Universität Stuttgart

Some New Bounds on the Entropy Numbers of Diagonal Operators

Simon Fischer

Stuttgarter Mathematische Berichte

2019-004

mathematik

Fachbereich Mathematik
Fakultät Mathematik und Physik
Universität Stuttgart
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: preprints@mathematik.uni-stuttgart.de
WWW: http://www.mathematik.uni-stuttgart.de/preprints

ISSN **1613-8309**

Some New Bounds on the Entropy Numbers of Diagonal Operators

Simon Fischer

March 1, 2019
Institute for Stochastics and Applications
Faculty 8: Mathematics and Physics
University of Stuttgart
D-70569 Stuttgart Germany
simon.fischer@mathematik.uni-stuttgart.de

Abstract

Entropy numbers are an important tool for quantifying the compactness of operators. Besides establishing new upper bounds on the entropy numbers of diagonal operators D_σ from ℓ_p to ℓ_q, where $p \neq q$, we investigate the optimality of these bounds. In case of $p < q$ optimality is proven for fast decaying diagonal sequences, which include exponentially decreasing sequences. In case of $p > q$ we show optimality under weaker assumption than previously used in the literature. In addition, we illustrate the benefit of our results with examples not covered in the literature so far.

Keywords Diagonal Operators, Entropy Numbers

1. Introduction and Main Results

For $1 \leq p, q \leq \infty$ and a non-increasing sequence $\sigma = (\sigma_k)_{k\geq 1}$ we write $D_\sigma : \ell_p \to \ell_q$ for the diagonal operator between the sequence spaces ℓ_p and ℓ_q, i.e. $D_\sigma(x_k)_{k\geq 1} := (\sigma_k x_k)_{k\geq 1}$. If we denote the closed unit ball of ℓ_p by B_{ℓ_p} then the entropy numbers of the operator $D_\sigma : \ell_p \to \ell_q$ are defined by

$$\varepsilon_n(D_\sigma) := \inf\Big\{\varepsilon > 0 : \ \exists y_1, \ldots, y_n \in \ell_q \text{ with } D_\sigma B_{\ell_p} \subseteq \bigcup_{i=1}^{n} y_i + \varepsilon B_{\ell_q}\Big\}$$

for all $n \geq 1$. In case of $p = q$ Gordon et al. [8, Proposition 1.7] give a complete description of the asymptotic behavior of the entropy numbers $\varepsilon_n(D_\sigma)$ for *all* diagonal sequences σ. In case of $p \neq q$ – as far as we know – there are only partial answers, see e.g. [11, 12, 4]. The present work is a further contribution to this problem: Our first theorem fills a gap in the literature by providing an upper bound in case of $p < q$, which is optimal for sequences that decay at least exponentially in the sense of (EXP). The second theorem considers the case $p > q$ and gives an upper bound, which is optimal for sequences that decrease at least polynomially in the sense of (ALP) as well as for sequences that decrease at most polynomially in the sense of (AMP). For the second type of sequences this recovers the optimal bound of Kühn [12], while the first type of sequences have not been considered so far. A more detailed comparison between our results and existing bounds can be found at the end of this section. The proofs of both our theorems combine the ideas of Gordon et al. [8, Proposition 1.7] and Oloff [15, Hilfsatz 2]. Moreover, in the appendix we summarize relations between the regularity conditions on σ we consider and some other common regularity conditions.

Before we proceed let us introduce some notation. For real sequences $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$ we write $x_n \preccurlyeq y_n$ iff there is a constant $c > 0$ with $x_n \leq c y_n$ for all $n \geq 1$ and $x_n \asymp y_n$ iff $x_n \preccurlyeq y_n$ as well as $x_n \succcurlyeq y_n$ hold. In the following, we declare an upper or lower bound $(x_n)_{n\geq 1}$ on the entropy numbers to be *optimal* if there is a corresponding lower resp. upper bound $(y_n)_{n\geq 1}$ with $x_n \asymp y_n$.

1.1 Theorem (Bound for $p < q$) *Let $1 \leq p < q \leq \infty$ with $\frac{1}{p} = \frac{1}{q} + \frac{1}{s}$ and $\sigma = (\sigma_k)_{k\geq 1}$ be a sequence with $\sigma_k > 0$ and $\sigma_k \searrow 0$. Then the entropy numbers of the diagonal operator $D_\sigma : \ell_p \to \ell_q$ satisfy*

$$\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\geq 1} k^{-1/s} \left(\frac{(\sigma_1 + k^{1/s}\sigma_k) \cdot \ldots \cdot (\sigma_k + k^{1/s}\sigma_k)}{n} \right)^{1/k}. \tag{1}$$

If, in addition, there is a real number $b > 1$ with

$$\sup_{k\leq n} \frac{\sigma_n b^n}{\sigma_k b^k} < \infty \tag{EXP}$$

then the bound in (1) is optimal and coincides with

$$\varepsilon_n(D_\sigma) \asymp \sup_{k\geq 1} k^{-1/s} \left(\frac{\sigma_1 \cdot \ldots \cdot \sigma_k}{n} \right)^{1/k}.$$

Note that the supremum in Equation (EXP) is taken over all tuples $(n, k) \in \mathbb{N}^2$ with $k \leq n$. Moreover, Condition (EXP) implies $\sigma_n \preccurlyeq b^{-n}$ and is independent of p and q.

To treat the case $p > q$ we recall that the diagonal operator D_σ is well-defined if and only if $\sigma \in \ell_r$ with $\frac{1}{q} = \frac{1}{p} + \frac{1}{r}$. For this reason we restricted our considerations in this case to $\sigma \in \ell_r$ and define the *tail sequence* for $k \geq 1$

$$\tau_k := \Big(\sum_{n=k}^{\infty} \sigma_n^r \Big)^{1/r}. \tag{2}$$

1.2 Theorem (Bound for $p > q$) *Let $1 \leq q < p \leq \infty$ with $\frac{1}{q} = \frac{1}{p} + \frac{1}{r}$ and $\sigma = (\sigma_k)_{k\geq 1} \in \ell_r$ be a sequence with $\sigma_k > 0$ and $\sigma_k \searrow 0$. Then the entropy numbers of the diagonal operator $D_\sigma : \ell_p \to \ell_q$ satisfy*

$$\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\geq 1} \left(\frac{(\tau_k + k^{1/r}\sigma_1) \cdot \ldots \cdot (\tau_k + k^{1/r}\sigma_k)}{n} \right)^{1/k}. \tag{3}$$

Moreover, under each of the following additional assumptions the bound in (3) is optimal:

(i) Assumption (ALP): $\tau_n \preccurlyeq \sigma_n n^{1/r}$. In this case the bound in (3) coincides with

$$\varepsilon_n(D_\sigma) \asymp \sup_{k\geq 1} k^{1/r} \left(\frac{\sigma_1 \cdot \ldots \cdot \sigma_k}{n} \right)^{1/k}.$$

(ii) Assumption (AMP): $\tau_n \succcurlyeq \sigma_n n^{1/r}$. In this case the bound in (3) *coincides with*

$$\varepsilon_n(D_\sigma) \asymp \tau_{\lfloor \log_2(n) \rfloor + 1}.$$

According to Lemma A.3 (i) the Condition (ALP) implies $\sigma_n \preccurlyeq n^{-\alpha}$ for some $\alpha > 1/r$. Moreover, Lemma A.3 (ii) and Lemma A.2 say that the Condition (AMP) is equivalent to $\tau_n \asymp \tau_{2n}$ and that this implies $\tau_n \succcurlyeq n^{-\alpha}$ for some $\alpha > 0$. Furthermore, if we combine Lemma A.1 (iv) with (b) and (d) of Lemma A.3 we get (EXP)⊆(ALP) resp. (EXP)∩(AMP)= ∅.

Let us now compare our results to the bounds previously obtained in the literature. Since essentially all previously established results on the entropy (or covering) numbers of D_σ, see e.g. [9, 14, 13, 15, 3, 10] and the references therein, are contained in [11, 12, 4], we restrict our comparison to the latter three articles.

In case of $p < q$ the most general entropy bounds are derived by Kühn in [11]. Namely, he obtained optimal bounds under each of the following set of assumptions:

(i) polynomial: $\sup_{k \le n} \frac{\sigma_n n^\alpha}{\sigma_k k^\alpha} < \infty$ for some $\alpha > 0$ and $\sigma_n \asymp \sigma_{2n}$,

(ii) fast logarithmic: $\sup_{k \le n} \frac{\sigma_n}{\sigma_k} \big(\frac{1+\log n}{1+\log k}\big)^{1/s} < \infty$ and $\sigma_{n^2} \asymp \sigma_n$,

(iii) slow logarithmic: $\inf_{k \le n} \frac{\sigma_n}{\sigma_k} \big(\frac{1+\log n}{1+\log k}\big)^{1/s} > 0$.

Note that Scenario (i) and (ii) both exclude sequences that decrease too slow as well as sequences that decrease too fast. In contrast, (iii) only excludes sequences that decrease too fast. In comparison, the optimal bounds we obtain in Theorem 1.1 require sequences that decay at least exponentially in the sense of (EXP). Since all of the Scenarios (i)–(iii) imply $\sigma_n \asymp \sigma_{2n}$, we easily see that they all exclude (EXP), that is, (EXP) is not covered by the results in [11].

In case of $p > q$, [11] also provides optimal bounds for sequences σ satisfying $\sigma_n \asymp \sigma_{2n}$ and

$$\sup_{k \le n} \frac{\sigma_n n^\alpha}{\sigma_k k^\alpha} < \infty$$

for some $\alpha > 1/r$. According to Lemma A.3 the combination of both assumptions is equivalent to the combination of (AMP) *and* (ALP), i.e. $\tau_n \asymp \sigma_n n^{1/r}$. In [12], Kühn generalizes the results of [11] by establishing optimal bounds under Assumption (AMP), only. Consequently, Theorem 1.2 recovers the upper bounds of [12] and additionally provides optimal bounds for σ that only satisfy (ALP).

Table 1 lists three types of sequences σ that are not covered by the literature, but for which we obtain optimal bounds. Compared to [11, 12], another advantage of our results is that they actually provide bounds for *all* $1 \le p \ne q \le \infty$ and *all* sequences σ. However, in some cases the question of optimality is not answered yet.

Finally, there is another strand of research, see e.g. [3, 4], that describes the asymptotic of the entropy numbers in terms of *(generalized) Lorentz spaces*. The most general result in this direction is [4, Corollary 1.2]:

$$\sigma \in \ell_{t,v,\varphi} \quad \Longleftrightarrow \quad \varepsilon_{2^{n-1}}(D_\sigma) \in \ell_{u,v,\varphi}\,,$$

where $\ell_{u,v,\varphi}$ is a generalized Lorentz space with slowly varying function φ, see [4, Section 2] for a definition, and the parameters satisfy $1 \le p, q \le \infty$, $0 < t, v \le \infty$, $1/t > (1/q - 1/p)_+$, and

$\sigma_n \asymp$	$\tau_n \asymp$	(AMP)	(ALP)	(EXP)
$\exp\bigl(-a\log^{\lambda}(n)\bigr)$	$\sigma_n\, n^{1/r}\log^{(1-\lambda)/r}(n)$	no	yes if $\lambda > 1$	no
$\exp\bigl(-an^{\lambda}\bigr)$	$\sigma_n\, n^{(1-\lambda)_+/r}$	no	yes	yes if $\lambda \geq 1$
$\exp\bigl(-ae^{\lambda n}\bigr)$	σ_n	no	yes	yes

Table 1: Three types of sequences for which our results provide optimal bounds and which are not covered by the existing literature. For all examples we assume $a > 0$ and $\lambda > 0$. In addition, the conditions (AMP) and (ALP) are only considered in the case $p > q$, whereas (EXP) is actually independent of p and q. Note some subtleties of the first example: For $\lambda = 1$ it reduces to a plain polynomial decay, which is already well understood. Moreover, for $\lambda < 1$ the operator D_σ is not even bounded in case of $p > q$. Finally, for $\lambda < 1$ and $p < q$, Kühn [11] leaves the behavior of $\varepsilon_n(D_\sigma)$ as an open question, which our results cannot address, either.

$1/u = 1/t - (1/q - 1/p)$. Note that ($\Leftarrow$) is contained in Lemma 2.3 and ($\Rightarrow$) is contained in Theorem 1.2 if $p > q$ and $v = \infty$.

Acknowledgment

I am especially grateful to Prof. Dr. Ingo Steinwart for carefully proofreading preliminary versions of this manuscript and pointing out some errors.

2. Proofs

2.1. Preliminaries

Before we prove the main theorems we summarize some preparatory results. Because we will reduce the investigation of diagonal operators to the case of finite dimensional diagonal operators on $\mathbb{R}^k$ we will include this case in the following. To this end, we consider sequences over an index set $I \subseteq \mathbb{N}$ and define, for $1 \leq p \leq \infty$, the sequence space $\ell_p(I) := \{x = (x_i)_{i\in I} \in \mathbb{R}^I : \ \|x\|_{\ell_p(I)} < \infty\}$ with norm

$$\|x\|_{\ell_p(I)} := \left(\sum_{i\in I} |x_i|^p\right)^{1/p}$$

and closed unit ball $B_{\ell_p(I)}$. With this notation we have $\ell_p = \ell_p(\mathbb{N})$ and for $k \geq 1$ we introduce the abbreviation $\ell_p^k := \ell_p(\{1, \ldots, k\})$. In the following, we fix $1 \leq p, q \leq \infty$, a sequence $\sigma = (\sigma_i)_{i\in I} \in \mathbb{R}^I$, and the corresponding diagonal operator $D_\sigma : \ell_p(I) \to \ell_q(I)$ defined by $D_\sigma(x_i)_{i\in I} := (\sigma_i x_i)_{i\in I}$. As a consequence of Hölder's inequality the operator norm of D_σ satisfies

$$\|D_\sigma\| = \begin{cases} \|\sigma\|_{\ell_r(I)}, & p > q,\ \frac{1}{q} = \frac{1}{p} + \frac{1}{r} \\ \|\sigma\|_{\ell_\infty(I)}, & p \leq q\ . \end{cases} \tag{4}$$

Next, we introduce some concepts related to entropy numbers. For $\varepsilon > 0$ the *covering number* of D_σ is defined by

$$\mathcal{N}(D_\sigma, \varepsilon) := \min\Big\{n \geq 1 : \ \exists y_1, \ldots, y_n \in \ell_q(I) \text{ with } D_\sigma B_{\ell_p(I)} \subseteq \bigcup_{i=1}^{n} y_i + \varepsilon B_{\ell_q(I)}\Big\}. \tag{5}$$

The next result establishes a comparison between covering and entropy numbers.

2.1 Lemma *Let $1 \leq p, q \leq \infty$, $(a_k)_{k\geq 1}$ be a positive sequence and $D_\sigma : \ell_p \to \ell_q$ be a diagonal operator with $\|D_\sigma\| < \infty$. If we have the covering number estimate*

$$\mathcal{N}(D_\sigma, \varepsilon) \leq \sup_{k\geq 1} a_k \Big(\frac{1}{\varepsilon}\Big)^k$$

for all $0 < \varepsilon < \|D_\sigma\|$, then for all $n \geq 1$ the n-th entropy number satisfies

$$\varepsilon_n(D_\sigma) \leq \sup_{k\geq 1} \Big(\frac{a_k}{n}\Big)^{1/k}.$$

Proof. Let $n \geq 1$ be a natural number. In case of $\varepsilon_n(D_\sigma) = 0$ there is nothing to prove. Hence we assume $\varepsilon_n(D_\sigma) > 0$ and choose $0 < \varepsilon < \varepsilon_n(D_\sigma)$. By the definition of entropy and covering numbers we have $n < \mathcal{N}(D_\sigma, \varepsilon)$. Moreover, by our assumption there is, for every $\delta > 0$, a $k_\delta \geq 1$ with

$$n \leq \mathcal{N}(D_\sigma, \varepsilon) \leq (1 + \delta)\, a_{k_\delta} \Big(\frac{1}{\varepsilon}\Big)^{k_\delta}.$$

This implies

$$\varepsilon \leq \Big(\frac{(1 + \delta)\, a_{k_\delta}}{n}\Big)^{1/k_\delta} \leq (1 + \delta) \Big(\frac{a_{k_\delta}}{n}\Big)^{1/k_\delta} \leq (1 + \delta) \sup_{k\geq 1} \Big(\frac{a_k}{n}\Big)^{1/k}.$$

Letting $\delta \searrow 0$ and $\varepsilon \nearrow \varepsilon_n(D_\sigma)$ we get the assertion. □

In the following, λ^k denotes the k-dimensional Lebesgue measure.

2.2 Lemma *Let $1 \leq p, q \leq \infty$, $k \geq 1$ and $\sigma_1, \ldots, \sigma_k > 0$. Then for all $\varepsilon > 0$ the diagonal operator $D_\sigma : \ell_p^k \to \ell_q^k$ satisfies*

$$\mathcal{N}(D_\sigma, 2\varepsilon) \leq 2^k \frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})} \Big(\|\operatorname{id}_{q,p}^k\| + \frac{\sigma_1}{\varepsilon}\Big) \cdot \ldots \cdot \Big(\|\operatorname{id}_{q,p}^k\| + \frac{\sigma_k}{\varepsilon}\Big), \tag{6}$$

where $\operatorname{id}_{q,p}^k : \ell_q^k \to \ell_p^k$ denotes the identity operator.

In case of $p = q$ the bound in (6) originates from Oloff [15, Hilfsatz 2]. Furthermore, note that the proof of Kolmogorov and Tikhomirov [9, Theorem XVI] contains the case $p = q = 2$ and $\sigma_n = n^{-\alpha}$.

Proof. For this proof we use *packing numbers*, which for $\varepsilon > 0$ are defined by

$$\mathcal{P}(D_\sigma, \varepsilon) := \max\Big\{n \geq 1 : \ \exists y_1, \ldots, y_n \in D_\sigma B_{\ell_p^k} \text{ with } \|y_i - y_j\|_{\ell_q^k} > 2\varepsilon \ \forall i \neq j\Big\}.$$

Recall from [9, Theorem IV] that $\mathcal{P}(D_\sigma, 2\varepsilon) \le \mathcal{N}(D_\sigma, 2\varepsilon) \le \mathcal{P}(D_\sigma, \varepsilon)$ holds for all $\varepsilon > 0$. Therefore it is enough to prove that $\mathcal{P}(D_\sigma, \varepsilon)$ is bounded by the right hand side of (6).

Now, for $\varepsilon > 0$ and $n := \mathcal{P}(D_\sigma, \varepsilon)$ we choose $x_1, \ldots, x_n \in D_\sigma B_{\ell_p^k}$ with $\|x_i - x_j\|_{\ell_q^k} > 2\varepsilon$ for all $i \neq j$. Then $x_i + \varepsilon B_{\ell_q^k}$ are disjoint sets contained in $D_\sigma B_{\ell_p^k} + \varepsilon B_{\ell_q^k}$. Hence their volume satisfies

$$n\varepsilon^k \lambda^k(B_{\ell_q^k}) = \lambda^k\Big(\bigcup_{i=1}^n (x_i + \varepsilon B_{\ell_q^k})\Big) \le \lambda^k(D_\sigma B_{\ell_p^k} + \varepsilon B_{\ell_q^k}). \tag{7}$$

Before we continue to estimate (7) we prove the following auxiliary result: For a second diagonal operator $D_\omega : \ell_p^k \to \ell_q^k$ with $\omega_i > 0$ for all $i = 1, \ldots, k$ we have

$$D_\sigma B_{\ell_p^k} + D_\omega B_{\ell_p^k} \subseteq 2D_{\sigma+\omega} B_{\ell_p^k}. \tag{8}$$

Note that since $D_{\sigma+\omega}$ is invertible (8) is equivalent to $D_{\sigma+\omega}^{-1}(D_\sigma B_{\ell_p^k} + D_\omega B_{\ell_p^k}) \subseteq 2B_{\ell_p^k}$. Now, to show (8) we fix $x, y \in B_{\ell_p^k}$ and observe

$$\|D_{\sigma+\omega}^{-1}(D_\sigma x + D_\omega y)\|_{\ell_p^k} \le \|D_{\sigma+\omega}^{-1} D_\sigma x\|_{\ell_p^k} + \|D_{\sigma+\omega}^{-1} D_\omega y\|_{\ell_p^k} \le \|D_{\sigma+\omega}^{-1} D_\sigma\| + \|D_{\sigma+\omega}^{-1} D_\omega\|.$$

Since $D_{\sigma+\omega}^{-1} D_\sigma$ is an operator from ℓ_p^k to ℓ_p^k the operator norm is given by $\|D_{\sigma+\omega}^{-1} D_\sigma\| = \max_{i=1,\ldots,k} \frac{\sigma_i}{\sigma_i+\omega_i} \le 1$. Analogously we have $\|D_{\sigma+\omega}^{-1} D_\omega\| = \max_{i=1,\ldots,k} \frac{\omega_i}{\sigma_i+\omega_i} \le 1$ and therefore (8) is proven.

By the definition of the operator norm we have $B_{\ell_q^k} \subseteq \|\operatorname{id}_{q,p}^k\| B_{\ell_p^k}$. Together with (8) we get

$$D_\sigma B_{\ell_p^k} + \varepsilon B_{\ell_q^k} \subseteq D_\sigma B_{\ell_p^k} + \varepsilon \|\operatorname{id}_{q,p}^k\| B_{\ell_p^k} \subseteq 2D_{\sigma+\varepsilon\|\operatorname{id}_{q,p}^k\|} B_{\ell_p^k}.$$

Continuing estimate (7) with this inclusion yields (6). □

2.2. Entropy Bounds

In this subsection we provide lower and upper bounds on the entropy numbers. To this end, we define, for $k \ge 1$, the auxiliary operators

$$\begin{aligned} D_{p,q}^k &: \ell_p^k \to \ell_q^k,\ (x_n)_{n=1}^k \mapsto (\sigma_1 x_1, \ldots, \sigma_k x_k), \\ P_p^k &: \ell_p \to \ell_p^k,\ (x_n)_{n\ge1} \mapsto (x_1, \ldots, x_k), \\ I_p^k &: \ell_p^k \to \ell_p,\ (x_n)_{n=1}^k \mapsto (x_1, \ldots, x_k, 0, 0, \ldots). \end{aligned}$$

Note that these operators satisfy $D_{p,q}^k = P_q^k D_\sigma I_p^k$ and $\|I_p^k\| = \|P_p^k\| = 1$.

2.3 Lemma (Lower Bound) *Let $1 \le p, q \le \infty$ and $\sigma = (\sigma_k)_{k\ge1}$ with $\sigma_k > 0$ and $\sigma_k \searrow 0$ such that the diagonal operator $D_\sigma : \ell_p \to \ell_q$ is bounded. Then for all $n \ge 1$ we have*

$$\varepsilon_n(D_\sigma) \ge \sup_{k\ge1} \bigg(\frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})} \, \frac{\sigma_1 \cdot \ldots \cdot \sigma_k}{n} \bigg)^{1/k}.$$

Note that this lower bound holds without any additional assumption on σ. Moreover, a combination of Wang [16] with Stirling's formula yields

$$\left(\frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})}\right)^{1/k} \asymp k^{1/q-1/p}. \tag{9}$$

Proof. By the multiplicativity of entropy numbers, see [5, p. 11], we find $\varepsilon_n(D_{p,q}^k) = \varepsilon_n(P_q^k D_\sigma I_p^k) \leq \varepsilon_n(D_\sigma)$, and hence it remains to give a lower bound for $\varepsilon_n(D_{p,q}^k)$. To this end, choose for $\varepsilon > \varepsilon_n(D_{p,q}^k)$ some $x_1, \dots, x_n \in \mathbb{R}^k$ with $D_\sigma B_{\ell_p^k} \subseteq \bigcup_{i=1}^n (x_i + \varepsilon B_{\ell_q^k})$. Consequently, the volume of these sets satisfy

$$\sigma_1 \cdot \ldots \cdot \sigma_k \lambda^k(B_{\ell_p^k}) = \lambda^k(D_\sigma B_{\ell_p^k}) \leq \sum_{i=1}^n \lambda^k(x_i + \varepsilon B_{\ell_q^k}) = n\varepsilon^k \lambda^k(B_{\ell_q^k}),$$

and hence we find

$$\varepsilon \geq \left(\frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})} \frac{\sigma_1 \cdot \ldots \cdot \sigma_k}{n}\right)^{1/k}.$$

Letting $\varepsilon \searrow \varepsilon_n(D_{p,q}^k)$ and taking the supremum over $k \geq 1$ we get the assertion. □

Since the upper bounds in (1) and (3) are based on the same decomposition we first introduce this decomposition. To this end, recall that the covering numbers have an additivity and multiplicativity property analogously to the entropy numbers, see [5, p. 11]. Using these properties yields

$$\begin{aligned}\mathcal{N}(D_\sigma, \varepsilon) = \mathcal{N}\Big(I_q^k D_{p,q}^k P_p^k + (D_\sigma - I_q^k D_{p,q}^k P_p^k), \varepsilon\Big) &\leq \mathcal{N}\Big(I_q^k D_{p,q}^k P_p^k, \varepsilon/2\Big) \cdot \mathcal{N}\Big(D_\sigma - I_q^k D_{p,q}^k P_p^k, \varepsilon/2\Big) \\ &\leq \mathcal{N}\Big(D_{p,q}^k, \varepsilon/2\Big) \cdot \mathcal{N}\Big(D_\sigma - I_q^k D_{p,q}^k P_p^k, \varepsilon/2\Big).\end{aligned}$$

In the following, we will choose a suitable k with $\|D_\sigma - I_q^k D_{p,q}^k P_p^k\| \leq \varepsilon/2$. Since in this case we have $\mathcal{N}(D_\sigma - I_q^k D_{p,q}^k P_p^k, \varepsilon/2) = 1$ the estimate above reduces to

$$\mathcal{N}(D_\sigma, \varepsilon) \leq \mathcal{N}(D_{p,q}^k, \varepsilon/2). \tag{10}$$

Let us first treat the case $p < q$.

2.4 Lemma *Let $1 \leq p < q \leq \infty$ with $\frac{1}{p} = \frac{1}{q} + \frac{1}{s}$ and $\sigma = (\sigma_k)_{k\geq 1}$ with $\sigma_k > 0$ and $\sigma_k \searrow 0$. Then for all $n \geq 1$ the diagonal operator $D_\sigma : \ell_p \to \ell_q$ satisfies*

$$\varepsilon_n(D_\sigma) \leq 4 \sup_{k\geq 1}\left(\frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})} \frac{(2\sigma_1 + k^{1/s}\sigma_k) \cdot \ldots \cdot (2\sigma_k + k^{1/s}\sigma_k)}{n}\right)^{1/k}.$$

Proof. Because of the monotonicity of σ, for every $\varepsilon > 0$ with $\varepsilon < \|D_\sigma\| = \sigma_1$, there is a $k \geq 1$ with $\sigma_{k+1} \leq \varepsilon/2 < \sigma_k$. Equation (4) gives us $\|D_\sigma - I_q^k D_{p,q}^k P_p^k\| = \sigma_{k+1} \leq \varepsilon/2$. Using Equation (10) with this k, Lemma 2.2, and $\|\operatorname{id}_{q,p}^k\| = k^{1/s}$ we get

$$\mathcal{N}(D_\sigma, \varepsilon) \leq \mathcal{N}(D_{p,q}^k, \varepsilon/2) \leq 2^k \frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})}\Big(k^{1/s} + \frac{4\sigma_1}{\varepsilon}\Big) \cdot \ldots \cdot \Big(k^{1/s} + \frac{4\sigma_k}{\varepsilon}\Big).$$

Using $k^{1/s} < \frac{2\sigma_k k^{1/s}}{\varepsilon}$ and taking the supremum over k gives

$$\mathcal{N}(D_\sigma,\varepsilon) \le \sup_{k\ge1}\left\{\frac{\lambda^k(B_{\ell_p^k})}{\lambda^k(B_{\ell_q^k})}\big(\sigma_k k^{1/s}+2\sigma_1\big)\cdot\ldots\cdot\big(\sigma_k k^{1/s}+2\sigma_k\big)\Big(\frac{4}{\varepsilon}\Big)^k\right\}.$$

Finally, Lemma 2.1 yields the assertion. □

2.5 Lemma *Let $1 \le q < p \le \infty$ with $\frac{1}{q} = \frac{1}{p} + \frac{1}{r}$, $\sigma = (\sigma_k)_{k\ge1} \in \ell_r$ with $\sigma_k > 0$ and $\sigma_k \searrow 0$, and τ the tail sequence defined by (2). Then for all $n \ge 1$ the diagonal operator $D_\sigma : \ell_p \to \ell_q$ satisfies*

$$\varepsilon_n(D_\sigma) \le 4\sup_{k\ge1}\left(\frac{(\tau_k+2k^{1/r}\sigma_1)\cdot\ldots\cdot(\tau_k+2k^{1/r}\sigma_k)}{n}\right)^{1/k}.$$

Proof. Because of the monotonicity of τ, for every $0 < \varepsilon < \|D_\sigma\| = \tau_1$, there is a $k \ge 1$ with $\tau_{k+1} \le \varepsilon/2 < \tau_k$. Equation (4) gives us $\|D_\sigma - I_q^k D_{p,q}^k P_p^k\| = \tau_{k+1} \le \varepsilon/2$. Using Equation (10) with this k, the decomposition $D_{p,q}^k = \mathrm{id}_{p,q}^k \circ D_{p,p}^k$, and $\|\mathrm{id}_{p,q}^k\| = k^{1/r}$ we get

$$\mathcal{N}(D_\sigma,\varepsilon) \le \mathcal{N}(D_{p,q}^k,\varepsilon/2) \le \mathcal{N}(D_{p,p}^k,k^{-1/r}\varepsilon/2)\cdot\mathcal{N}(\mathrm{id}_{p,q}^k,k^{1/r}) = \mathcal{N}(D_{p,p}^k,k^{-1/r}\varepsilon/2).$$

Using Lemma 2.2 and $1 < \frac{2\tau_k}{\varepsilon}$ gives

$$\begin{aligned}\mathcal{N}(D_\sigma,\varepsilon) &\le 2^k\Big(1+\frac{4k^{1/r}\sigma_1}{\varepsilon}\Big)\cdot\ldots\cdot\Big(1+\frac{4k^{1/r}\sigma_k}{\varepsilon}\Big)\\ &\le \big(\tau_k+2k^{1/r}\sigma_1\big)\cdot\ldots\cdot\big(\tau_k+2k^{1/r}\sigma_k\big)\Big(\frac{4}{\varepsilon}\Big)^k.\end{aligned}$$

Finally, taking the supremum over k and using Lemma 2.1 gives the assertion. □

2.3. Optimality

Proof of Theorem 1.1. The upper bound in (1) is a consequence of Lemma 2.4 and Equation (9). It remains to prove the optimality under the additional Assumption (EXP). To this end, we continue the estimate of the upper bound as follows

$$\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\ge1} k^{-1/s}\Big(\frac{\sigma_1\cdot\ldots\cdot\sigma_k}{n}\Big)^{1/k}\bigg(\Big(1+\frac{k^{1/s}\sigma_k}{\sigma_1}\Big)\cdot\ldots\cdot\Big(1+\frac{k^{1/s}\sigma_k}{\sigma_k}\Big)\bigg)^{1/k}.$$

Applying that the geometric mean is bounded by the arithmetic mean as well as the triangle inequality in ℓ_s^k (since $s \ge p \ge 1$) yields

$$\bigg(\Big(1+\frac{k^{1/s}\sigma_k}{\sigma_1}\Big)\cdot\ldots\cdot\Big(1+\frac{k^{1/s}\sigma_k}{\sigma_k}\Big)\bigg)^{1/k} \le \bigg(1/k\sum_{i=1}^k\Big(1+\frac{k^{1/s}\sigma_k}{\sigma_i}\Big)^s\bigg)^{1/s} \le 1+\sigma_k\Big(\sum_{i=1}^k\sigma_i^{-s}\Big)^{1/s}.$$

According Lemma A.1 (iii) the right hand side is bounded in k and we get the claimed upper bound. If we combine Lemma 2.3 with Equation (9) we get the corresponding lower bound. □

Proof of Theorem 1.2. The upper bound in (3) directly follows from Lemma 2.5 and it thus remains to prove the optimality under the additional Assumption (i) and (ii).

(i) The upper bound (3) can be transformed into

$$\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\geq 1} k^{1/r}\Big(\frac{\sigma_1\cdot\ldots\cdot\sigma_k}{n}\Big)^{1/k}\Big(\Big(\frac{\tau_k}{k^{1/r}\sigma_1}+1\Big)\cdot\ldots\cdot\Big(\frac{\tau_k}{k^{1/r}\sigma_k}+1\Big)\Big)^{1/k}.$$

Since the last factor is bounded in k according to our additional Assumption (ALP) this yields the claimed upper bound. If we combine Lemma 2.3 with Equation (9) we get the corresponding lower bound.

(ii) Because of Lemma A.3 (ii) we have $\tau_n \asymp \tau_{2n}$. Hence Kühn [12, Theorem 1] yields $\varepsilon_n(D_\sigma) \asymp \tau_{\lfloor \log_2(n)\rfloor+1}$ and it is enough to show that upper bound (3) is asymptotically bounded by $\tau_{\lfloor \log_2(n)\rfloor+1}$. According to (AMP) and Lemma A.2 (iii) applied to $(\tau_n)_{n\geq 1}$ there are constants $c_1, c_2, \beta > 0$ with $\sigma_i \leq c_1\tau_i i^{-1/r}$ and $\tau_i \leq c_2\tau_k k^\beta i^{-\beta}$ for all $k \geq i$. Together we get for $\alpha = 1/r+\beta$

$$\tau_k + k^{1/r}\sigma_i \leq \tau_k + c_1c_2\tau_k\frac{k^{1/r+\beta}}{i^{1/r+\beta}} \leq \tau_k\frac{k^\alpha}{i^\alpha}(1+c_1c_2)$$

and all $k \geq i$. Plugging this into bound (3) we get

$$\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\geq 1}\Big(\frac{(\tau_k+k^{1/r}\sigma_1)\cdot\ldots\cdot(\tau_k+k^{1/r}\sigma_k)}{n}\Big)^{1/k} \preccurlyeq \sup_{k\geq 1}\frac{\tau_k}{n^{1/k}}\frac{k^\alpha}{(k!)^{\alpha/k}}.$$

From Stirling's formula we know $(k!)^{1/k} \asymp k$. Hence we have $\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\geq 1}\frac{\tau_k}{n^{1/k}}$ and it remains to show, that this supremum behaves asymptotically like $\tau_{\lfloor \log_2(n)\rfloor+1}$. To this end, let $c > 0$ be the doubling constant of τ, i.e. $\tau_{2n} \geq c\tau_n$ for all $n \geq 1$. Without loss of generality we can assume $c < 1$ and define $\alpha := \frac{\log(2)}{2\log(1/c)} > 0$. For $k \leq \alpha\log_2(n)$ we have

$$n^{\frac{1}{2k}-\frac{1}{k}} = n^{-\frac{1}{2k}} = \exp\Big(-\frac{\log(n)}{2k}\Big) \leq \exp\Big(-\frac{\log(n)}{2\alpha\log_2(n)}\Big) = \exp(-\log(1/c)) = c \leq \frac{\tau_{2k}}{\tau_k}$$

and this implies

$$\frac{\tau_k}{n^{\frac{1}{k}}} \leq \frac{\tau_{2k}}{n^{\frac{1}{2k}}}.$$

A recursive application of this inequality enables us to restrict our supremum to $k > \alpha\log_2(n)$. Moreover, for such k we have

$$1 \geq n^{-1/k} = \exp\Big(-\frac{\log(n)}{k}\Big) \geq \exp\Big(-\frac{\log(n)}{\alpha\log_2(n)}\Big) = 2^{-1/\alpha}.$$

Combining this with Lemma A.2 (ii) we get the assertion

$$\varepsilon_n(D_\sigma) \preccurlyeq \sup_{k\geq 1}\frac{\tau_k}{n^{1/k}} = \sup_{k>\alpha\log_2(n)}\frac{\tau_k}{n^{1/k}} \asymp \sup_{k>\alpha\log_2(n)}\tau_k = \tau_{\lfloor\alpha\log_2(n)\rfloor+1} \asymp \tau_{\lfloor\log_2(n)\rfloor+1}. \qquad \square$$

A. Conditions on Sequences

In this section we collect some characterizations of the conditions used on the diagonal sequence. Most of them are consequences of the general theory of $\mathcal{O}$-regular varying functions/sequences, but for convenience we include the proofs, respectively give detailed references. These results enable us to compare our findings with [11, 12]. In the following, all supremums $\sup_{k\le n}$ and infimums $\inf_{k\le n}$ are taken over all tuples $(n,k)\in\mathbb{N}^2$ with $k\le n$.

A.1 Lemma ((EXP) Sequences) *Let $r,s>0$, $\sigma=(\sigma_k)_{k\ge1}$ with $\sigma_k>0$ and $\sigma_k\searrow 0$, τ be the tail sequence given by (2), and $v_n := \left(\sum_{k=1}^n \sigma_k^{-s}\right)^{1/s}$ the partial sum sequence. Then the following statements are equivalent:*

(i) There is a real number $b>1$ with $\sup_{k\le n}\frac{\sigma_n b^n}{\sigma_k b^k}<\infty$.

(ii) There is an $n_0\ge1$ and a real number $0<a<1$ with $\sigma_{k+n_0}\le a\,\sigma_k$ for all $k\ge1$.

(iii) $\sigma_n \asymp 1/v_n$.

(iv) $\sigma_n\asymp\tau_n$.

Note that Condition (i) and (ii) are independent of $r>0$ and $s>0$. Consequently, if σ satisfies Condition (iii) or (iv) for some $s>0$ resp. $r>0$ then σ satisfies both conditions for all $r,s>0$.

Proof. (i)⇒(iii) For $c:=\sup_{k\le n}\frac{\sigma_n b^n}{\sigma_k b^k}<\infty$ we get

$$v_n^s\sigma_n^s=\sum_{k=1}^n\left(\frac{\sigma_n}{\sigma_k}\right)^s\le c^s\sum_{k=1}^n b^{-s(n-k)}=c^s\sum_{k=0}^{n-1}b^{-sk}\le\frac{(bc)^s}{b^s-1}$$

for all $n\ge1$. Moreover, $v_n\sigma_n\ge1$ always holds. By considering $(\tau_k/\sigma_k)^r$ we can analogously prove (i)⇒(iv).

(iii)⇒(ii) Let $c>0$ be a constant with $v_n\sigma_n\le c$ for all $n\ge1$. Because of the monotonicity of σ we get for $k,n_0\ge1$

$$c^s\ge v_{k+n_0}^s\sigma_{k+n_0}^s=\sum_{i=1}^{k+n_0}\left(\frac{\sigma_{k+n_0}}{\sigma_i}\right)^s\ge\sum_{i=k}^{k+n_0}\left(\frac{\sigma_{k+n_0}}{\sigma_i}\right)^s\ge\left(\frac{\sigma_{k+n_0}}{\sigma_k}\right)^s(n_0+1).$$

Choosing $n_0:=\lceil c^s\rceil$ yields

$$\frac{\sigma_{k+n_0}}{\sigma_k}\le\frac{c}{(n_0+1)^{1/s}}\le\frac{c}{(c^s+1)^{1/s}}<1$$

for all $k\ge1$.

(iv)⇒(ii) Let $c>0$ be a constant with $\tau_k\le c\sigma_k$ for all $k\ge1$. Because of the monotonicity of σ we get for $k,n_0\ge1$

$$c^r\ge\frac{\tau_k^r}{\sigma_k^r}=\sum_{n=k}^{\infty}\left(\frac{\sigma_n}{\sigma_k}\right)^r\ge\sum_{n=k}^{k+n_0}\left(\frac{\sigma_n}{\sigma_k}\right)^r\ge\left(\frac{\sigma_{k+n_0}}{\sigma_k}\right)^r(n_0+1).$$

Hence Statement (ii) follows along the same line as (iii)⇒(ii).

(ii)⇒(i) For $k\le n$ there is a unique $m\ge0$ with $k+mn_0\le n<k+(m+1)n_0$. Using the monotonicity of σ and Assumption (ii) m-times we get

$$\sigma_n\le\sigma_{k+mn_0}\le\sigma_k a^m\le\frac{\sigma_k}{a}a^{\frac{n-k}{n_0}}=\frac{\sigma_k}{a}b^{k-n}$$

with $b=a^{-1/n_0}>1$. Hence the supremum is bounded by a^{-1}. ☐

A.2 Lemma (Doubling Condition) *Let $\sigma = (\sigma_k)_{k\geq 1}$ with $\sigma_k > 0$ and $\sigma_k \searrow 0$. Then the following statements are equivalent:*

(i) $\sigma_n \asymp \sigma_{2n}$.

(ii) $\sigma_{\lfloor x\rfloor+1} \asymp \sigma_{\lfloor \lambda x\rfloor+1}$ *as function in $x > 0$ for all $\lambda > 0$.*

(iii) $\inf_{k\leq n} \frac{\sigma_n n^\alpha}{\sigma_k k^\alpha} > 0$ *for some $\alpha > 0$.*

(iv) $\sigma_n \asymp (\sigma_1 \cdot \ldots \cdot \sigma_n)^{1/n}$.

Note that the symbol $\asymp$ in Statement (ii) means that there are constants $c_1, c_2 > 0$ with $c_1 \leq \sigma_{\lfloor \lambda x\rfloor+1}/\sigma_{\lfloor x\rfloor+1} \leq c_2$ for all $x > 0$. Moreover, Statement (iii) implies $\sigma_n \succcurlyeq n^{-\alpha}$ and hence σ decreases at most polynomial.

Proof. (i)⇔(iii) This has already been pointed out by Kühn [11, p. 482] and is a direct consequence of the monotonicity. (i)⇔(ii) There are some closely related results in the literature, see e.g. [7, Theorem 1], but we did not exactly find this one. For this reason we present a proof. Obviously (ii) implies (i) and for the inverse implication we first show that the set

$$G_\sigma := \left\{\lambda > 0 :\ \exists a_\lambda, b_\lambda > 0 :\ \forall x > 0 :\ a_\lambda \leq \frac{\sigma_{\lfloor \lambda x\rfloor+1}}{\sigma_{\lfloor x\rfloor+1}} \leq b_\lambda\right\}$$

is a subgroup of the multiplicative group $(0,\infty)$. Clearly, $1 \in G_\sigma$ and if $\lambda, \mu \in G_\sigma$ then

$$\frac{\sigma_{\lfloor \lambda\mu x\rfloor+1}}{\sigma_{\lfloor x\rfloor+1}} = \frac{\sigma_{\lfloor \lambda\mu x\rfloor+1}}{\sigma_{\lfloor \mu x\rfloor+1}} \frac{\sigma_{\lfloor \mu x\rfloor+1}}{\sigma_{\lfloor x\rfloor+1}} \begin{cases} \leq b_\lambda b_\mu \\ \geq a_\lambda a_\mu \end{cases}$$

holds for all $x > 0$. Hence $\lambda\mu \in G_\sigma$. If $\lambda \in G_\sigma$ then

$$\frac{\sigma_{\lfloor x/\lambda\rfloor+1}}{\sigma_{\lfloor x\rfloor+1}} = \frac{\sigma_{\lfloor x/\lambda\rfloor+1}}{\sigma_{\lfloor \lambda(x/\lambda)\rfloor+1}} \begin{cases} \leq \frac{1}{a_{1/\lambda}} \\ \geq \frac{1}{b_{1/\lambda}} \end{cases}$$

holds for all $x > 0$. Hence $\lambda^{-1} \in G_\sigma$ and G_σ is indeed a subgroup of $(0,\infty)$. Now, because of the monotonicity of σ we have for $1 \leq \lambda \leq 2$

$$1 \geq \frac{\sigma_{\lfloor \lambda x\rfloor+1}}{\sigma_{\lfloor x\rfloor+1}} \geq \frac{\sigma_{2(\lfloor x\rfloor+1)}}{\sigma_{\lfloor x\rfloor+1}} \geq c$$

for all $x > 0$, where $c > 0$ is a constant satisfying $\sigma_{2n} \geq c\sigma_n$ for all $n \geq 1$. Hence $[1,2] \subseteq G_\sigma$ and this implies $G_\sigma = (0,\infty)$.

(iii)⇒(iv) Because of the monotonicity of σ we always have $(\sigma_1 \cdot \ldots \cdot \sigma_n)^{\frac{1}{n}} \geq \sigma_n$. For $c := \inf_{k\leq n} \frac{\sigma_n n^\alpha}{\sigma_k k^\alpha} > 0$ we then have $\sigma_k \leq c^{-1}\sigma_n n^\alpha k^{-\alpha}$ for all $k \leq n$. Since Stirling's formula yields $(n!)^{1/n} \asymp n$ we get

$$(\sigma_1 \cdot \ldots \cdot \sigma_n)^{1/n} \leq c^{-1}\sigma_n \left(\frac{n}{(n!)^{1/n}}\right)^\alpha \asymp \sigma_n.$$

(iv)⇒(i) Let $c > 0$ with $\sigma_n \le (\sigma_1 \cdot \ldots \cdot \sigma_n)^{1/n} \le c\sigma_n$ for all $n \ge 1$. Then we have

$$c\sigma_{2n} \ge (\sigma_1 \cdot \ldots \cdot \sigma_{2n})^{\frac{1}{2n}} = (\sigma_1 \cdot \ldots \cdot \sigma_n)^{\frac{1}{2n}}(\sigma_{n+1} \cdot \ldots \cdot \sigma_{2n})^{\frac{1}{2n}} \ge \sqrt{\sigma_n \sigma_{2n}}.$$

for all $n \ge 1$. Hence $c^2\sigma_{2n} \ge \sigma_n \ge \sigma_{2n}$ for all $n \ge 1$. □

A.3 Lemma (Tail Sequence) *Let $r > 0$, $\sigma = (\sigma_k)_{k\ge1}$ with $\sigma_k > 0$ and $\sigma_k \searrow 0$ and τ be the tail sequence given by* (2)*. Then the following statements hold:*

(i) The following statements are equivalent:

(a) $\sup_{k\le n} \frac{\sigma_n n^\alpha}{\sigma_k k^\alpha} < \infty$ *for some* $\alpha > 1/r$.

(b) Condition (ALP): $\tau_n \preccurlyeq \sigma_n n^{1/r}$.

(ii) The following statements are equivalent:

(c) $\tau_n \asymp \tau_{2n}$.

(d) Condition (AMP): $\tau_n \succcurlyeq \sigma_n n^{1/r}$.

(iii) Condition $\sigma_n \asymp \sigma_{2n}$ *implies* $\tau_n \asymp \tau_{2n}$*, and if we additionally assume (a) then we have equivalence.*

Proof. (a)⇒(b) For $c := \sup_{k\le n} \frac{\sigma_n n^\alpha}{\sigma_k k^\alpha} < \infty$ we get

$$\frac{\tau_k^r}{k\sigma_k^r} = \frac{1}{k}\sum_{n=k}^{\infty}\Big(\frac{\sigma_n}{\sigma_k}\Big)^r \le c^r k^{\alpha r-1}\sum_{n=k}^{\infty} n^{-\alpha r}$$

for all $k \ge 1$. Estimating the remaining sum using integrals we get the assertion

$$k^{\alpha r-1}\sum_{n=k}^{\infty} n^{-\alpha r} \le k^{\alpha r-1}\Big(k^{-\alpha r} + \int_k^\infty t^{-\alpha r}\ \mathrm{d}t\Big) = k^{\alpha r-1}\Big(k^{-\alpha r} + \frac{k^{1-\alpha r}}{\alpha r-1}\Big) \le \frac{\alpha r}{\alpha r-1}.$$

(b)⇒(a) is a consequence of Bingham et al. [1, Theorem 2.6.3] to the positive and measurable function $f(x) := x\sigma_{\lfloor x\rfloor}^r$ for $x \ge 1$. To this end, we recall the definition of *almost decreasing* functions from [1, Section 2.2.1] and the *Matuszewska index* $\alpha(f)$ of f, defined in [1, Section 2.1.2]. Moreover, we have

$$\alpha(f) = \inf\big\{\alpha \in \mathbb{R}:\ x^{-\alpha}f(x) \text{ is almost decreasing}\big\}.$$

according to [1, Theorem 2.2.2]. Since $x^{-1}f(x)$ is decreasing we have $\alpha(f) \le 1 < \infty$ and hence f is of *bounded increase*, i.e. $f \in \mathrm{BI}$, see [1, p. 71] for a definition. Consequently, [1, Theorem 2.6.3 (d)] is applicable to the function f. For the $\tilde f(x) := \int_x^\infty f(t)/t\,\mathrm{d}t$ we have

$$\frac{f(x)}{\tilde f(x)} = \frac{x\sigma_{\lfloor x\rfloor}^r}{\tau_{\lfloor x\rfloor}^r - (x-\lfloor x\rfloor)\sigma_{\lfloor x\rfloor}^r} \ge \frac{x\sigma_{\lfloor x\rfloor}^r}{\tau_{\lfloor x\rfloor}^r} \ge \frac{\lfloor x\rfloor\sigma_{\lfloor x\rfloor}^r}{\tau_{\lfloor x\rfloor}^r} \ge c^{-r}$$

for all $x \ge 1$, where $c > 0$ is a constant satisfying $\tau_n \le c\sigma_n n^{1/r}$ for all $n \ge 1$. Therefore, $\liminf_{x\to\infty} f(x)/\tilde f(x) > 0$ and [1, Theorem 2.6.3 (d)] yields $\alpha(f) < 0$. Accordingly, there is a $\alpha_0 < 0$ such that $x^{-\alpha_0}f(x)$ is almost decreasing. The definition of almost decreasing functions, see [1, Section 2.2.1], gives us the assertion with $\alpha = \frac{1-\alpha_0}{r} > 1/r$.

(c)⇒(d) This is from [12, first equation on p. 45]. (d)⇒(c) The following idea is from Bojanic and Seneta [2, proof of Theorem 4]. According to our assumption the sequence

$$\rho_n := n\Big(1 - \frac{\tau_{n+1}^r}{\tau_n^r}\Big) = n\frac{\tau_n^r - \tau_{n+1}^r}{\tau_n^r} = \frac{n\sigma_n^r}{\tau_n^r}$$

is positive and bounded. Building a telescope product we get

$$\frac{\tau_n^r}{\tau_1^r} = \prod_{k=1}^{n-1} \frac{\tau_{k+1}^r}{\tau_k^r} = \prod_{k=1}^{n-1}\Big(1 - \frac{\rho_k}{k}\Big).$$

Since $0 < 1 - \frac{\rho_k}{k} < 1$ this gives the representation $\tau_n^r = \exp\circ\log(\tau_n^r) = \exp(\gamma_n - \sum_{k=1}^{n-1}\rho_k/k)$ with

$$\gamma_n := \log \tau_1^r + \sum_{k=1}^{n-1}\Big[\log\Big(1 - \frac{\rho_k}{k}\Big) + \frac{\rho_k}{k}\Big].$$

Below we will prove that $(\gamma_n)_{n\geq 1}$ converges and hence the assertion is a consequence of this representation of τ_n^r according to [6, Theorem 2]. Now, to the convergence of $(\gamma_n)_{n\geq 1}$. Since $(\rho_k)_{k\geq 1}$ is bounded the sequence $a_k := \rho_k/k$ is square summable. Without loss of generality we assume that there is a $0 < q < 1$ with $a_n < q$ for all $n \geq 1$. Using the Taylor series of the logarithm we get

$$\log(1 - a_k) + a_k = -\sum_{\ell=1}^{\infty}\frac{a_k^\ell}{\ell} + a_k = -\sum_{\ell=2}^{\infty}\frac{a_k^\ell}{\ell}.$$

Additionally, for $\ell \geq 2$ we have the estimate $\sum_{k=1}^{\infty} a_k^\ell \leq \|a\|_{\ell_2}^2 q^{\ell-2}$. Together we get the absolute convergence of the series

$$\sum_{k=1}^{\infty}|\log(1 - a_k) + a_k| = \sum_{k=1}^{\infty}\sum_{\ell=2}^{\infty}\frac{a_k^\ell}{\ell} = \sum_{\ell=2}^{\infty}\frac{1}{\ell}\sum_{k=1}^{\infty}a_k^\ell \leq \frac{\|a\|_{\ell_2}^2}{q^2}\sum_{\ell=2}^{\infty}\frac{q^\ell}{\ell} < \infty.$$

(iii) According to our assumption there is a constant $c > 0$ with $\sigma_{2n} \geq c\sigma_n$ for all $n \geq 1$. Then the assertion follows by

$$\tau_{2n}^r \geq \sum_{k=n}^{\infty}\sigma_{2k}^r \geq c^r\sum_{k=n}^{\infty}\sigma_k^r = c^r\tau_n^r.$$

For the inverse we additionally assume (a) and hence we have also (b) and (d), i.e. $\tau_n \asymp \sigma_n n^{1/r}$. Consequently, σ satisfies the doubling condition $\sigma_{2n} \asymp \tau_{2n}(2n)^{-1/r} \asymp \tau_n n^{-1/r} \asymp \sigma_n$. □

References

[1] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, Cambridge University Press, Cambridge, 1989.

[2] R. Bojanic and E. Seneta. A unified theory of regularly varying sequences. *Math. Z.*, 134:91–106, 1973.

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

[4] B. Carl and P. Rudolph. Entropy numbers of operators factoring through general diagonal operators. *Rev. Mat. Complut.*, 27:623–639, 2014.

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

[6] D. Djurčić and A. Torgašev. Representation theorems for the sequences of the classes CR_c and ER_c. *Siberian Math. J.*, 45:855–859, 2004.

[7] D. Djurčić and V. Božin. A proof of an Aljančić hypothesis on $\mathcal{O}$-regularly varying sequences. *Publ. Inst. Math. (Beograd) (N.S.)*, 62(76):46–52, 1997.

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

[9] A. N. Kolmogorov and V. M. Tikhomirov. ε-entropy and ε-capacity of sets in functional spaces. *Uspekhi Mat. Nauk*, 17, 1961.

[10] T. Kühn. Entropy numbers of diagonal operators of logarithmic type. *Georgian Math. J.*, 8: 307–318, 2001.

[11] T. Kühn. Entropy numbers of general diagonal operators. *Rev. Mat. Complut.*, 18:479–491, 2005.

[12] T. Kühn. Entropy numbers in sequence spaces with an application to weighted function spaces. *J. Approx. Theory*, 153:40–52, 2008.

[13] M. B. Marcus. The ε-entropy of some compact subsets of ℓ^p. *J. Approx. Theory*, 10:304–312, 1974.

[14] B. S. Mitjagin. Approximate dimension and bases in nuclear spaces. *Uspehi Mat. Nauk*, 16:63–132, 1961.

[15] R. Oloff. Entropieeigenschaften von Diagonaloperatoren. *Math. Nachr.*, 86:157–165, 1978.

[16] X. Wang. Volumes of generalized unit balls. *Math. Mag.*, 78:390–395, 2005.

Simon Fischer
Fachbereich Mathematik, Universität Stuttgart, Stuttgart, Germany
E-Mail: simon.fischer@mathematik.uni-stuttgart.de

Erschienene Preprints ab Nummer 2013-001

Komplette Liste:
www.f08.uni-stuttgart.de/mathematik/forschung/publikationen/mathematische_berichte

2019-004 *Fischer, S.:* Some New Bounds on the Entropy Numbers of Diagonal Operators

2019-003 *Braun, A.; Kohler, M.; Walk, H.:* On the rate of convergence of a neural network regression estimate learned by gradient descent

2019-002 *Györfi, L.; Henze, N.; Walk, H.:* The limit distribution of the maximum probability nearest neighbor ball

2019-001 *Gaspoz, F.D.; Heine, C.-J.; Siebert, K.G.:* An Alternative Proof of H^1-Stability of the L_2-Projection on Graded Meshes

2018-003 *Kollross, A.:* Octonions, triality, the exceptional Lie algebra F4, and polar actions on the Cayley hyperbolic plane

2018-002 *Diaź-Ramos, J.C.; Domínguez-Vázquez, M.; Kollross, A.:* On homogeneous manifolds whose isotropy actions are polar

2018-001 *Grundhöfer, T.; Stroppel, M.; Van Maldeghem, H.:* Embeddings of hermitian unitals into pappian projective planes

2017-011 *Hansmann, M.; Kohler, M.; Walk, H.:* On the strong universal consistency of local averaging regression estimates

2017-010 *Devroye, L.; Györfi, L.; Lugosi, G.; Walk, H.:* A nearest neighbor estimate of a regression functional

2017-009 *Steinke, G.; Stroppel, M.:* On elation Laguerre planes with a two-transitive orbit on the set of generators

2017-008 *Steinke, G.; Stroppel, M.:* Laguerre planes and shift planes

2017-007 *Blunck, A.; Knarr, N.; Stroppel, B.; Stroppel, M.:* Transitive groups of similitudes generated by octonions

2017-006 *Blunck, A.; Knarr, N.; Stroppel, B.; Stroppel, M.:* Clifford parallelisms defined by octonions

2017-005 *Knarr, N.; Stroppel, M.:* Subforms of Norm Forms of Octonion Fields

2017-004 *Apprich, C.; Dieterich, A.; Höllig, K.; Nava-Yazdani, E.:* Cubic Spline Approximation of a Circle with Maximal Smoothness and Accuracy

2017-003 *Fischer, S.; Steinwart, I.:* Sobolev Norm Learning Rates for Regularized Least-Squares Algorithm

2017-002 *Farooq, M.; Steinwart, I.:* Learning Rates for Kernel-Based Expectile Regression

2017-001 *Bauer, B.; Devroye, L; Kohler, M.; Krzyzak, A.; Walk, H.:* Nonparametric Estimation of a Function From Noiseless Observations at Random Points

2016-006 *Devroye, L.; Györfi, L.; Lugosi, G.; Walk, H.:* On the measure of Voronoi cells

2016-005 *Kohls, C.; Kreuzer, C.; Rösch, A.; Siebert, K.G.:* Convergence of Adaptive Finite Elements for Optimal Control Problems with Control Constraints

2016-004 *Blaschzyk, I.; Steinwart, I.:* Improved Classification Rates under Refined Margin Conditions

2016-003 *Feistauer, M.; Roskovec, F.; Sändig, AM.:* Discontinuous Galerkin Method for an Elliptic Problem with Nonlinear Newton Boundary Conditions in a Polygon

2016-002 *Steinwart, I.:* A Short Note on the Comparison of Interpolation Widths, Entropy Numbers, and Kolmogorov Widths

2016-001 *Köster, I.:* Sylow Numbers in Spectral Tables

2015-016 *Hang, H.; Steinwart, I.:* A Bernstein-type Inequality for Some Mixing Processes and Dynamical Systems with an Application to Learning

2015-015 *Steinwart, I.:* Representation of Quasi-Monotone Functionals by Families of Separating Hyperplanes

2015-014 *Muhammad, F.; Steinwart, I.:* An SVM-like Approach for Expectile Regression

2015-013 *Nava-Yazdani, E.:* Splines and geometric mean for data in geodesic spaces

2015-012 *Kimmerle, W.; Köster, I.:* Sylow Numbers from Character Tables and Group Rings

2015-011 *Györfi, L.; Walk, H.:* On the asymptotic normality of an estimate of a regression functional

2015-010 *Gorodski, C, Kollross, A.:* Some remarks on polar actions

2015-009 *Apprich, C.; Höllig, K.; Hörner, J.; Reif, U.:* Collocation with WEB–Splines

2015-008 *Kabil, B.; Rodrigues, M.:* Spectral Validation of the Whitham Equations for Periodic Waves of Lattice Dynamical Systems

2015-007 *Kollross, A.:* Hyperpolar actions on reducible symmetric spaces

2015-006 *Schmid, J.; Griesemer, M.:* Well-posedness of Non-autonomous Linear Evolution Equations in Uniformly Convex Spaces

2015-005 *Hinrichs, A.; Markhasin, L.; Oettershagen, J.; Ullrich, T.:* Optimal quasi-Monte Carlo rules on higher order digital nets for the numerical integration of multivariate periodic functions

2015-004 *Kutter, M.; Rohde, C.; Sändig, A.-M.:* Well-Posedness of a Two Scale Model for Liquid Phase Epitaxy with Elasticity

2015-003 *Rossi, E.; Schleper, V.:* Convergence of a numerical scheme for a mixed hyperbolic-parabolic system in two space dimensions

2015-002 *Döring, M.; Györfi, L.; Walk, H.:* Exact rate of convergence of kernel-based classification rule

2015-001 *Kohler, M.; Müller, F.; Walk, H.:* Estimation of a regression function corresponding to latent variables

2014-021 *Neusser, J.; Rohde, C.; Schleper, V.:* Relaxed Navier-Stokes-Korteweg Equations for Compressible Two-Phase Flow with Phase Transition

2014-020 *Kabil, B.; Rohde, C.:* Persistence of undercompressive phase boundaries for isothermal Euler equations including configurational forces and surface tension

2014-019 *Bilyk, D.; Markhasin, L.:* BMO and exponential Orlicz space estimates of the discrepancy function in arbitrary dimension

2014-018 *Schmid, J.:* Well-posedness of non-autonomous linear evolution equations for generators whose commutators are scalar

2014-017 *Margolis, L.:* A Sylow theorem for the integral group ring of $PSL(2, q)$

2014-016 *Rybak, I.; Magiera, J.; Helmig, R.; Rohde, C.:* Multirate time integration for coupled saturated/unsaturated porous medium and free flow systems

2014-015 *Gaspoz, F.D.; Heine, C.-J.; Siebert, K.G.:* Optimal Grading of the Newest Vertex Bisection and H^1-Stability of the L_2-Projection

2014-014 *Kohler, M.; Krzyżak, A.; Walk, H.:* Nonparametric recursive quantile estimation

2014-013 *Kohler, M.; Krzyżak, A.; Tent, R.; Walk, H.:* Nonparametric quantile estimation using importance sampling

2014-012 *Györfi, L.; Ottucsák, G.; Walk, H.:* The growth optimal investment strategy is secure, too.

2014-011 *Györfi, L.; Walk, H.:* Strongly consistent detection for nonparametric hypotheses

2014-010 *Köster, I.:* Finite Groups with Sylow numbers $\{q^x, a, b\}$

2014-009 *Kahnert, D.:* Hausdorff Dimension of Rings

2014-008 *Steinwart, I.:* Measuring the Capacity of Sets of Functions in the Analysis of ERM

2014-007 *Steinwart, I.:* Convergence Types and Rates in Generic Karhunen-Loève Expansions with Applications to Sample Path Properties

2014-006 *Steinwart, I.; Pasin, C.; Williamson, R.; Zhang, S.:* Elicitation and Identification of Properties

2014-005 *Schmid, J.; Griesemer, M.:* Integration of Non-Autonomous Linear Evolution Equations

2014-004 *Markhasin, L.:* L_2- and $S^r_{p,q}B$-discrepancy of (order 2) digital nets

2014-003 *Markhasin, L.:* Discrepancy and integration in function spaces with dominating mixed smoothness

2014-002 *Eberts, M.; Steinwart, I.:* Optimal Learning Rates for Localized SVMs

2014-001 *Giesselmann, J.:* A relative entropy approach to convergence of a low order approximation to a nonlinear elasticity model with viscosity and capillarity

2013-016 *Steinwart, I.:* Fully Adaptive Density-Based Clustering

2013-015 *Steinwart, I.:* Some Remarks on the Statistical Analysis of SVMs and Related Methods

2013-014 *Rohde, C.; Zeiler, C.:* A Relaxation Riemann Solver for Compressible Two-Phase Flow with Phase Transition and Surface Tension

2013-013 *Moroianu, A.; Semmelmann, U.:* Generalized Killling spinors on Einstein manifolds

2013-012 *Moroianu, A.; Semmelmann, U.:* Generalized Killing Spinors on Spheres

2013-011 *Kohls, K; Rösch, A.; Siebert, K.G.:* Convergence of Adaptive Finite Elements for Control Constrained Optimal Control Problems

2013-010 *Corli, A.; Rohde, C.; Schleper, V.:* Parabolic Approximations of Diffusive-Dispersive Equations

2013-009 *Nava-Yazdani, E.; Polthier, K.:* De Casteljau's Algorithm on Manifolds

2013-008 *Bächle, A.; Margolis, L.:* Rational conjugacy of torsion units in integral group rings of non-solvable groups

2013-007 *Knarr, N.; Stroppel, M.J.:* Heisenberg groups over composition algebras

2013-006 *Knarr, N.; Stroppel, M.J.:* Heisenberg groups, semifields, and translation planes

2013-005 *Eck, C.; Kutter, M.; Sändig, A.-M.; Rohde, C.:* A Two Scale Model for Liquid Phase Epitaxy with Elasticity: An Iterative Procedure

2013-004 *Griesemer, M.; Wellig, D.:* The Strong-Coupling Polaron in Electromagnetic Fields

2013-003 *Kabil, B.; Rohde, C.:* The Influence of Surface Tension and Configurational Forces on the Stability of Liquid-Vapor Interfaces

2013-002 *Devroye, L.; Ferrario, P.G.; Györfi, L.; Walk, H.:* Strong universal consistent estimate of the minimum mean squared error

2013-001 *Kohls, K.; Rösch, A.; Siebert, K.G.:* A Posteriori Error Analysis of Optimal Control Problems with Control Constraints