \leq 3e^2 \Gamma_n(E) \Gamma_n(X) \left( \prod_{k=0}^{n-1} d_k(T) \right)^{1/n}, \quad n \geq 1.
\]

We have the following consequence.

**Corollary 3.6.** Let \( X \) be a Banach space and \( K \) be a compact set in \( X \). Assume that there exists a compact operator \( T \in \mathcal{L}(\ell_2, X) \) such that \( K \subset T(B_{\ell_2}) \). Then we have
\[
\sigma_{3n-1}(K)_X \leq 3e^2 \Gamma_n(X) \left( \prod_{k=0}^{n-1} d_k(T) \right)^{1/n}, \quad n \geq 1. \tag{3.19}
\]

In addition, if \( X = L_p \) for \( 1 \leq p \leq \infty \) and \( d_n(T) \leq C_0 n^{-s} \) for some \( s > \left| \frac{1}{2} - \frac{1}{p} \right|, \quad (n \geq 1) \), then there exists a constant \( C > 0 \) such that
\[
\sigma_{3n-1}(K)_X \leq C 3e^2 n^{1/p - 1/2 - s}, \quad n \geq 1.
\]
We proceed by considering the case $\mathcal{K} = T(B_{\ell_2})$ for some compact operator $T \in \mathcal{L}(\ell_2, X)$. In this situation we can replace $\sigma_{3n-1}(\mathcal{K})_X$ in (3.19) by $\sigma_{2n-1}(\mathcal{K})_X$. We have the following.

**Theorem 3.7.** Let $X$ be a Banach space and $T \in \mathcal{L}(\ell_2, X)$ be a compact operator. Assume that $\mathcal{K} = T(B_{\ell_2})$. Then we have

$$
\sigma_{2n-1}(\mathcal{K})_X \leq e\sqrt{2\Gamma_n(X)} \left( \prod_{k=0}^{n-1} d_k(\mathcal{K})_X \right)^{1/n}, \quad n \geq 1.
$$

**Proof.** Recall that $T(B_{\ell_2})$ is the closed set in $X$. For $n \in \mathbb{N}$ fixed, running the greedy algorithm for $\mathcal{K}$ we get $\{f_0, \ldots, f_{2n-1}\}$ and $V_k = \text{span}\{f_0, \ldots, f_{k-1}\}$. First we show that we can select $e_k \in B_{\ell_2}$ such that $Te_k = f_k$ for $k = 0, \ldots, 2n-1$ and $\{e_k\}_{k=0}^{2n-1}$ is an orthonormal system in $\ell_2$. Indeed, if $Te_0 = f_0$ with

$$
\|f_0\|_X = \max_{f \in \mathcal{K}} \|f\|_X = \max_{e \in B_{\ell_2}} \|Te\|_X,
$$

then $\|e_0\|_{\ell_2} = 1$. Assume that we have chosen the orthonormal system $\{e_0, \ldots, e_{k-1}\}$ in $\ell_2$ with $Te_i = f_i$ for $i = 0, \ldots, k-1$. Let $\{e_j\}_{j \geq 0}$ be an orthonormal basis of $\ell_2$ constructed from the system $\{e_0, \ldots, e_{k-1}\}$ and let $f_k = Te$ where $e = \sum_{j \geq 0} c_je_j$ with $\|\{c_j\}_{j \geq 0}\|_{\ell_2} \leq 1$. We consider

$$
f_k^* = \frac{1}{\|c^*\|_{\ell_2}} \sum_{j \geq k} c_jTe_j, \quad \text{with} \quad c^* = (0, \ldots, 0, c_k, c_{k+1}, \ldots).
$$

Here we assume that $c^* \neq 0$ otherwise $\sigma_k(\mathcal{K})_X = 0$. We have $f_k^* \in T(B_{\ell_2})$ and

$$
\text{dist}(f_k, V_k) \geq \text{dist}(f_k^*, V_k)_X
$$

$$
= \frac{k}{\|c^*\|_{\ell_2}} \inf_{a_0, \ldots, a_{k-1}} \left\| \sum_{j \geq 0} c_jTe_j - \sum_{j = 0}^{k-1} a_jTe_j \right\|_X
$$

$$
= \frac{1}{\|c^*\|_{\ell_2}} \inf_{c^* \neq 0} \left\| \sum_{j \geq 0} c_jTe_j - \sum_{j = 0}^{k-1} (a_j\|c^*\|_{\ell_2} + c_j)Te_j \right\|_X
$$

$$
= \frac{1}{\|c^*\|_{\ell_2}} \text{dist}(f_k, V_k).
$$

Hence $\|c^*\|_{\ell_2} = 1$ and $f_k = f_k^*$ which implies $e$ is orthogonal to $\{e_0, \ldots, e_{k-1}\}$ and $\|e\|_{\ell_2} = 1$. Similar to (3.15) we define the operators $A \in \mathcal{L}(\ell_2^{2n}, \ell_2)$ and $B \in \mathcal{L}(X, \ell_2^{2n})$ by

$$
A := \sum_{k=0}^{2n-1} u_k \otimes e_k \quad \text{and} \quad B := \sum_{k=0}^{2n-1} b_k \otimes u_k.
$$

Note that $\|B|B_x\| \leq \sqrt{2n}$ and $\|A\| \leq 1$. We have

$$
\left( \prod_{k=0}^{2n-1} \sigma_k(\mathcal{K})_X \right)^{1/2n} \leq \left( \prod_{k=0}^{2n-1} d_k(B(A)) \right)^{1/2n}
$$

$$
\leq \left( \prod_{k=0}^{2n-1} \|A\|d_k(BT) \right)^{1/2n} \leq \left( \prod_{k=0}^{2n-1} d_k(BT) \right)^{1/2n},
$$

see (3.16). By the same argument as in the proof of Theorem 3.4 we obtain the desired estimate. \qed
In some situations, the estimate given in Corollary 3.6 is sharp. Let us consider the following example which is borrowed from [3], see also [13]. Let \( \mathcal{K} = \{ n^{-\alpha}u_n \} \subset \ell_q \) with \( 2 < q < \infty \) where \( \{ u_n \}_{n \geq 1} \) is the canonical basis of \( \ell_2 \). It is clear that \( \sigma_n(\mathcal{K})_{\ell_q} = \frac{1}{(n+1)^{\frac{1}{q}}} \). We consider the diagonal operator \( D_\alpha : \ell_2 \to \ell_q \) defined by \( u_n \to n^{-\alpha}u_n \). Then \( \mathcal{K} \subset D_\alpha(B_\ell_2) \). We know that \( d_n(D_\alpha) \leq Cn^{-\alpha + \frac{1}{q} - \frac{1}{2}} \), see, e.g., [11, Section 6.2.5.3] which implies
\[
\sigma_{3n-1}(\mathcal{K})_{\ell_q} \leq Cn^{\frac{1}{2} - \frac{1}{q}}n^{-\alpha + \frac{1}{q} - \frac{1}{2}} = Cn^{-\alpha}.
\]
Hence the operator \( D_\alpha \) give the sharp estimate in the rate of convergence in this example.

Acknowledgements The author would like to thank Markus Bachmayr for fruitful discussions. Moreover, the author acknowledges the Hausdorff Center of Mathematics, University of Bonn for financial support.

References

[1] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis methods, SIAM J. Math. Anal. 43 (2011), 1457 – 1472.

[2] A. Buffa, Y. Maday, A.T. Patera, C. Prudhomme, and G. Turinici, A priori convergence of the greedy algorithm for the parametrized reduced basis, Model. Math. Anal. Numer. 46 (2012), 595 – 603.

[3] R. DeVore, G. Petrova, and P. Wojtaszczyk, Greedy algorithms for reduced bases in Banach spaces, Constr. Approx. 37 (2013), 455 – 466.

[4] G.H. Hardy, J.E. Littlewood, and G. Pólya, Inequalities, Cambridge University Press, London, 1934.

[5] E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer, Berlin, 1969.

[6] Y. Maday, A. T. Patera, and G. Turinici, A priori convergence theory for reduced-basis approximations of single-parametric elliptic partial differential equations, J. Sci. Comput. 17 (2002), 437 – 446.

[7] ———, Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations, C. R. Acad. Sci., Paris, Ser. I, Math 335 (2002), 289 –294.

[8] A. Pietsch, Operator Ideals, North-Holland, Amsterdam, 1980.

[9] ———, Weyl Numbers and Eigenvalues of Operators in Banach Spaces, Math. Ann. 247 (1980), 149 – 168.

[10] ———, Eigenvalue distributions and geometry of Banach space, Math. Nachr. 160 (1991), 41 – 81.

[11] ———, History of Banach Spaces and Linear Operators, Birkhäuser, Boston, 2007.

[12] A. Pinkus, n-Widths in Approximation Theory, Springer, Berlin, 1985.

[13] P. Wojtaszczyk, On greedy algorithm approximating Kolmogorov widths in Banach spaces, J. Math. Anal. Appl. 424 (2015), 685 – 695.