A Short Note on the Comparison of Interpolation Widths, Entropy Numbers, and Kolmogorov Widths

Ingo Steinwart

March 27, 2022

Abstract

We compare the Kolmogorov and entropy numbers of compact operators mapping from a Hilbert space into a Banach space. We then apply these general findings to embeddings between reproducing kernel Hilbert spaces and $L_\infty(\mu)$. Here we provide a sufficient condition for a gap of the order $n^{1/2}$ between the associated interpolation and Kolmogorov n-widths. Finally, we show that in the multi-dimensional Sobolev case, this gap actually occurs between the Kolmogorov and approximation widths.

1 Introduction

Let $(X, \mathcal{A}, \mu)$ be a measure space and H be a a reproducing kernel Hilbert space (RKHS) over X. Moreover, assume that the kernel k of H is measurable and that for all $p \in [2, \infty]$, the map $I_{k,\mu} : H \to L_p(\mu)$ defined by $I_{k,\mu} f := [f]_\sim$, where $[f]_\sim$ denotes the μ-equivalence class of f in $L_p(\mu)$, is compact. Now consider the linear interpolation n-width of H in $L_2(\mu)$, that is

$$I_n\bigl(H, L_p(\mu)\bigr) := \inf_{D\subset X, |D|\le n} \left(\int_X \sup_{f\in B_H} \bigl|f(x) - A_D f(x)\bigr|^p \, d\mu(x) \right)^{1/p},$$

with the usual modification for $p = \infty$. Here, $A_D : H \to H$ is the bounded linear operator defined by $A_D f(x) := \sum_{i=1}^n \alpha_i^*(x) f(x_i)$, where $D = (x_1, \dots, x_n)$ and $\alpha^*(x) \in \mathbb{R}^n$ is the unique solution of

$$\alpha^*(x) = \arg\min_{\alpha\in\mathbb{R}^n} \Bigl\| \delta_x - \sum_{i=1}^n \alpha_i \delta_{x_i} \Bigr\|_{H'}^2 .$$

For later use we note that we always have

$$\inf_{D\subset X, |D|\le n} \sup_{f\in B_H} \|f - A_D f\|_{L_p(\mu)} \le I_n\bigl(H, L_p(\mu)\bigr) \tag{1}$$

and equality holds in the extreme case $p = \infty$. Moreover, consider the classical Kolmogorov n-width

$$d_n\bigl(H, L_p(\mu)\bigr) = \inf_{F_n\subset L_p(\mu)} \sup_{f\in B_H} \inf_{g\in F_n} \|f - g\|_{L_p(\mu)} ,$$

where the left most infimum runs over all subspaces F_n of $L_p(\mu)$ with $\dim F_n \le n$. Note that the lower bound of I_n in (1) measures, how well f can be approximated by a very particular linear and n-dimensional scheme, whereas d_n measures how well f can be approximated by the best n-dimensional scheme. Consequently, the approximation n-width

$$a_n\bigl(H, L_p(\mu)\bigr) := \inf_{A: H\to L_p(\mu)} \sup_{f\in B_H} \|f - Af\|_{L_p(\mu)} ,$$

where the infimum is taken over all bounded linear operators $A : H \to L_p(\mu)$ with $\operatorname{rank} A \leq n$, satisfies $d_n(H, L_p(\mu)) \leq a_n(H, L_p(\mu)) \leq I_n(H, L_p(\mu))$.

In the Hilbert space case, that is, $p = 2$, these quantities are well understood. Indeed, the general theory of s-numbers [12] shows, see e.g. Section 2, that

$$d_n\big(H, L_2(\mu)\big) = a_n\big(H, L_2(\mu)\big) = \sqrt{\lambda_{n+1}}\,, \tag{2}$$

where (λ_n) denotes the (extended and) ordered sequence of eigenvalues of the integral operator $T_k : L_2(\mu) \to L_2(\mu)$ associated with the kernel k. Moreover, if H is a Sobolev space, then $I_n(H, L_p(\mu))$ shares the asymptotic behavior of (2) and this can actually be achieved by taking quasi-uniform points $D \subset X$, see [16]. Unfortunately, the situation changes in the other extreme, namely $p = \infty$. Indeed, if μ is a finite measure, then (2) immediately yields

$$\sqrt{\lambda_{n+1}} \leq \sqrt{\mu(X)}\, d_n\big(H, L_\infty(\mu)\big)\,,$$

while [14, Theorem 3] shows that

$$\sqrt{\sum_{i=n+1}^{\infty} \lambda_i} \leq \sqrt{\mu(X)}\, I_n\big(H, L_\infty(\mu)\big)\,, \tag{3}$$

and in the Sobolev case, this lower bound is matched by an upper bound of the same asymptotic behavior, see [15]. In the case of an algebraic decay of the eigenvalues, it is not hard to see that there is a gap of the order $n^{-1/2}$ between the *lower bounds* for $d_n(H, L_\infty(\mu))$ and $I_n(H, L_\infty(\mu))$, and this naturally raises the question, whether this gap actually occurs between the quantities of interest, that is, between $d_n(H, L_\infty(\mu))$ and $I_n(H, L_\infty(\mu))$. So far, a positive answer only exists for the 1-dimensional Sobolev case, see [13]. The goal of this note is to provide a positive answer in a more general framework. To be more precise, we show that for algebraically decaying eigenvalues we have $d_n(H, L_\infty(\mu)) \asymp \sqrt{\lambda_{n+1}}$ if and only if the entropy numbers of the embedding $I_{k,\mu} : H \to L_\infty(\mu)$ behave like $\sqrt{\lambda_{n+1}}$. Using (3) this characterization gives a sufficient condition for the existence of the gap. In addition, we present a result that highlights the role of the eigenfunctions of T_k. For the multi-dimensional Sobolev case we then show with the help of well-known asymptotics of the entropy and approximation numbers that the gap $n^{-1/2}$ actually occurs between $d_n(H, L_\infty(\mu))$ and $a_n(H, L_\infty(\mu))$, that is, between arbitrary n-dimensional approximation and linear n-dimensional approximation. In addition, the cases $p \in (2, \infty)$ are treated simultaneously.

The rest of this note is organized as follows: In Section 2 we recall the definition of entropy numbers and also introduce some examples of s-number scales. Section 3 summarizes the relationship between entropy numbers and the different s-number scales. In Section 4 two general results comparing entropy and Kolmogorov numbers of compact operators are presented and based upon these results the RKHS situation is investigated in more detail. In Section 5 we then apply these findings to the multi-dimensional Sobolev case.

2 Preliminaries: Entropy Numbers, s-Numbers, and Eigenvalues

We write $a_n \prec b_n$ for two positive sequences (a_n) and (b_n) if there exists a constant $c \in (0, \infty)$ such that $a_n \leq c b_n$ for all $n \geq 1$. Similarly, we write $a_n \asymp b_n$ if both $a_n \prec b_n$ and $b_n \prec a_n$. Finally, a positive sequence is called regular if there exists a constant $c \in (0, \infty)$ such that $a_n \leq c a_{2n}$ and $a_m \leq c a_n$ for all $1 \leq m \leq n$. Probably the most interesting examples of regular sequences are $a_n = n^{-p}(1 + \ln n)^{-q}$ for $p > 0$ and $q \in \mathbb{R}$, or $p = 0$ and $q > 0$.

Given a Banach space E, we denote its closed unit ball by B_E and its dual by E'. Moreover, we write $I_F : F \to \ell_\infty(B_{F'})$ for the canonical embedding and $Q_E : \ell_1(B_E) \to E$ for the canonical surjection. Furthermore, we write $E \hookrightarrow F$ if $E \subset F$ and the inclusion map is continuous. Finally, the adjoint of a bounded linear operator S acting between two Hilbert spaces is denoted by S^*.

Now, let E and F be Banach spaces and $T : E \to F$ be a bounded, linear operator. Then the n-th *(dyadic) entropy number* of T is defined by

$$e_n(T) := \inf\Big\{\varepsilon > 0 : \exists x_1, \dots, x_{2^{n-1}} \in F : TB_E \subset \bigcup_{i=1}^{2^{n-1}} x_i + \varepsilon B_F\Big\}.$$

Some elementary properties of entropy numbers can be found in [6, Chapter 1]. In particular, we have $e_n(T) \to 0$ if and only if T is compact. Since T is compact if and only if its dual T' is compact, this immediately raises the question how the entropy numbers of T and T' are related to each other. This question, known as the duality problem for entropy numbers has, so far, no complete answer. Partial answers, however, do exist. The one we will need is the following inequality taken from [3]

$$\frac{1}{d_p} \sup_{k\le n} k^{1/p}\, e_k(T) \le \sup_{k\le n} k^{1/p}\, e_k(T') \le d_p \sup_{k\le n} k^{1/p}\, e_k(T)\,, \tag{4}$$

which holds for all $n \ge 1$ and all compact $T : E \to F$, whenever E or F are B-convex. Here, $d_p \in (0,\infty)$ is a constant, which depends on $p \in (0,\infty)$ and the geometry of the involved spaces E and F, but which is independent of both n and T. Moreover, recall from e.g. [7, Theorem 13.10] that a Banach space is B-convex if and only if it has non-trivial type. In particular, Hilbert spaces are B-convex, and so are the spaces $L_p(\mu)$ for $p \in (1,\infty)$ since these spaces have type $\min\{2, p\}$, see e.g. [7, Chapter 11]. Moreover, if E or F is a Hilbert space, it was shown in [20] that we may choose $d_p = 32$ for all $p \in (0,\infty)$. Finally note that from the inequalities in (4) we can derive the following equivalences, which hold for all regular sequences (α_n) and all compact operators T:

$$\begin{aligned} e_n(T) \prec \alpha_n \qquad &\Longleftrightarrow \qquad e_n(T') \prec \alpha_n \\ e_n(T) \asymp \alpha_n \qquad &\Longleftrightarrow \qquad e_n(T') \asymp \alpha_n\,. \end{aligned}$$

For a proof, which is based on a little trick originating from Carl [5], we refer to the proof of [18, Corollary 1.19] or, in a slightly simplified version, to the proof of [17, Proposition 2].

Besides entropy numbers, we are also interested in some so-called s-numbers. Namely, if $T : E \to F$ is a bounded linear operator, we are interested in the n-th *approximation number* of T, defined by

$$a_n(T) := \inf\{\, \|T - A\| \mid A : E \to F \text{ bounded, linear with } \operatorname{rank} A < n \,\}\,,$$

in the n-th *Gelfand number* of T defined by

$$c_n(T) := \inf\{\, \|T I^E_{E_0}\| \;:\; E_0 \text{ subspace of } E \text{ with } \operatorname{codim} E_0 < n \,\}\,,$$

where $I^E_{E_0}$ denotes the canonical inclusion of E_0 into E, and in the n-th *Kolmogorov number* of T defined by

$$d_n(T) := \inf\{\, \|Q^F_{F_0} T\| \;:\; F_0 \text{ subspace of } F \text{ with } \dim F_0 < n \,\}\,,$$

where $Q^F_{F_0}$ denotes the canonical surjection from the Banach space F onto the quotient space F/F_0. Recall from [6, Proposition 2.2.2] that the latter quantity can also be expressed by

$$d_n(T) = \inf\Big\{\varepsilon > 0 : \exists F_\varepsilon \text{ subspace of } F \text{ with } \dim F_\varepsilon < n \text{ and } TB_E \subset F_\varepsilon + \varepsilon B_F\Big\}\,,$$

and consequently, we have

$$d_{n+1}(T) = \inf_{F_n \subset F} \sup_{y\in TB_E} \inf_{z\in F_n} \|y - z\|_F\,,$$

where the left most infimum runs over all subspaces F_n of F with $\dim F_n \leq n$. In other words, $d_{n+1}(T)$ equals the classical Kolmogorov n-width of the set TB_E in F, cf. [9, Chapter 13], and therefore we have $d_{n+1}(I_{k,\mu} : H \to L_p(\mu)) = d_n(H, L_p(\mu))$, where H and μ are as in the introduction. In addition, it is not hard to see that we also have $a_{n+1}(I_{k,\mu} : H \to L_p(\mu)) = a_n(H, L_p(\mu))$, and consequently we will consider the operator versions in the remaining parts of this note. Furthermore, recall e.g. from [6, Theorems 2.3.1 and 2.2.1, and Proposition 2.5.5] that we always have

$$\begin{aligned} c_n(T) &= a_n(I_F T) \\ d_n(T) &= a_n(T Q_E) \\ d_n(T') &= c_n(T)\,, \end{aligned}$$

and for compact T its dual operator T' additionally satisfies $c_n(T') = d_n(T)$, see [6, Proposition 2.5.6]. Moreover, the approximation, Gelfand, and Kolmogorov numbers are s-numbers in the sense of [12, Definition 2.2.1], and the same is true for the *Tikhomirov numbers* of T, which are defined by

$$t_n(T) := a_n(I_F T Q_E)\,, \qquad\qquad n \geq 1.$$

In addition, we always have

$$\begin{aligned} t_n(T) \leq c_n(T) \leq a_n(T) \leq \sqrt{2n}\, c_n(T) \\ t_n(T) \leq d_n(T) \leq a_n(T) \leq \sqrt{2n}\, d_n(T)\,, \end{aligned}$$

where we note that in both cases the first two inequalities follow from s-number properties and the right most inequalities can be found in [6, Propositions 2.4.3 and 2.4.6]. In addition, the factor $\sqrt{2n}$ can be sharpened to $1 + \sqrt{n-1}$.

The two chains of inequalities above show that the gap between the asymptotic behavior of (a_n), (c_n) and (d_n) is at most of the order $\sqrt{n}$. It is well-known that this gap is sometimes attained, see e.g. Section 5, while in other cases the gap vanishes. For example, we have

$$a_n(T) = c_n(T)\,, \tag{5}$$

if E is a Hilbert space, see [6, Proposition 2.4.1], or F has the metric extension property, see [6, Proposition 2.3.3], and

$$a_n(T) = d_n(T)\,, \tag{6}$$

if F is a Hilbert space, see [6, Proposition 2.4.4], or E has the metric lifting property, see [6, Proposition 2.2.3]. In this respect recall that the spaces $\ell_\infty(J)$ and $L_\infty(\mu)$, where μ is some finite measure, have the metric extension property, see [6, p. 60] and [7, Theorem 4.14], respectively. Moreover, the spaces $\ell_1(I)$ have the metric lifting property, see [6, p. 51]. By combining all these relations we further see that we have $t_n(T) = a_n(T)$ if either $E = \ell_1(I)$ and $F = \ell_1(J)$, or E and F are Hilbert spaces. The latter case also follows from a general result showing that there is only one s-scale for operators between Hilbert spaces, see e.g. [12, Theorem 2.11.9].

Our next goal is to relate the s-numbers introduced above to eigenvalues. To this end, let $S : H_1 \to H_2$ be a compact operator acting between two Hilbert spaces. Then $S^*S : H_1 \to H_1$ is compact, self-adjoint and positive, and therefore the classical spectral theorem shows that there is an at most countable family $(\lambda_i(S^*S))_{i\in I}$ of eigenvalues of S^*S, which in addition are non-negative and have at most one limit point, namely 0. In the following, we always assume that either $I = \{1, \ldots, n\}$ or $I = \mathbb{N}$, and that the eigenvalues are ordered decreasingly without excluding (geometric) multiplicities. Then, the *singular numbers* of S are defined by

$$s_i(S) := \begin{cases} \sqrt{\lambda_i(S^*S)} = \lambda_i(\sqrt{S^*S}) & \text{if } i \in I \\ 0 & \text{if } i \in \mathbb{N} \setminus I\,. \end{cases}$$

Recall that this gives $s_i(S) = s_i(S^*)$ for all $i \geq 1$, and $s_i(T) = \lambda_i(T)$ for all $i \in I$ if $T : H \to H$ is compact, self-adjoint and positive. Moreover, we have, see e.g. [12, Chapter 2.11]

$$s_n(S) = a_n(S)$$

for all $n \geq 1$ and all compact operators $S : H_1 \to H_2$ between Hilbert spaces H_1, H_2.

3 Carl's Inequality and some Inverse Versions

In this section we recall some inequalities between s-numbers and entropy numbers. We begin with Carl's inequality, which states that for all $p \in (0, \infty)$, there exists a constant $C_p \in (0, \infty)$ such that for all bounded, linear $T : E \to F$ and all $n \geq 1$, we have

$$\sup_{k \leq n} k^{1/p} e_k(T) \leq C_p \sup_{k \leq n} k^{1/p} a_k(T) \,. \tag{7}$$

We refer to [6, Theorem 3.1.1], where it is also shown that a possible value for the constant is $C_P = 128(32 + 16/p)^{1/p}$. Recall from e.g. [6, Chapter 1.3] that entropy numbers are surjective and weakly injective, and therefore we have

$$e_k(T) \leq 2e_k(I_F T Q_E) \leq 2e_k(T)$$

for all bounded, linear $T : E \to F$ and all $k \geq 1$. In particular, we may replace the approximation numbers in (7) by the Gelfand, Kolmogorov, or Tikhomirov numbers for the price of an additional factor of 2 in the constant. Moreover, like for the entropy numbers of T and T', we further have

$$a_n(T) \prec \alpha_n \qquad \Longrightarrow \qquad e_n(T) \prec \alpha_n$$

for all regular sequences (α_n) and all bounded linear $T : E \to F$. It is needless to say that the approximation numbers in this implication may be replaced by the Gelfand, Kolmogorov, or Tikhomirov numbers.

Let us now recall some inequalities that describe how certain s-numbers are dominated by entropy numbers. We begin with compact operators $S : H_1 \to H_2$ acting between two Hilbert spaces. Then [6, Inequality (3.0.9)] shows

$$a_n(S) \leq 2e_n(S) \tag{8}$$

for all $n \geq 1$. By an adaptation of the proof of [18, Corollary 1.19] we can then see that (8) in combination with (7) leads to the following equivalences, which hold for all regular sequences (α_n) and all compact operators $S : H_1 \to H_2$ acting between two Hilbert spaces:

$$a_n(S) \prec \alpha_n \qquad \Longleftrightarrow \qquad e_n(S) \prec \alpha_n \tag{9}$$

$$a_n(S) \asymp \alpha_n \qquad \Longleftrightarrow \qquad e_n(S) \asymp \alpha_n \,. \tag{10}$$

Again, the approximation numbers in these equivalences may be replaced by the Gelfand, Kolmogorov, or Tikhomirov numbers. Finally, let us consider the compact, self-adjoint and positive operator $T : H_1 \to H_1$ defined by S^*S. Then we have

$$s_i(T) = \lambda_i(T) = \lambda_i(S^*S) = s_i^2(S^*) \tag{11}$$

if $i \in I$ and $s_i(T) = 0 = s_i^2(S^*)$ if $i \in \mathbb{N} \setminus I$. The two equivalences above then lead to

$$s_n(T) \prec \alpha_n \qquad \Longleftrightarrow \qquad e_n(S^*) \prec \sqrt{\alpha_n} \tag{12}$$

$$s_n(T) \asymp \alpha_n \qquad \Longleftrightarrow \qquad e_n(S^*) \asymp \sqrt{\alpha_n} \tag{13}$$

for all regular sequences (α_n). Note that s_n can be replaced by any s-number scale, and in particular by the approximation, Gelfand, Kolmogorov, and Tikhomirov numbers. Moreover, we may replace $e_n(S^*)$ by $e_n(S)$ using the duality results for entropy numbers mentioned above.

Let us now consider the situation in which only one of the involved spaces is a Hilbert space, that is, we consider compact operators of the form $S : E \to H$ or $S : H \to F$, where H is a Hilbert space and E or F is an arbitrary Banach space. Then (7) still holds, but in general, we may no longer have (8). To compare the s-numbers of T to the entropy numbers of T, we thus need a surrogate for (8). Fortunately, there are a few such results. For example, [11, Lemme 1] shows that there exist constants $A, B \in (0, \infty)$ such that for all compact $S : E \to H$ and all $n \geq 1$ we have

$$n^{1/2} c_n(S) \leq B \sum_{k > An} k^{-1/2} e_k(S) \,. \tag{14}$$

With the help of this inequality it is easy to show that for all $p \in (0, 2)$ there exists another constant $B_p \in (0, \infty)$ such that

$$n^{1/p} c_n(S) \leq B_p \sup_{k > An} k^{1/p} e_k(S) \tag{15}$$

for all compact $S : E \to H$ and all $n \geq 1$. We refer to the very short proof of [11, Théorème A]. Complementary, [18, Theorem 5.12] shows that for all $p \in (2, \infty)$ there exists a constant $K_p \in (0, \infty)$ such that

$$\sup_{k \leq n} k^{1/p} t_k(S) \leq K_p \sup_{k \leq n} k^{1/p} e_k(S) \,. \tag{16}$$

for all compact operators $S : E \to H$ or $S : H \to F$ and all $n \geq 1$. Last but not least we like to mention that [5, Theorem 6] showed an inequality of the form (16) with t_k replaced by d_k or c_k for all $p \in (0, \infty)$ and all compact $S : E \to F$ for which E and F' are type 2 spaces.

4 Main Results

The goal of this section is to compare the entropy and Kolmogorov numbers of the embedding $I_{k,\mu} : H \to L_\infty(\mu)$. To this end, our first auxiliary result combines Carl's inequality with its inversed versions mentioned in Section 3.

Lemma 4.1. *Let H be a Hilbert space, F be a Banach space $S : H \to F$ be a compact operator, and $p \in (0, 2)$. Then, the following equivalence holds:*

$$d_n(S) \prec n^{-1/p} \qquad \Longleftrightarrow \qquad e_n(S) \prec n^{-1/p} \,. \tag{17}$$

Moreover, if F has the metric extension property, the equivalence is also true for $p \in (2, \infty)$, and in addition, we have

$$d_n(S) \asymp n^{-1/p} \qquad \Longleftrightarrow \qquad e_n(S) \asymp n^{-1/p} \,. \tag{18}$$

Finally, if F' has type 2, then (17) *and* (18) *hold for all $p \in (0, \infty)$.*

Proof of Lemma 4.1: Independent of p and F, the implication "$\Rightarrow$" in (17) is a direct consequence of Carl's inequality (7). For the proof of the converse implication we first consider the case $p \in (0, 2)$. By (4) we then know that $e_n(S') \prec n^{-1/p}$, and consequently (15) shows that $c_n(S') \prec n^{-1/p}$. Using $c_n(S') = d_n(S)$, which holds for compact operators S, we then obtain the assertion. In the case $p \in (2, \infty)$, we conclude by (16) that $t_n(S) \prec n^{-1/p}$. Moreover, F has the metric extension property, and therefore we have $c_n(SQ_E) = a_n(SQ_E)$ by (5). This leads to

$$t_n(S) = a_n(I_F S Q_E) = c_n(SQ_E) = a_n(SQ_E) = d_n(S) \,,$$

and hence we find $d_n(S) \prec n^{-1/p}$. In addition, (18) follows from combining (7) and (16) as in the proof of [17, Proposition 2]. Finally, the last assertion can be shown analogously using [5, Theorem 6] instead of (16). $\square$

Note that the equivalences obtained in Lemma 4.1 also holds for regular sequences of the form $\alpha_n = n^{-1/p}(\log n)^\beta$, where p satisfies the constraints of Lemma 4.1 and $\beta \in \mathbb{R}$. Indeed, for the second and third case this can be deduced from (7) and (16), respectively [5, Theorem 6], while in the first case this follows from (7) , (15), and [12, G.3.2].

Clearly, Lemma 4.1 in particular holds for compact operators $S : H \to L_p(\mu)$. Our next result shows that for some spaces $L_p(\mu)$ even more information can be obtained.

Theorem 4.2. *Let H be a Hilbert space, μ be a finite measure, and $p \in [2, \infty]$. Assume that we have a compact operator $S : H \to L_p(\mu)$ such that*

$$e_n\big(S : H \to L_2(\mu)\big) \asymp n^{-1/\alpha} \tag{19}$$

for some $\alpha \in (0, 2)$. Then, for all $q \in [2, p]$, the following equivalence hold:

$$d_n\big(S : H \to L_q(\mu)\big) \asymp n^{-1/\alpha} \qquad \Longleftrightarrow \qquad e_n\big(S : H \to L_q(\mu)\big) \asymp n^{-1/\alpha}\,.$$

Proof of Theorem 4.2: "$\Rightarrow$": By Lemma 4.1, or more precisely, Carl's inequality, we already know that $e_n(S : H \to L_q(\mu)) \prec n^{-1/\alpha}$. Moreover, using $L_q(\mu) \hookrightarrow L_2(\mu)$ we find

$$n^{-1/\alpha} \asymp e_n\big(S : H \to L_2(\mu)\big) \leq \|\,\mathrm{id} : L_q(\mu) \to L_2(\mu)\|\, e_n\big(S : H \to L_q(\mu)\big)\,,$$

and thus $e_n(S : H \to L_q(\mu)) \asymp n^{-1/\alpha}$.

"$\Leftarrow$": By Lemma 4.1, we already know that $d_n(S : H \to L_q(\mu)) \prec n^{-1/\alpha}$. Moreover, by (19), (10), and (6) we obtain $d_n(S : H \to L_2(\mu)) \asymp n^{-1/\alpha}$, and hence we find

$$n^{-1/\alpha} \asymp d_n\big(S : H \to L_2(\mu)\big) \leq \|\,\mathrm{id} : L_q(\mu) \to L_2(\mu)\|\, d_n\big(S : H \to L_q(\mu)\big)\,,$$

that is $d_n(S : H \to L_q(\mu)) \asymp n^{-1/\alpha}$. $\square$

Note that the entropy numbers in condition (19) can be replaced by the Kolmogorov numbers. Indeed, (10) shows that (19) is equivalent to $a_n(S : H \to L_2(\mu)) \asymp n^{-1/\alpha}$, and since we further have $a_n(S) = d_n(S)$, we see that condition (19) can be replaced by

$$d_n\big(S : H \to L_2(\mu)\big) \asymp n^{-1/\alpha}\,. \tag{20}$$

In addition, if H is an RKHS with kernel k and T_k denotes the integral operator associated with k, then (20), or (19), can be replaced by

$$\lambda_n\big(T_k : L_2(\mu) \to L_2(\mu)\big) \asymp n^{-2/\alpha} \tag{21}$$

with the help of (11). The following corollary summarizes our findings in this situation in view of the gap discussed in the introduction.

Corollary 4.3. *Let H be an RKHS of a bounded measurable kernel k on $(X, \mathcal{A})$ and μ be a finite measure on the σ-algebra $\mathcal{A}$. If, in addition, we have*

$$e_n\big(I_{k,\mu} : H \to L_2(\mu)\big) \asymp e_n\big(I_{k,\mu} : H \to L_\infty(\mu)\big) \asymp n^{-1/\alpha}$$

for some $\alpha \in (0, 2)$, then we have $d_n(H, L_\infty(\mu)) \asymp n^{-1/\alpha}$ and $n^{-1/\alpha+1/2} \prec I_n(H, L_\infty(\mu))$.

Proof of Corollary 4.3: The behavior $d_n(I_{k,\mu}: H \to L_\infty(\mu)) \asymp n^{-1/\alpha}$ follows from Theorem 4.2. Moreover, we know $\lambda_i \asymp i^{-2/\alpha}$ by (11) and (13), and therefore, (3) shows

$$n^{-1/\alpha+1/2} \prec \sqrt{\sum_{i=n+1}^{\infty} i^{-2/\alpha}} \prec \sqrt{\sum_{i=n+1}^{\infty} \lambda_i} \leq \sqrt{\mu(X)}\, I_n\big(H, L_\infty(\mu)\big)\,,$$

that is, we have shown the second assertion, too. $\square$

Our last result in this section shows that in the RKHS case and $q = \infty$ the asymptotic behavior $e_n(I_{k,\mu}: H \to L_2(\mu)) \asymp n^{-1/\alpha}$ is inherited from certain interpolation spaces between H and $L_2(\mu)$. For its formulation we need the scale of interpolation spaces of the real method, see e.g. [2, Chapter 5], as well as the notation $[H]_\sim := \{[f]_\sim : f \in H\}$.

Theorem 4.4. *Let H be an RKHS of a bounded measurable kernel k on $(X, \mathcal{A})$ and μ be a finite measure on the σ-algebra $\mathcal{A}$ such that $\mathcal{A}$ is μ-complete and that*

$$e_n\big(I_{k,\mu}: H \to L_2(\mu)\big) \asymp n^{-1/\alpha} \tag{22}$$

for some $\alpha \in (0, 2)$. In addition assume that the interpolation space $[L_2(\mu), [H]_\sim]_{\beta,2}$ is compactly embedded into $L_\infty(\mu)$ for some $\beta \in (\alpha/2, 1]$ with

$$e_n\big([L_2(\mu), [H]_\sim]_{\beta,2} \hookrightarrow L_\infty(\mu)\big) \prec n^{-\beta/\alpha} \tag{23}$$

Then we have

$$e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow L_\infty(\mu)\big) \asymp e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow L_2(\mu)\big) \asymp n^{-\gamma/\alpha} \tag{24}$$

for all $\gamma \in [\beta, 1]$, as well as $e_n(I_{k,\mu}: H \to L_\infty(\mu)) \asymp n^{-1/\alpha}$.

Before we prove this theorem we note that the interpolation spaces $[L_2(\mu), [H]_\sim]_{\gamma,2}$ can be identified as RKHSs provided that the assumptions of Theorem 4.4 are satisfied. For details we refer to the last part of the following proof.

Proof of Theorem 4.4: We first consider the case $\beta \in (\alpha/2, 1)$. Since $\gamma \geq \beta$, we then know $[L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow [L_2(\mu), [H]_\sim]_{\beta,2}$ by e.g. [19, Theorems 4.3 and 4.6], and consequently we have the following diagram of bounded linear embeddings:

$$\begin{tikzcd}
{[L_2(\mu), [H]_\sim]_{\gamma,2}} \arrow[rr] \arrow[dr] & & L_\infty(\mu) \\
& {[L_2(\mu), [H]_\sim]_{\beta,2}} \arrow[ur] &
\end{tikzcd}$$

The multiplicativity of entropy numbers thus yields

$$\begin{aligned}
&e_{2n}\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow L_\infty(\mu)\big) \\
&\leq e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow [L_2(\mu), [H]_\sim]_{\beta,2}\big) \cdot e_n\big([L_2(\mu), [H]_\sim]_{\beta,2} \hookrightarrow L_\infty(\mu)\big)\,.
\end{aligned} \tag{25}$$

Now recall from [19, Equation (36) and Theorem 4.6] that

$$[L_2(\mu), [H]_\sim]_{\gamma,2} = \left\{ \sum_{i \in I} a_i \lambda_i^{\gamma/2} [e_i]_\sim : (a_i) \in \ell_2(I) \right\}, \tag{26}$$

where (λ_i) is the sequence of eigenvalues of the integral operator $T_k : L_2(\mu) \to L_2(\mu)$ and $([e_i]_\sim)$ is a corresponding ONS of eigenfunctions. Moreover, the system $(\lambda_i^{\gamma/2}[e_i]_\sim)$ is an ONB of $[L_2(\mu), [H]_\sim]_{\gamma,2}$ with respect to an equivalent Hilbert space norm on $[L_2(\mu), [H]_\sim]_{\gamma,2}$. Consequently, we have the following diagram for the embedding $[L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow [L_2(\mu), [H]_\sim]_{\beta,2}$:

$$\begin{tikzcd}
{[L_2(\mu), [H]_\sim]_{\gamma,2}} \arrow[r] \arrow[d, "\Phi_\gamma"'] & {[L_2(\mu), [H]_\sim]_{\beta,2}} \\
\ell_2(I) \arrow[r, "D_\Lambda^{(\gamma-\beta)/2}"'] & \ell_2(I) \arrow[u, "\Phi_\beta^{-1}"']
\end{tikzcd}$$

where Φ_γ and Φ_β are the coordinate mappings and $D_\Lambda^{(\gamma-\beta)/2}$ is the diagonal operator associated to the sequence $(\lambda_i^{(\gamma-\beta)/2})$. By (22), (11), and (13) we conclude that $\lambda_i^{(\gamma-\beta)/2} \asymp i^{-(\gamma-\beta)/\alpha}$. Using [4, Proposition 2], which estimates entropy numbers of diagonal operators, and the diagram above, we thus find

$$e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow [L_2(\mu), [H]_\sim]_{\beta,2}\big) \prec n^{-(\gamma-\beta)/\alpha} \,.$$

Combining this with (25), (23), and the fact that μ is a finite measure we obtain

$$e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow L_2(\mu)\big) \prec e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow L_\infty(\mu)\big) \prec n^{-\gamma/\alpha} \,.$$

To establish the lower bound, we recall from [19, Proposition 4.2 and Theorems 5.3 and 4.6] that, for a suitable suitable μ-zero set N, $[L_2(\mu), [H]_\sim]_{\gamma,2}$ can be identified with the RKHS over $X \setminus N$, whose kernel is given by

$$k_\mu^\gamma(x, x') := \sum_{i \in I} \lambda_i^\gamma e_i(x) e_i(x') \,, \qquad\qquad x, x' \in X \setminus N \,.$$

Since the eigenvalues of the corresponding integral operator are $\lambda_i^\gamma \asymp i^{-2\gamma/\alpha}$, we conclude from (11) and (13) that $e_n\big([L_2(\mu), [H]_\sim]_{\gamma,2} \hookrightarrow L_2(\mu)\big) \asymp n^{-\gamma/\alpha}$.

Finally, using $\operatorname{ran} I_{k,\mu} = [H]_\sim$ and Theorem 4.2 the remaining assertions, namely the case $\beta = \gamma = 1$ as well as the assertion for $I_{k,\mu} : H \to L_\infty(\mu)$ can be proven analogously. □

Theorem 4.4 essentially states that the property (24) is passed down from the large spaces in the scale of spaces $[L_2(\mu), [H]_\sim]_{s,2}$ to the smaller ones. Moreover, using the spaces on the right hand side of (26) instead of the interpolation spaces, it can easily be seen that the result is also true for $\gamma > 1$. In addition, the representation (26) suggests that the eigenfunctions may play a crucial role in determining whether (24) holds. In this respect note that [10, Lemma 5.1] essentially showed the continuous embedding $[L_2(\mu), [H]_\sim]_{\alpha/2,1} \hookrightarrow L_\infty(\mu)$ provided that (22) holds and that the eigenfunctions are not only bounded but uniformly bounded. From this it is easy to conclude that $[L_2(\mu), [H]_\sim]_{\beta,2} \hookrightarrow L_\infty(\mu)$ holds for all $\beta \in (\alpha/2, 1)$. In addition, the case $[L_2(\mu), [H]_\sim]_{\beta,2} \hookrightarrow L_\infty(\mu)$ for $\beta \in (0, \alpha/2]$ can always be excluded, since [19, Theorem 5.3] shows that such an inclusion would imply $\sum_{i \geq 1} \lambda_i^\beta < \infty$ for the eigenvalues of the integral operator T_k, and this summability clearly contradicts (22) by (11) and (13). Summarizing, we think that understanding when the asymptotic equivalence (24) holds for some γ close to $\alpha/2$ is an interesting question for future research.

5 An Example: Sobolev Spaces

The goal of this section is to illustrate the consequences of Lemma 4.1 and Theorem 4.2 by applying them to embeddings of the form $\mathrm{id} : H \to L_p(\mu)$, where H is a Sobolev space, $X \subset \mathbb{R}^d$ is a suitable subset, μ is the Lebesgue measure on X, and $p \in [2, \infty]$.

We begin by recalling some basics on Sobolev spaces. To this end let $X \subset \mathbb{R}^d$ be a non-empty, open, and conncected subset satisfying the strong local Lipschitz condition in the sense of [1, p. 83]. For $m \geq 1$ being an integer, we denote the classical Sobolev space on X that is defined by weak derivatives, see e.g. [1, p. 59-60], by $W^m(X) := W^{m,2}(X)$.

For $m > d/2$, it is well-known that the embedding $\mathrm{id} : W^m(X) \to C_B(X)$ is compact, see e.g. [1, Theorem 6.3] in combination with [1, p. 84]. Therefore, the embeddings $\mathrm{id} : W^m(X) \to L_\infty(X)$ are compact, and if X has finite Lebesgue measure, we also obtain the compactness of the the embeddings $\mathrm{id} : W^m(X) \to L_p(X)$, where we followed the standard notation $L_p(X) := L_p(\mu)$. Note that an immediate consequence of this is that the approximation and entropy numbers of these embeddings converge to zero. Let us recall some results from [8] that describe the asymptotic behavior of these numbers. To this end, note that a consequence of Stein's extension theorem, see [1, Theorem 5.24] is that

$$\|f\| := \inf\bigl\{\|g\|_{W^m(\mathbb{R}^d)} : g \in W^m(\mathbb{R}^d) \text{ with } g_{|X} = f\bigr\}\,, \tag{27}$$

where $f \in W^m(X)$, defines an equivalent norm on $W^m(X)$. Moreover, if for $s \in [0,\infty)$ and $p, q, \in (0,\infty]$ we write $B^s_{p,q}(\mathbb{R}^d)$ and $F^s_{p,q}(\mathbb{R}^d)$ for the Besov and Triebel-Lizorkin spaces in the sense of [8, p. 24f], then we have $B^m_{2,2}(\mathbb{R}^d) = F^m_{2,2}(\mathbb{R}^d) = W^m(\mathbb{R}^d)$ by [8, p. 44 and p. 25]. By (27) we conclude that the spaces $B^m_{2,2}(X)$ defined by restrictions as in [8, p. 57] satisfy

$$B^m_{2,2}(X) = W^m(X) \tag{28}$$

up to equivalent norms. Moreover, [8, p. 25] shows that $F^0_{p,2}(\mathbb{R}^d) = L_p(\mathbb{R}^d)$, and by [8, p. 44] we find continuous embeddings $B^0_{p,2}(\mathbb{R}^d) \hookrightarrow L_p(\mathbb{R}^d) \hookrightarrow B^0_{p,p}(\mathbb{R}^d)$ for all $p \in [2,\infty)$. By (27) we conclude that

$$B^0_{p,2}(X) \hookrightarrow L_p(X) \hookrightarrow B^0_{p,p}(X)\,. \tag{29}$$

Similarly, recall that we have continuous embeddings $B^0_{\infty,1}(\mathbb{R}^d) \hookrightarrow L_\infty(\mathbb{R}^d) \hookrightarrow B^0_{\infty,\infty}(\mathbb{R}^d)$, see e.g. [8, p. 44], and thus we also have the continuous embeddings

$$B^0_{\infty,1}(X) \hookrightarrow L_\infty(X) \hookrightarrow B^0_{\infty,\infty}(X)\,. \tag{30}$$

Let us now assume that X is open, connected, and bounded, and that it has a C^∞-boundary. Moreover, we fix some $s_1, s_2 \in [0,\infty)$ and $p_1, p_2, q_1, q_2 \in (0,\infty]$ such that $s_1 - s_2 > d\bigl(\frac{1}{p_1} - \frac{1}{p_2}\bigr)_+$. Then [8, Theorem 2 on p. 118] shows that

$$e_n\bigl(\mathrm{id} : B^{s_1}_{p_1,q_1}(X) \to B^{s_2}_{p_2,q_2}(X)\bigr) \asymp n^{-(s_1-s_2)/d}\,, \tag{31}$$

and if we additionally assume that $2 \leq p_1 \leq p_2 \leq \infty$, then [8, p. 119] shows that

$$a_n\bigl(\mathrm{id} : B^{s_1}_{p_1,q_1}(X) \to B^{s_2}_{p_2,q_2}(X)\bigr) \asymp n^{-(s_1-s_2)/d+1/p_1-1/p_2}\,. \tag{32}$$

In particular, for $s_1 = s$, $p_1 = q_1 = 2$, $s_2 = 0$, $p_2 = p \in [2,\infty]$, and $q_2 = q \in [1,\infty]$ with $s > d(\frac{1}{2} - \frac{1}{p})$ we obtain

$$\begin{aligned} e_n\bigl(\mathrm{id} : B^s_{2,2}(X) \to B^0_{p,q}(X)\bigr) &\asymp n^{-s/d} \\ a_n\bigl(\mathrm{id} : B^s_{2,2}(X) \to B^0_{p,q}(X)\bigr) &\asymp n^{-s/d+1/2-1/p}\,. \end{aligned}$$

By (28), (29), and (30) we conclude that

$$e_n\bigl(\mathrm{id} : W^m(X) \to L_p(X)\bigr) \asymp n^{-m/d} \tag{33}$$

$$a_n\bigl(\mathrm{id} : W^m(X) \to L_p(X)\bigr) \asymp n^{-m/d+1/2-1/p} \tag{34}$$

for all $m \in \mathbb{N}$ with $m > d(\frac{1}{2} - \frac{1}{p})$. In other words, the gap between the entropy and approximation numbers is of the order $n^{1/2-1/p}$. Note that for the Hilbert space case, i.e. $p = 2$, the gap vanishes as already observed in Section 3, while in the other extreme $p = \infty$, the gap is of the order $n^{1/2}$. Finally, we see by (27) that these asymptotics still hold, if we only assume that X is an open, connected, and bounded subset of $\mathbb{R}^d$ satisfying the strong local Lipschitz condition.

To illustrate these findings, we now consider the linear interpolation n-width mentioned in the introduction. To this end, we fix an $m \in \mathbb{N}$ with $m > d/2$ and let $H = W^m(X)$ with equivalent norms. Then (34) shows

$$n^{-m/d+1/2-1/p} \asymp a_{n+1}\big((\mathrm{id} : H \to L_p(X)\big) \leq I_n\big(H, L_p(X)\big)$$

for all $p \in [2, \infty]$. Here we note that in the case $p = \infty$, the lower bound $n^{-m/d+1/2} \prec I_n\big(H, L_\infty(X)\big)$ already follows from (3). Moreover, (33) in combination with Theorem 4.2 yields

$$d_n\big(\mathrm{id} : H \to L_p(X)\big) \asymp n^{-m/d} \tag{35}$$

for all $p \in [2, \infty]$. In other words, the gap of $1/2 - 1/p$ actually occurs between the non-linear approximation described by d_n and the linear approximation described by a_n. Moreover, the gap is maximal for $p = \infty$ and vanishes in the other extreme case $p = 2$.

We like to mention that (35) appears to be new, since it is not contained in the list of known asymptotics compiled in [21]. In addition, the gap between d_n and a_n is solely derived from the same gap between e_n and a_n, that is, from (33) and (34). In other words, we will observe a gap between d_n and a_n if and only if there is a gap between e_n and a_n. Fortunately, the latter two quantities have been considered for various other spaces H and measures μ, so that it should be possible to compile a list of cases, in which the gap occurs.

For convenience, the following corollary summarizes our findings for sufficiently large subspaces of $W^m(X)$. It particularly applies to kernels of many standard Gaussian processes, such as the (iterated) Brownian motion and -bridge, see e.g. the numerical example in [14].

Corollary 5.1. *Let $X \subset \mathbb{R}^d$ be an open, connected, and bounded subset satisfying the strong Lipschitz condition. Moreover, let H be an RKHS over X with kernel k such that $H \hookrightarrow W^m(X)$ for some integer $m > d/2$. Assume, in addition, that*

$$e_n\big(I_{k,\mu} : H \to L_2(X)\big) \asymp n^{-m/d}$$

holds. Then we have

$$d_n\big(H, L_\infty(\mu)\big) \asymp n^{-m/d} \qquad \textit{and} \qquad n^{-m/d+1/2} \prec I_n\big(H, L_\infty(\mu)\big)\,.$$

In addition, if $H = W^m(X)$ with equivalent norms, then, for all $p \in [2, \infty]$, we have

$$d_n\big(H, L_p(\mu)\big) \asymp n^{-m/d} \qquad \textit{and} \qquad n^{-m/d+1/2-1/p} \asymp a_n\big(H, L_p(\mu)\big)\,.$$

Proof of Corollary 5.1: We first note that the sequence of estimates

$$n^{-m/d} \asymp e_n\big(I_{k,\mu} : H \to L_2(X)\big) \prec e_n\big(I_{k,\mu} : H \to L_\infty(X)\big) \prec e_n\big(I_{k,\mu} : W^m \to L_\infty(X)\big) \asymp n^{-m/d}$$

yields $e_n\big(I_{k,\mu} : H \to L_\infty(X)\big) \asymp n^{-m/d}$, and therefore Corollary 4.3 shows the first two assertions. The second set of asymptotic equivalences immediately follows from our findings in the text above together with the muliplicativity of the approximation numbers. □

Acknowledgement. I deeply thank R. Schaback, for pointing me to the question regarding the gap between Kolmogorov and interpolation widths as well as for his feedback and his patience at the Shanghai airport. I also thank G. Santin for giving valuable hints.

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. Academic Press, New York, 2nd edition, 2003.

[2] C. Bennett and R. Sharpley. *Interpolation of Operators*. Academic Press, Boston, 1988.

[3] J. Bourgain, A. Pajor, S. J. Szarek, and N. Tomczak-Jaegermann. On the duality problem for entropy numbers of operators. *Lecture Notes in Math. 1376*, pages 50–63, 1989.

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

[5] B. Carl. Inequalities of Bernstein-Jackson-type and the degree of compactness of operators in Banach spaces. *Ann. Inst. Fourier*, 35:79–118, 1985.

[6] B. Carl and I. Stephani. *Entropy, Compactness and the Approximation of Operators*. Cambridge University Press, Cambridge, 1990.

[7] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely Summing Operators*. Cambridge University Press, Cambridge, 1995.

[8] D. E. Edmunds and H. Triebel. *Function Spaces, Entropy Numbers, Differential Operators*. Cambridge University Press, Cambridge, 1996.

[9] G. G. Lorentz, M. v. Golitschek, and Y. Makovoz. *Constructive Approximation. Advanced Problems*. Springer-Verlag, Berlin, 1996.

[10] S. Mendelson and J. Neeman. Regularization in kernel learning. *Ann. Statist.*, 38:526–565, 2010.

[11] A. Pajor and N. Tomczak-Jaegermann. Remarques sur les nombres d'entropie d'un opérateur et de son transposé. *C. R. Acad. Sci. Paris, Ser. I*, 301:743–746, 1985.

[12] A. Pietsch. *Eigenvalues and s-Numbers*. Geest & Portig K.-G., Leipzig, 1987.

[13] A. Pinkus. *n-Widths in Approximation Theory*. Springer-Verlag, Berlin, 1985.

[14] G. Santin and R. Schaback. Approximation of eigenfunctions in kernel-based spaces. *Adv. Comput. Math.*, pages 1–21, 2016.

[15] R. Schaback and H. Wendland. Inverse and saturation theorems for radial basis function interpolation. *Math. Comp.*, 71:669–681, 2002.

[16] R. Schaback and H. Wendland. Approximation by positive definite kernels. In M.D. Buhmann and D.H. Mache, editors, *Advanced Problems in Constructive Approximation: 3rd International Dortmund Meeting on Approximation Theory (IDoMAT) 2001*, pages 203–222. Birkhäuser Basel, Basel, 2003.

[17] I. Steinwart. Entropy of $C(K)$-valued operators. *J. Approx. Theory*, 103:302–328, 2000.

[18] I. Steinwart. *Entropy of $C(K)$-valued operators and some applications*. PhD thesis, Friedrich-Schiller Universität Jena, Fakultät für Mathematik und Informatik, 2000.

[19] I. Steinwart and C. Scovel. Mercer's theorem on general domains: on the interaction between measures, kernels, and RKHSs. *Constr. Approx.*, 35:363–417, 2012.

[20] N. Tomczak-Jaegermann. Dualite des nombres d'entropie pour des opérateurs a valeurs dans un espace de Hilbert. *C. R. Acad. Sci. Paris, Ser. I*, 305:299–301, 1987.

[21] J. Vybíral. Widths of embeddings in function spaces. *J. Complexity*, 24:545–570, 2008.