Tensor power sequences and the approximation of tensor product operators

David Krieg (david.krieg@uni-jena.de)

October 10, 2017

Abstract

The approximation numbers of the L_2-embedding of mixed order Sobolev functions on the d-torus are well studied. They are given as the nonincreasing rearrangement of the dth tensor power of the approximation number sequence in the univariate case. I present results on the asymptotic and preasymptotic behavior for tensor powers of arbitrary sequences of polynomial decay. This can be used to study the approximation numbers of many other tensor product operators, like the embedding of mixed order Sobolev functions on the d-cube into $L_2\left([0,1]^d\right)$ or the embedding of mixed order Jacobi functions on the d-cube into $L_2\left([0,1]^d, w_d\right)$ with Jacobi weight w_d.

1 Introduction and Results

Let $\sigma : \mathbb{N} \to \mathbb{R}$ be a nonincreasing zero sequence. For any natural number d, its dth tensor power is the sequence $\sigma_d : \mathbb{N}^d \to \mathbb{R}$, where

$$\sigma_d(n_1, \ldots, n_d) = \prod_{j=1}^{d} \sigma(n_j). \tag{1.1}$$

Any such sequence σ_d can then be uniquely rearranged to a nonincreasing zero sequence $\tau : \mathbb{N} \to \mathbb{R}$. Tensor power sequences like this occur naturally in the study of approximation numbers of tensor power operators. If σ is the sequence of approximation numbers of a compact operator between two Hilbert spaces, then τ is the sequence of approximation numbers of the compact dth tensor power operator between the tensor power spaces.

What can we say about the behavior of τ based on the behavior of σ? A classical result of Babenko [B60] and Mityagin [M62] is concerned with the speed of decay of these sequences:

Theorem 1. *Let σ be a nonincreasing zero sequence and τ be the nonincreasing rearrangement of its dth tensor power. For any $s > 0$, the following holds.*

(i) If $\sigma(n) \preccurlyeq n^{-s}$, then $\tau(n) \preccurlyeq n^{-s} (\log n)^{s(d-1)}$.

(ii) If $\sigma(n) \succcurlyeq n^{-s}$, then $\tau(n) \succcurlyeq n^{-s} (\log n)^{s(d-1)}$.

Here, the symbol $\preccurlyeq$ (respectively $\succcurlyeq$) means that the left (right) hand side is bounded above by a constant multiple of the right (left) hand side for all $n \in \mathbb{N}$. Of course, other decay assumptions on σ may be of interest. For instance, Pietsch [P82] and König [K84] study the decay of τ, if σ lies in the Lorentz sequence space $\ell_{p,q}$ for positive indices p and q, which is a stronger assumption than (i) for $s = 1/p$ but weaker than (i) for any $s > 1/p$. However, since we are motivated by the example of Sobolev embeddings, we will stick to the assumptions of Theorem 1. One of the problems with this theorem is that it does not provide explicit estimates for $\tau(n)$, even if n is huge. This is because of the constants hidden in the notation. But Theorem 1 can be sharpened.

Theorem 2. *Let σ be a nonincreasing zero sequence and τ be the nonincreasing rearrangement of its dth tensor power. For $c > 0$ and $s > 0$, the following holds.*

(i) If $\sigma(n) \lesssim c\, n^{-s}$, then $\tau(n) \lesssim \frac{c^d}{(d-1)!^s}\, n^{-s} (\log n)^{s(d-1)}$.

(ii) If $\sigma(n) \gtrsim c\, n^{-s}$, then $\tau(n) \gtrsim \frac{c^d}{(d-1)!^s}\, n^{-s} (\log n)^{s(d-1)}$.

We write $f(n) \lesssim g(n)$ for positive sequences f and g and say that $f(n)$ is asymptotically smaller or equal than $g(n)$, if the limit superior of $f(n)/g(n)$ is at most one as n tends to infinity. Analogously, $f(n)$ is asymptotically greater than or equal to $g(n)$, write $f(n) \gtrsim g(n)$, if the limit inferior of this ratio is at least one. Finally, we say $f(n)$ is asymptotically equal to $g(n)$ and write $f(n) \simeq g(n)$ if the limit of the ratio equals one. In particular, we obtain that $\sigma(n) \simeq c\, n^{-s}$ implies that $\tau(n) \simeq \frac{c^d}{(d-1)!^s}\, n^{-s} (\log n)^{s(d-1)}$. Theorem 2 is due to Theorem 4.3 in [KSU15]. There, Kühn, Sickel and Ullrich prove this asymptotic equality in an interesting special case: τ is the sequence of approximation numbers for the L_2-embedding of the tensor power space $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ on the d-torus $[0, 2\pi]^d$, equipped with a tensor product norm. The statement can be deduced from this special case with the help of their Lemma 4.14. However, we prefer to give a direct proof in Section 2 by generalizing the proof of Theorem 4.3 in [KSU15].

Theorem 2 gives us a pretty good understanding of the asymptotic behavior of the dth tensor power τ of a sequence σ of polynomial decay. If $\sigma(n)$ is roughly $c\, n^{-s}$ for large n, then $\tau(n)$ is roughly $c^d \left(\frac{(\log n)^{d-1}}{(d-1)!}\right)^s n^{-s}$ for n larger than a certain threshold. But even for modest values of d, the size of this threshold may go far beyond the scope of computational capabilities. Indeed, while τ decreases, the function $n^{-s} (\log n)^{s(d-1)}$ grows rapidly as n goes from 1 to e^{d-1}. For $n^{-s} (\log n)^{s(d-1)}$ to become less than one, n even has to be super exponentially large in d. Thus, any estimate for the sequence τ in terms of $n^{-s} (\log n)^{s(d-1)}$ is useless to describe its behavior in the range $n \leq 2^d$, its so called preasymptotic behavior. As a replacement, we will prove the following estimate in Section 3.

Theorem 3. *Let σ be a nonincreasing zero sequence and τ be the nonincreasing rearrangement of its dth tensor power. Let $\sigma(1) > \sigma(2) > 0$ and assume that $\sigma(n) \leq C\, n^{-s}$ for some $s, C > 0$ and all $n \geq 2$. For any $n \in \left\{2, \ldots, 2^d\right\}$,*

$$\frac{\sigma(2)}{\sigma(1)} \cdot \left(\frac{1}{n}\right)^{\frac{\log(\sigma(1)/\sigma(2))}{\log\left(1+\frac{d}{\log_2 n}\right)}} \leq \frac{\tau(n)}{\tau(1)} \leq \left(\frac{\exp\left(\left(C/\sigma(1)\right)^{2/s}\right)}{n}\right)^{\frac{\log(\sigma(1)/\sigma(2))}{\log\left((\sigma(1)/\sigma(2))^{2/s}\, d\right)}}.$$

Let us assume the power (or dimension) d to be large. Then the tensor power sequence, which roughly decays like n^{-s} for huge values of n, roughly decays like n^{-t_d} with $t_d = \log\left(\sigma(1)/\sigma(2)\right) / \log d$ for small values of n. This is why I will refer to t_d as preasymptotic rate of the tensor power sequence. The preasymptotic rate is much worse than the asymptotic rate. This is not an unusual phenomenon for high-dimensional problems. Comparable estimates for the case of τ being the sequence of approximation numbers of the embedding $H^s_{\text{mix}}\left(\mathbb{T}^d\right) \hookrightarrow L_2\left(\mathbb{T}^d\right)$ are established in Theorem 4.9, 4.10, 4.17 and 4.20 of [KSU15]. See [CW16], [KMU16] or [CW17] for other examples. An interesting consequence of these preasymptotic estimates is the following tractability result. For each $d \in \mathbb{N}$, let T_d be a compact norm-one operator between two Hilbert spaces and let T_d^d be its dth tensor power. Assume that the corresponding approximation numbers $a_n\left(T_d\right)$ are nonincreasing in d and that $a_n\left(T_1\right)$ decays polynomially in n. Then the problem of approximating T_d^d by linear functionals is strongly polynomially tractable, iff it is polynomially tractable, iff $a_2\left(T_d\right)$ decays polynomially in d.

In Section 4, these results will be applied to the L_2-approximation of mixed order Sobolev functions on the d-torus, as well as mixed order Jacobi and Sobolev functions on the d-cube, taking different normalizations into account. For instance, we will consider the L_2-embedding

$$T_s^d : H^s_{\text{mix}}\left([0,1]^d\right) \hookrightarrow L_2\left([0,1]^d\right) \tag{1.2}$$

of the d-variate Sobolev space $H^s_{\text{mix}}\left([0,1]^d\right)$ with dominating mixed smoothness $s \in \mathbb{N}$, equipped with the scalar product

$$\langle f, g \rangle = \sum_{\alpha \in \{0,\ldots,s\}^d} \langle D^\alpha f, D^\alpha g \rangle_{L_2}. \tag{1.3}$$

Let $\widetilde{T}_s^d$ be the restriction of T_s^d to the subspace $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ of periodic functions. Theorem 2 yields that the approximation numbers of these embeddings satisfy

$$\lim_{n\to\infty} \frac{a_n(T_s^d) \cdot n^s}{(\log n)^{s(d-1)}} = \lim_{n\to\infty} \frac{a_n(\widetilde{T}_s^d) \cdot n^s}{(\log n)^{s(d-1)}} = \left(\pi^d \cdot (d-1)!\right)^{-s}. \tag{1.4}$$

In particular, they do not only have the same rate of convergence, but even the limit of their ratio is one. This means that the L_2-approximation of mixed order Sobolev functions on the d-cube with n linear functionals is just as hard for nonperiodic functions as for periodic functions, if n is large enough. The preasymptotic rate $\tilde{t}_d$ for the periodic case satisfies

$$\frac{s \cdot \log(2\pi)}{\log d} \leq \tilde{t}_d \leq \frac{s \cdot \log(2\pi) + 1}{\log d}. \tag{1.5}$$

Although this is significantly worse than the asymptotic main rate s, it still grows linearly with the smoothness. An increasing dimension can hence be neutralized by increasing the smoothness of the functions. In contrast, the preasymptotic rate t_d for the nonperiodic case satisfies

$$\frac{1.2803}{\log d} \leq t_d \leq \frac{1.2825}{\log d} \tag{1.6}$$

for any $s \geq 2$. This means that increasing the smoothness of the functions beyond $s = 2$ in the nonperiodic setting is a very ineffective way of reducing the approximation error. The L_2-approximation of mixed order Sobolev functions on the d-cube with less than 2^d linear functionals is hence much harder for nonperiodic functions than for periodic functions. This is also reflected in the corresponding tractability results: The approximation problem $\{\widetilde{T}^d_{s_d}\}$ is (strongly) polynomially tractable, iff the smoothness s_d grows at least logarithmically with the dimension, whereas the approximation problem $\{T^d_{s_d}\}$ is never (strongly) polynomially tractable. A similar effect for functions with coordinatewise increasing smoothness has already been observed by Papageorgiou and Woźniakowski in [PW10]. However, the tractability result for the space of periodic functions heavily depends on the side length $b-a$ of the torus $\mathbb{T}^d = [a,b]^d$. If it is less than 2π, (strong) polynomial tractability is equivalent to logarithmic increase of the smoothness. If it equals 2π, (strong) polynomial tractability is equivalent to polynomial increase of the smoothness. If it is larger than 2π, there cannot be (strong) polynomial tractability. These tractability results and interpretations can be found in Section 5.

2 Asymptotic Behavior of Tensor Power Sequences

Let σ be a nonincreasing zero sequence and τ be the nonincreasing rearrangement of its dth tensor power. Fix some $s>0$ and let us consider the quantities

$$C_1 = \limsup_{n\to\infty} \sigma(n) n^s, \qquad c_1 = \liminf_{n\to\infty} \sigma(n) n^s,$$
$$C_d = \limsup_{n\to\infty} \frac{\tau(n)\cdot n^s}{(\log n)^{s(d-1)}}, \qquad c_d = \liminf_{n\to\infty} \frac{\tau(n)\cdot n^s}{(\log n)^{s(d-1)}}.$$

These limits may be both infinite or zero. They can be interpreted as asymptotic or optimal constants for the bounds

$$\tau(n) \leq C\cdot n^{-s} (\log n)^{s(d-1)} \quad \text{and} \tag{2.1}$$
$$\tau(n) \geq c\cdot n^{-s} (\log n)^{s(d-1)}. \tag{2.2}$$

For any $C > C_d$ respectively $c < c_d$ there is a threshold $n_0 \in \mathbb{N}$ such that (2.1) respectively (2.2) holds for all $n \geq n_0$, whereas for any $C < C_d$ respectively $c > c_d$ there is no such threshold. Theorem 1 states that C_d is finite, whenever C_1 is finite, whereas c_d is positive, whenever c_1 is positive. Theorem 2 is more precise. It states that

$$\frac{c_1^d}{(d-1)!^s} \leq c_d \leq C_d \leq \frac{C_1^d}{(d-1)!^s}. \tag{2.3}$$

In this section, we will give its proof. We will also show that equality can but does not always hold. Note that the proof provides a possibility to track down admissible thresholds n_0 for any $C > \frac{C_1^d}{(d-1)!^s}$ respectively any $c < \frac{c_1^d}{(d-1)!^s}$.

For the proof, it will be essential to study the asymptotics of the cardinalities

$$A_N(r,l) = \#\left\{ \boldsymbol{n} \in \{N, N+1, \ldots\}^l \mid \prod_{j=1}^{l} n_j \leq r \right\} \tag{2.4}$$

for $l \in \{1, \ldots, d\}$ and $N \in \mathbb{N}$ as $r \to \infty$. In [KSU15, Lemma 3.2], it is shown that

$$r\left(\frac{\left(\log \frac{r}{2^l}\right)^{l-1}}{(l-1)!} - \frac{\left(\log \frac{r}{2^l}\right)^{l-2}}{(l-2)!} \right) \leq A_2(r,l) \leq r\frac{(\log r)^{l-1}}{(l-1)!} \tag{2.5}$$

for $l \geq 2$ and $r \in \left\{4^l, 4^l+1, \ldots\right\}$, see also [CD16, Theorem 3.4]. Consequently, we have

$$\lim_{r\to\infty} \frac{A_N(r,l)}{r\,(\log r)^{l-1}} = \frac{1}{(l-1)!} \tag{2.6}$$

for $N = 2$. In fact, (2.6) holds true for any $N \in \mathbb{N}$. This can be derived from the case $N = 2$, but for the reader's convenience, I will give a complete proof.

Lemma 1.

$$\lim_{r\to\infty} \frac{A_N(r,l)}{r\,(\log r)^{l-1}} = \frac{1}{(l-1)!}.$$

Proof. Note that for all values of the parameters,

$$A_N(r, l+1) = \sum_{k=N}^{\infty} A_N\left(\frac{r}{k}, l\right), \tag{2.7}$$

where $A_N\left(\frac{r}{k}, l\right) = 0$ for $k > \frac{r}{N^l}$. This allows a proof by induction on $l \in \mathbb{N}$.

Like in estimate (2.5), we first show that

$$A_2(r,l) \leq r\frac{(\log r)^{l-1}}{(l-1)!} \tag{2.8}$$

for any $l \in \mathbb{N}$ and $r \geq 1$. This is obviously true for $l = 1$. On the other hand, if this relation holds for some $l \in \mathbb{N}$ and if $r \geq 1$, then

$$\begin{aligned} A_2(r, l+1) &= \sum_{k=2}^{\lfloor r \rfloor} A_2\left(\frac{r}{k}, l\right) \leq \sum_{k=2}^{\lfloor r \rfloor} \frac{r\left(\log \frac{r}{k}\right)^{l-1}}{k\,(l-1)!} \\ &\leq \frac{r}{(l-1)!} \int_1^r \frac{\left(\log \frac{r}{x}\right)^{l-1}}{x}\,\mathrm{d}x = \frac{r}{(l-1)!}\left[-\frac{1}{l}\left(\log \frac{r}{x}\right)^l\right]_1^r = r\frac{(\log r)^l}{l!} \end{aligned} \tag{2.9}$$

and (2.8) is proven. In particular, we have

$$\limsup_{r\to\infty} \frac{A_N(r,l)}{r\,(\log r)^{l-1}} \le \frac{1}{(l-1)!} \tag{2.10}$$

for $l \in \mathbb{N}$ and $N = 2$. Clearly, the same holds for $N \ge 2$, since $A_N(r,l)$ is decreasing in N. Relation (2.10) for $N = 1$ follows from the case $N = 2$ by the identity

$$\begin{aligned} A_1(r,l) &= \sum_{m=0}^{l} \# \left\{ \boldsymbol{n} \in \mathbb{N}^l \mid \# \{1 \le j \le l \mid n_j \ne 1\} = m \wedge \prod_{j=1}^{d} n_j \le r \right\} \\ &= \mathbf{1}_{r\ge 1} + \sum_{m=1}^{l} \binom{l}{m} \cdot A_2(r,m). \end{aligned} \tag{2.11}$$

It remains to prove

$$\liminf_{r\to\infty} \frac{A_N(r,l)}{r\,(\log r)^{l-1}} \ge \frac{1}{(l-1)!} \tag{2.12}$$

for $N \in \mathbb{N}$ and $l \in \mathbb{N}$. Again, this is obvious for $l = 1$. Suppose, (2.12) holds for some $l \in \mathbb{N}$ and let $b < 1$. Then there is some $r_0 \ge 1$ such that

$$A_N(r,l) \ge br \frac{(\log r)^{l-1}}{(l-1)!} \tag{2.13}$$

for all $r \ge r_0$ and hence

$$\begin{aligned} &A_N(r,l+1) \ge \sum_{k=N}^{\lfloor r/r_0 \rfloor} A_N\left(\frac{r}{k}, l\right) \ge \sum_{k=N}^{\lfloor r/r_0 \rfloor} \frac{br\left(\log \frac{r}{k}\right)^{l-1}}{k\,(l-1)!} \\ &\ge \frac{br}{(l-1)!} \int_N^{\frac{r}{r_0}} \frac{\left(\log \frac{r}{x}\right)^{l-1}}{x}\,\mathrm{d}x = \frac{br}{l!}\left(\left(\log \frac{r}{N}\right)^l - (\log r_0)^l\right) \ge b^2 r \frac{(\log r)^l}{l!} \end{aligned} \tag{2.14}$$

for large r. Since this is true for any $b < 1$, the induction step is complete. □

Proof of Theorem 2. Without loss of generality, we can assume that $s = 1$ and $\sigma(1) = 1$. If $\sigma(1) \ne 0$, the stated inequalities follow from the corresponding inequalities for the sequence $\tilde{\sigma} = (\sigma/\sigma(1))^{1/s}$. If $\sigma(1) = 0$, they are trivial.

Proof of (i): Let $c_3 > c_2 > c_1 > c$. There is some $N \in \mathbb{N}$ such that for any $n \ge N$, we have

$$\sigma(n) \le c_1\, n^{-1}. \tag{2.15}$$

We want to prove

$$\limsup_{n\to\infty} \frac{\tau(n)\, n}{(\log n)^{d-1}} \le \frac{c^d}{(d-1)!}. \tag{2.16}$$

Since $n/(\log n)^{d-1}$ is finally increasing, instead of giving an upper bound for $\tau(n)$ in terms of n, we can just as well give an upper bound for n in terms of $\tau(n)$ to obtain (2.16). Clearly, there are at least n elements in the tensor power sequence greater than or equal to $\tau(n)$ and hence

$$\begin{aligned} n &\le \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \sigma_d(\boldsymbol{n}) \ge \tau(n)\right\} \\ &= \sum_{l=0}^{d} \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \#\{1 \le j \le d \mid n_j \ge N\} = l \wedge \sigma_d(\boldsymbol{n}) \ge \tau(n)\right\} \\ &\overset{\sigma(1)=1}{\le} \sum_{l=0}^{d} \binom{d}{l} N^{d-l}\, \#\left\{\boldsymbol{n} \in \{N, N+1, \dots\}^l \mid \sigma_d(\boldsymbol{n}) \ge \tau(n)\right\}. \end{aligned} \tag{2.17}$$

For every $\boldsymbol{n}$ in the last set, relation (2.15) implies that $\prod_{j=1}^{d} n_j \le c_1^l\, \tau(n)^{-1}$. Thus,

$$n \le \sum_{l=0}^{d} \binom{d}{l} N^{d-l} A_N\left(c_1^l\, \tau(n)^{-1}, l\right). \tag{2.18}$$

Lemma 1 yields that, if n and hence $c_1^l\, \tau(n)^{-1}$ is large enough,

$$A_N\left(c_1^l\, \tau(n)^{-1}, l\right) \le \frac{c_2^l\, \tau(n)^{-1}}{(l-1)!} \left(\log\left(c_2^l\, \tau(n)^{-1}\right)\right)^{l-1} \tag{2.19}$$

for $l \in \{1, \dots, d\}$. Letting $n \to \infty$, the term for $l = d$ is dominant and hence

$$n \le \frac{c_3^d\, \tau(n)^{-1}}{(d-1)!} \left(\log\left(c_3^d\, \tau(n)^{-1}\right)\right)^{d-1} \tag{2.20}$$

for large values of n. By the monotonicity of $n/(\log n)^{d-1}$, we obtain

$$\frac{\tau(n)\, n}{(\log n)^{d-1}} \le \frac{c_3^d}{(d-1)!} \cdot \left(\frac{\log\left(c_3^d \tau(n)^{-1}\right)}{\log\left(\tau(n)^{-1} \cdot \frac{c_3^d}{(d-1)!} \left(\log\left(c_3^d \tau(n)^{-1}\right)\right)^{d-1}\right)} \right)^{d-1}. \tag{2.21}$$

The fraction in brackets tends to one as n and hence $\tau(n)^{-1}$ tends to infinity and thus

$$\limsup_{n\to\infty} \frac{\tau(n)\, n}{(\log n)^{d-1}} \le \frac{c_3^d}{(d-1)!}. \tag{2.22}$$

Since this is true for any $c_3 > c$, the proof of (2.16) is complete.

Proof of (*ii*): Let $0 < c_3 < c_2 < c_1 < c$. There is some $N \in \mathbb{N}$ such that for any $n \geq N$, we have

$$\sigma(n) \geq c_1 \, n^{-1}. \tag{2.23}$$

We want to prove

$$\liminf_{n\to\infty} \frac{\tau(n)\,n}{(\log n)^{d-1}} \geq \frac{c^d}{(d-1)!^s} \tag{2.24}$$

for any $d \in \mathbb{N}$. Clearly, there are at most $n-1$ elements in the tensor power sequence greater than $\tau(n)$ and hence

$$n > \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \sigma_d(\boldsymbol{n}) > \tau(n)\right\} \geq \#\left\{\boldsymbol{n} \in \{N, N+1, \ldots\}^d \mid \sigma_d(\boldsymbol{n}) > \tau(n)\right\}. \tag{2.25}$$

Relation (2.23) implies that every $\boldsymbol{n} \in \{N, N+1, \ldots\}^d$ with $\prod_{j=1}^d n_j < c_1^d \, \tau(n)^{-1}$ is contained in the last set. This observation and Lemma 1 yield that

$$n > A_N\left(c_2^d \, \tau(n)^{-1}, d\right) \geq \frac{c_3^d \, \tau(n)^{-1}}{(d-1)!} \left(\log\left(c_3^d \, \tau(n)^{-1}\right)\right)^{d-1} \tag{2.26}$$

for sufficiently large n. By the monotonicity of $n/\left(\log n\right)^{d-1}$ for large n, we obtain

$$\frac{\tau(n)\,n}{(\log n)^{d-1}} \geq \frac{c_3^d}{(d-1)!} \cdot \left(\frac{\log\left(c_3^d \tau(n)^{-1}\right)}{\log\left(\frac{c_3^d}{(d-1)!}\left(\log\left(c_3^d \tau(n)^{-1}\right)\right)^{d-1} \tau(n)^{-1}\right)}\right)^{d-1}. \tag{2.27}$$

The fraction in brackets tends to one as n and hence $\tau(n)^{-1}$ tends to infinity and thus

$$\liminf_{n\to\infty} \frac{\tau(n)\,n}{(\log n)^{d-1}} \geq \frac{c_3^d}{(d-1)!}. \tag{2.28}$$

Since this is true for any $c_3 < c$, the proof of (2.24) is complete. □

This proves the relations (2.3) of the asymptotic constants. Obviously, there must be equality in all these relations, if the limit of $\sigma(n)\,n^s$ for $n \to \infty$ exists. It is natural to ask, whether any of these equalities always holds true. The answer is no, as shown by the following example.

Example 1. The sequence σ, defined by $\sigma(n) = 2^{-k}$ for $n \in \left\{2^k, \ldots, 2^{k+1}-1\right\}$ and $k \in \mathbb{N}_0$, decays linearly in n, but is constant on segments of length 2^k. It satisfies

$$C_1 = \limsup_{n\to\infty} \sigma(n) n = \lim_{k\to\infty} 2^{-k} \cdot \left(2^{k+1}-1\right) = 2 \tag{2.29}$$

and

$$c_1 = \liminf_{n\to\infty} \sigma(n)n = \lim_{k\to\infty} 2^{-k} \cdot 2^k = 1. \tag{2.30}$$

Also the values of the nonincreasing rearrangement τ of its dth tensor power are of the form 2^{-k} for some $k \in \mathbb{N}_0$, where

$$\begin{aligned}
&\#\left\{n \in \mathbb{N} \mid \tau(n) = 2^{-k}\right\} = \sum_{|\boldsymbol{k}|=k} \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \sigma(n_j) = 2^{-k_j} \text{ for } j = 1 \dots d\right\} \\
&= \sum_{|\boldsymbol{k}|=k} 2^k = 2^k \cdot \binom{k+d-1}{d-1} = \frac{2^k}{(d-1)!} \cdot (k+1) \cdot \ldots \cdot (k+d-1).
\end{aligned} \tag{2.31}$$

Hence, $\tau(n) = 2^{-k}$ for $N(k-1,d) < n \leq N(k,d)$ with $N(-1,d) = 0$ and

$$N(k,d) = \sum_{j=0}^{k} \frac{2^j}{(d-1)!} \cdot (j+1) \cdot \ldots \cdot (j+d-1) \tag{2.32}$$

for $k \in \mathbb{N}_0$. The monotonicity of $n/\left(\log n\right)^{d-1}$ for large n implies

$$C_d = \limsup_{n\to\infty} \frac{\tau(n)\cdot n}{\left(\log n\right)^{d-1}} = \lim_{k\to\infty} \frac{2^{-k}\cdot N(k,d)}{\left(\log N(k,d)\right)^{d-1}} \tag{2.33}$$

and

$$c_d = \liminf_{n\to\infty} \frac{\tau(n)\cdot n}{\left(\log n\right)^{d-1}} = \lim_{k\to\infty} \frac{2^{-k}\cdot N(k-1,d)}{\left(\log N(k-1,d)\right)^{d-1}}. \tag{2.34}$$

We insert the relations

$$N(k,d) \leq \frac{(k+d)^{d-1}}{(d-1)!} \sum_{j=0}^{k} 2^j \leq \frac{2^{k+1}\cdot (k+d)^{d-1}}{(d-1)!} \tag{2.35}$$

and

$$N(k,d) \geq \frac{(k-l)^{d-1}}{(d-1)!} \sum_{j=k-l+1}^{k} 2^j = \frac{2^{k+1}(k-l)^{d-1}}{(d-1)!} \left(1-2^{-l}\right) \tag{2.36}$$

for arbitrary $l \in \mathbb{N}$ in (2.33) and (2.34) and obtain

$$C_d = 2 \cdot \frac{\left(\log_2 e\right)^{d-1}}{(d-1)!} \qquad \text{and} \qquad c_d = \frac{\left(\log_2 e\right)^{d-1}}{(d-1)!}. \tag{2.37}$$

In particular,

$$\frac{c_1^d}{(d-1)!} < c_d < C_d < \frac{C_1^d}{(d-1)!} \quad \text{for } d \neq 1. \tag{2.38}$$

More generally, the tensor product of d nonincreasing zero sequences $\sigma^{(j)} : \mathbb{N} \to \mathbb{R}$ is the sequence $\sigma_d : \mathbb{N}^d \to \mathbb{R}$, where $\sigma_d(n_1, \dots, n_d) = \prod_{j=1}^d \sigma^{(j)}(n_j)$. It can be rearranged to a nonincreasing zero sequence τ. An example of such a sequence is given by the L_2-approximation numbers of Sobolev functions on the d-torus with mixed order $(s_1, \dots, s_d) \in \mathbb{R}_+^d$. They are generated by the L_2-approximation numbers of the univariate Sobolev spaces $H^{s_j}(\mathbb{T})$, which are of order n^{-s_j}. It is known that τ has the order $n^{-s} (\log n)^{s(l-1)}$ in this case, where s is the minimum among all numbers s_j and l is its multiplicity. This was proven by Mityagin [M62] for integer vectors $(s_1, \dots, s_d)$ and by Nikol'skaya [N74] in the general case. See [T86, pp. 32, 36, 72] and [DTU16] for more details. It is not hard to deduce that the order of decay of τ is at least (at most) $n^{-s} (\log n)^{s(l-1)}$, whenever the order of the factor sequences $\sigma^{(j)}$ is at least (at most) n^{-s_j}. But in contrast to the tensor power case, asymptotic constants of tensor product sequences in general are not determined by the asymptotic constants of the factor sequences.

Example 2. Consider the sequences $\sigma, \mu, \tilde{\mu} : \mathbb{N} \to \mathbb{R}$ with

$$\sigma(n) = n^{-1}, \quad \mu(n) = n^{-2}, \quad \tilde{\mu}(n) = \begin{cases} 1, & \text{for } n \leq N, \\ n^{-2}, & \text{for } n > N, \end{cases} \tag{2.39}$$

for some $N \in \mathbb{N}$. The tensor product $\sigma_2 : \mathbb{N}^2 \to \mathbb{R}$ of σ and μ has the form

$$\sigma_2(n_1, n_2) = n_1^{-1} n_2^{-2} \tag{2.40}$$

and its nonincreasing rearrangement τ satisfies for all $n \in \mathbb{N}$ that

$$\begin{aligned} n &\leq \# \left\{ (n_1, n_2) \in \mathbb{N}^2 \mid \sigma_2(n_1, n_2) \geq \tau(n) \right\} = \# \left\{ (n_1, n_2) \mid n_1 n_2^2 \leq \tau(n)^{-1} \right\} \\ &\leq \sum_{n_2=1}^{\infty} \# \left\{ n_1 \in \mathbb{N} \mid n_1 \leq a_n^{-1} n_2^{-2} \right\} \leq \tau(n)^{-1} \sum_{n_2=1}^{\infty} n_2^{-2} \leq 2\tau(n)^{-1}, \end{aligned} \tag{2.41}$$

and hence

$$\limsup_{n \to \infty} \tau(n) n \leq 2. \tag{2.42}$$

The tensor product $\tilde{\sigma}_2 : \mathbb{N}^2 \to \mathbb{R}$ of σ and $\tilde{\mu}$ takes the form

$$\tilde{\sigma}_2(n_1, n_2) = \begin{cases} n_1^{-1}, & \text{if } n_2 \leq N, \\ n_1^{-1} n_2^{-2}, & \text{else}, \end{cases} \tag{2.43}$$

and its nonincreasing rearrangement $\tilde{\tau}$ satisfies for all $n \in \mathbb{N}$ that

$$\begin{aligned} n &\geq \# \left\{ (n_1, n_2) \in \mathbb{N}^2 \mid \tilde{\sigma}_2(n_1, n_2) > \tilde{a}_n \right\} \geq N \# \left\{ n_1 \in \mathbb{N} \mid n_1^{-1} > \tilde{\tau}(n) \right\} \\ &\geq N \left(\tilde{\tau}(n)^{-1} - 1 \right) \end{aligned} \tag{2.44}$$

and thus

$$\liminf_{n \to \infty} \tilde{\tau}(n) n \geq N. \tag{2.45}$$

Hence, matching asymptotic constants of the factor sequences do not necessarily lead to matching asymptotic constants of the tensor product sequences.

3 Preasymptotic Behavior of Tensor Power Sequences

In order to estimate the size of $\tau(n)$ for small values of n, we give explicit estimates for $A_2(r, l)$ from (2.4) for $l \leq d$ and small values of r. The right asymptotic behavior of these estimates, however, is less important. Note that $A_2(r, l) = 0$ for $r < 2^l$.

Lemma 2. *Let $r \geq 0$ and $l \in \mathbb{N}$. For any $\delta > 0$ we have*

$$\begin{aligned} A_2(r, l) &\leq \frac{r^{1+\delta}}{\delta^{l-1}} && \text{and} \\ A_2(r, l) &\geq \frac{r}{3 \cdot 2^{l-1}} && \text{for } r \geq 2^l. \end{aligned}$$

Proof. Both estimates hold in the case $l = 1$, since

$$A_2(r, 1) = \begin{cases} 0, & \text{for } r < 2, \\ \lfloor r \rfloor - 1, & \text{for } r \geq 2. \end{cases} \tag{3.1}$$

If they hold for some $l \in \mathbb{N}$, then

$$\begin{aligned} A_2(r, l+1) = \sum_{k=2}^{\infty} A_2 \left(\frac{r}{k}, l \right) &\leq \frac{r^{1+\delta}}{\delta^{l-1}} \sum_{k=2}^{\infty} \frac{1}{k^{1+\delta}} \\ &\leq \frac{r^{1+\delta}}{\delta^{l-1}} \int_1^{\infty} \frac{1}{x^{1+\delta}} \, \mathrm{d}x = \frac{r^{1+\delta}}{\delta^l} \end{aligned} \tag{3.2}$$

and for $r \geq 2^{l+1}$

$$A_2(r, l+1) \geq A_2 \left(\frac{r}{2}, l \right) \geq \frac{r/2}{3 \cdot 2^{l-1}} = \frac{r}{3 \cdot 2^l}. \tag{3.3}$$

We have thus proven Lemma 2 by induction. $\square$

Theorem 4. *Let σ be a nonincreasing zero sequence with $1 = \sigma(1) > \sigma(2) > 0$ and let τ be the nonincreasing rearrangement of its dth tensor power.*

(i) Suppose that $\sigma(n) \leq C\, n^{-s}$ for some $s, C > 0$ and all $n \geq 2$ and let $\delta \in (0,1]$. For any $n \in \mathbb{N}$,

$$\tau(n) \leq \left(\frac{\tilde{C}(\delta)}{n}\right)^{\alpha(d,\delta)}, \quad \textit{where}$$

$$\tilde{C}(\delta) = \exp\left(\frac{C^{(1+\delta)/s}}{\delta}\right) \quad \textit{and} \quad \alpha(d,\delta) = \frac{\log \sigma(2)^{-1}}{\log\left(\sigma(2)^{-(1+\delta)/s} \cdot d\right)} > 0.$$

(ii) Let $v = \#\{n \geq 2 \mid \sigma(n) = \sigma(2)\}$. For any $n \in \left\{2, \ldots, (1+v)^d\right\}$,

$$\tau(n) \geq \sigma(2) \cdot \left(\frac{1}{n}\right)^{\beta(d,n)}, \quad \textit{where} \quad \beta(d,n) = \frac{\log \sigma(2)^{-1}}{\log\left(1 + \frac{v}{\log_{1+v} n} \cdot d\right)} > 0.$$

The assumption $\sigma(1) = 1$ merely reduces the complexity of the estimates. We can easily translate the above estimates for arbitrary $\sigma(1) > \sigma(2) > 0$ by applying Theorem 4 to the sequence $(\sigma(n)/\sigma(1))_{n\in\mathbb{N}}$. We simply have to replace $\sigma(2)$ by $\sigma(2)/\sigma(1)$, C by $C/\sigma(1)$ and $\tau(n)$ by $\tau(n)/\sigma(1)^d$. Theorem 3, as stated in the introduction, is an immediate consequence of Theorem 4. Obviously, $\sigma(2) = \sigma(1)$ implies $\tau(n) = \sigma(1)^d$ for every $n \leq (1+v)^d$, whereas $\sigma(2) = 0$ implies $\tau(n) = 0$ for every $n \geq 2$.

Proof. Part (i): Let $n \in \mathbb{N}$. There is some $L \geq 0$ with $\tau(n) = \sigma(2)^L$. If $\sigma_d(\boldsymbol{n}) \geq \tau(n)$, the number l of components of $\boldsymbol{n}$ not equal to one is at most $\lfloor L \rfloor$ and hence

$$\begin{aligned} n &\leq \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \sigma_d(\boldsymbol{n}) \geq \tau(n)\right\} \\ &= \sum_{l=0}^{\min\{\lfloor L \rfloor, d\}} \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \#\{1 \leq j \leq d \mid n_j \neq 1\} = l \wedge \sigma_d(\boldsymbol{n}) \geq \tau(n)\right\} \\ &= 1 + \sum_{l=1}^{\min\{\lfloor L \rfloor, d\}} \binom{d}{l} \#\left\{\boldsymbol{n} \in \{2,3,\ldots\}^l \mid \sigma_d(\boldsymbol{n}) \geq \tau(n)\right\}. \end{aligned} \tag{3.4}$$

Since $\sigma_d(\boldsymbol{n}) \leq C^l \prod_{j=1}^l n_j^{-s}$ for $\boldsymbol{n} \in \{2,3,\ldots\}^l$, Lemma 2 yields for $l \leq \min\{\lfloor L \rfloor, d\}$,

$$\begin{aligned} \#\left\{\boldsymbol{n} \in \{2,3,\ldots\}^l \mid \sigma_d(\boldsymbol{n}) \geq \tau(n)\right\} &\leq A_2\left(C^{l/s}\tau(n)^{-1/s}, l\right) \\ &\leq C^{(1+\delta)l/s}\tau(n)^{-(1+\delta)/s}\delta^{-l} \end{aligned} \tag{3.5}$$

Obviously,

$$1 \le \binom{d}{0} \cdot C^{0/s} \tau(n)^{-(1+\delta)/s} \delta^0. \tag{3.6}$$

Inserting these bounds in (3.4) yields

$$\begin{aligned} n &\le \sum_{l=0}^{\min\{\lfloor L \rfloor, d\}} \binom{d}{l} \cdot C^{(1+\delta)l/s} \tau(n)^{-(1+\delta)/s} \delta^{-l} \le \tau(n)^{-(1+\delta)/s} \sum_{l=0}^{\min\{\lfloor L \rfloor, d\}} \frac{d^l}{l!} C^{(1+\delta)l/s} \delta^{-l} \\ &\le \sigma(2)^{-(1+\delta)L/s} d^L \sum_{l=0}^{\min\{\lfloor L \rfloor, d\}} \frac{\left(\frac{C^{(1+\delta)/s}}{\delta}\right)^l}{l!} \le \left(\sigma(2)^{-(1+\delta)/s} \cdot d\right)^L \exp\left(\frac{C^{(1+\delta)/s}}{\delta}\right) \end{aligned} \tag{3.7}$$

and hence

$$L \ge \frac{\log n - \frac{C^{(1+\delta)/s}}{\delta}}{\log\left(\sigma(2)^{-(1+\delta)/s} \cdot d\right)}. \tag{3.8}$$

Thus

$$\tau(n) = \sigma(2)^L \le \exp\left(\frac{\left(\frac{C^{(1+\delta)/s}}{\delta} - \log n\right) \log \sigma(2)^{-1}}{\log\left(\sigma(2)^{-(1+\delta)/s} \cdot d\right)}\right) = \left(\frac{\exp\left(\frac{C^{(1+\delta)/s}}{\delta}\right)}{n}\right)^{\alpha(d,\delta)} \tag{3.9}$$

with

$$\alpha(d, \delta) = \frac{\log \sigma(2)^{-1}}{\log\left(\sigma(2)^{-(1+\delta)/s} \cdot d\right)}. \tag{3.10}$$

Part (ii): Let $n \in \{2, \ldots, (1+v)^d\}$. Then $\sigma(2)^d \le \tau(n) \le \sigma(2)$. If $\tau(n)$ equals $\sigma(2)$, the lower bound is trivial. Else, there is some $L \in \{1, \ldots, d-1\}$ such that $\tau(n) \in [\sigma(2)^{L+1}, \sigma(2)^L)$. Clearly,

$$n > \#\left\{\boldsymbol{n} \in \mathbb{N}^d \mid \sigma_d(\boldsymbol{n}) > \tau(n)\right\} \ge \sum_{l=1}^{L} \binom{d}{l} \#\left\{\boldsymbol{n} \in \{2, 3, \ldots\}^l \mid \sigma_d(\boldsymbol{n}) > \tau(n)\right\}. \tag{3.11}$$

If $l \le L$, we have $\sigma_d(\boldsymbol{n}) > \tau(n)$ for every $\boldsymbol{n} \in \{2, \ldots, 1+v\}^l$ and hence

$$n \ge \sum_{l=0}^{L} \binom{d}{l} v^l \ge \sum_{l=0}^{L} \binom{L}{l} \left(\frac{d}{L}\right)^l v^l = \left(1 + \frac{vd}{L}\right)^L. \tag{3.12}$$

Since d/L is bigger than one, this yields in particular that

$$L \le \log_{1+v} n. \tag{3.13}$$

We insert this auxiliary estimate on L in (3.12) and get

$$n \geq \left(1 + \frac{vd}{\log_{1+v} n}\right)^L, \tag{3.14}$$

or equivalently

$$L \leq \frac{\log n}{\log\left(1 + \frac{vd}{\log_{1+v} n}\right)}. \tag{3.15}$$

We recall that $\tau(n) \geq \sigma(2)^{L+1}$ and realize that the proof is finished. $\square$

The bounds of Theorem 4 are very explicit, but complex. One might be bothered by the dependence of the exponent in the lower bound on n. This can be overcome, if we restrict the lower bound to the case $n \leq (1+v)^{d^a}$ for some $0 < a < 1$ and replace $\beta(d,n)$ by

$$\tilde{\beta}(d) = \frac{\log \sigma(2)^{-1}}{\log\left(1 + v \cdot d^{1-a}\right)}. \tag{3.16}$$

Of course, we throw away information this way. Similarly, we get a worse but still valid estimate, if we replace v by one. Note that these lower bounds are valid for any zero sequence σ, independent of its rate of convergence.

The constants 1, $\sigma(2)$ and $\tilde{C}(\delta)$ are independent of the power d. The additional parameter δ in the upper bound was introduced to maximize the exponent $\alpha(d,\delta)$. If δ tends to zero, $\alpha(d,\delta)$ gets bigger, but also the constant $\tilde{C}(\delta)$ explodes.

For large values of d and if n is significantly smaller than $(1+v)^d$, the exponents in both the upper and the lower bound are close to $t_d = \frac{\log(\sigma(2)/\sigma(1))^{-1}}{\log d}$. In other words, the sequence τ preasymptotically roughly decays like n^{-t_d}.

These kinds of estimates are also closely related to those in [GW11, Section 3]. Using the language of generalized tractability, Gnewuch and Woźniakowski show that the supremum of all $p > 0$ such that there is a constant $C > 0$ with

$$\tau(n) \leq e \cdot (C/n)^{\frac{p}{1+\log d}} \tag{3.17}$$

for all $n \in \mathbb{N}$ and $d \in \mathbb{N}$ is $\min\{s, \log \sigma(2)^{-1}\}$.

4 Applications to some Tensor Power Operators

Let X and Y be Hilbert spaces and let $T : X \to Y$ be a compact linear operator. The nth approximation number of T is the quantity

$$a_n(T) = \inf_{\operatorname{rank}(A)<n} \|T - A\|. \tag{4.1}$$

It measures the power of approximating T in $\mathcal{L}(X,Y)$ by operators of rank less than n. Obviously, the first approximation number of T coincides with its norm. Since $W = T^*T \in \mathcal{L}(X)$ is positive semi-definite and compact, it admits a finite or countable orthonormal basis $\mathcal{B}$ of $N(T)^\perp$ consisting of eigenvectors $b \in \mathcal{B}$ to eigenvalues

$$\lambda(b) = \langle Wb, b\rangle_X = \|Tb\|_Y^2 > 0. \tag{4.2}$$

I will refer to $\mathcal{B}$ as the orthonormal basis associated with T. It can be characterized as the orthonormal basis of $N(T)^\perp$ whose image is an orthogonal basis of $\overline{R(T)}$. It is unique up to the choice of orthonormal bases in the finite-dimensional eigenspaces of W. Clearly,

$$Tf = \sum_{b\in\mathcal{B}} \langle f, b\rangle_X Tb \quad \text{for } f \in X. \tag{4.3}$$

The square-roots of the eigenvalues of W are called singular values of T. Let $\sigma(n)$ be the nth largest singular value of T, provided $n \leq |\mathcal{B}|$. Else, let $\sigma(n) = 0$. The algorithm

$$A_n f = \sum_{b\in\mathcal{B}_n} \langle f, b\rangle_X Tb \quad \text{for } f \in X \tag{4.4}$$

is an optimal approximation of T by operators of rank less than n, if $\mathcal{B}_n$ consists of all $b \in \mathcal{B}$ with $\|Tb\|_Y > \sigma(n)$. In particular, $a_n(T)$ and $\sigma(n)$ coincide and

$$a_n(T) = \min_{\substack{V\subseteq X \\ \dim(V)\leq n-1}} \max_{\substack{f\perp V \\ \|f\|_X=1}} \|Tf\|_Y. \tag{4.5}$$

We are concerned with the approximation numbers of tensor power operators, defined as follows. Let G be a set and G^d be its d-fold Cartesian product and let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The tensor product of $\mathbb{K}$-valued functions $f_1, \ldots, f_d$ on G is the function

$$f_1 \otimes \ldots \otimes f_d : \quad G^d \to \mathbb{K}, \quad x \mapsto f_1(x_1) \cdot \ldots \cdot f_d(x_d). \tag{4.6}$$

If X is a Hilbert space of $\mathbb{K}$-valued functions on G, its dth tensor power X^d is the smallest Hilbert space of $\mathbb{K}$-valued functions on G^d that contains any tensor product of functions in X and satisfies

$$\langle f_1 \otimes \ldots \otimes f_d, g_1 \otimes \ldots \otimes g_d \rangle = \langle f_1, g_1 \rangle \cdot \ldots \cdot \langle f_d, g_d \rangle \tag{4.7}$$

for any choice of functions $f_1, \ldots, f_d$ and $g_1, \ldots, g_d$ in X. Let Y be another Hilbert space of $\mathbb{K}$-valued functions and let $T \in \mathcal{L}(X, Y)$. The dth tensor power of T is the unique operator $T^d \in \mathcal{L}(X^d, Y^d)$ that satisfies

$$T^d (f_1 \otimes \ldots \otimes f_d) = T f_1 \otimes \ldots \otimes T f_d \tag{4.8}$$

for any choice of functions $f_1, \ldots, f_d$ in X. If T is compact, then so is T^d. Moreover, if $\mathcal{B}$ is the orthonormal basis associated with T, then

$$\mathcal{B}^d = \{b_1 \otimes \ldots \otimes b_d \mid b_1, \ldots, b_d \in \mathcal{B}\} \tag{4.9}$$

is the orthonormal basis associated with T^d. In particular, the singular values of T^d are given as the d-fold products of singular values of T. The sequence of approximation numbers $a_n\left(T^d\right)$ is hence given as the nonincreasing rearrangement of the dth tensor power of the sequence σ of singular values of T.

4.1 Approximation of Mixed Order Sobolev Functions on the Torus

Let $\mathbb{T}$ be the 1-torus, the circle, represented by the interval $[a, b]$, where the two end points $a < b$ are identified. By $L_2(\mathbb{T})$, we denote the Hilbert space of square-integrable functions on $\mathbb{T}$, equipped with the scalar product

$$\langle f, g \rangle = \frac{1}{L} \int_{\mathbb{T}} f(x) \overline{g(x)} \, \mathrm{d}x \tag{4.10}$$

and the induced norm $\|\cdot\|$ for some $L > 0$. Typical normalizations are $[a, b] \in \{[0,1], [-1,1], [0, 2\pi]\}$ and $L \in \{1, b-a\}$. The family $(b_k)_{k \in \mathbb{Z}}$ with

$$b_k(x) = \sqrt{\frac{L}{b-a}} \exp\left(2\pi i k \, \frac{x-a}{b-a}\right) \tag{4.11}$$

is an orthonormal basis of $L_2(\mathbb{T})$, its Fourier basis, and

$$\hat{f}(k) = \langle f, b_k \rangle \tag{4.12}$$

is the kth Fourier coefficient of $f \in L_2(\mathbb{T})$. By Parseval's identity,

$$\|f\|^2 = \sum_{k\in\mathbb{Z}} |\hat{f}(k)|^2 \quad \text{and} \quad \langle f, g\rangle = \sum_{k\in\mathbb{Z}} \hat{f}(k) \cdot \overline{\hat{g}(k)}. \tag{4.13}$$

Let $w = (w_k)_{k\in\mathbb{N}}$ be a nondecreasing sequence of real numbers with $w_0 = 1$ and let $w_{-k} = w_k$ for $k \in \mathbb{N}$ and so let $\tilde{w}$. The univariate Sobolev space $H^w(\mathbb{T})$ is the Hilbert space of functions $f \in L_2(\mathbb{T})$ for which

$$\|f\|_w^2 = \sum_{k\in\mathbb{Z}} w_k^2 \cdot |\hat{f}(k)|^2 \tag{4.14}$$

is finite, equipped with the scalar product

$$\langle f, g\rangle_w = \sum_{k\in\mathbb{Z}} w_k \hat{f}(k) \cdot \overline{w_k \hat{g}(k)}. \tag{4.15}$$

Note that $H^w(\mathbb{T})$ and $H^{\tilde{w}}(\mathbb{T})$ coincide and their norms are equivalent, if and only if $w \sim \tilde{w}$. In case $w_k \sim k^s$ for some $s \geq 0$, the space $H^w(\mathbb{T})$ is the classical Sobolev space of periodic univariate functions with fractional smoothness s, also denoted by $H^s(\mathbb{T})$. In particular, $H^w(\mathbb{T}) = L_2(\mathbb{T})$ for $w \equiv 1$.

In accordance with previous notation, let $X = H^w(\mathbb{T})$ and $Y = H^{\tilde{w}}(\mathbb{T})$. The embedding T of X into Y is compact, if and only if $w_k/\tilde{w}_k$ tends to infinity as k tends to infinity. The Fourier basis $(b_k)_{k\in\mathbb{Z}}$ is an orthogonal basis of X consisting of eigenfunctions of $W = T^*T$ with corresponding eigenvalues

$$\lambda(b_k) = \frac{\|b_k\|_Y^2}{\|b_k\|_X^2} = \frac{\tilde{w}_k^2}{w_k^2}. \tag{4.16}$$

The nth approximation number $\sigma(n)$ of this embedding is the square root of the nth biggest eigenvalue. Hence, replacing the Fourier weight sequences w and $\tilde{w}$ by equivalent sequences does not affect the order of convergence of the corresponding approximation numbers, but it may drastically affect their asymptotic constants and preasymptotic behavior. If $Y = L_2(\mathbb{T})$, we obtain

$$\sigma(n) = w_{k_n}^{-1}, \quad \text{where} \quad k_n = (-1)^n \lfloor n/2 \rfloor. \tag{4.17}$$

Note that $\sigma(1)$, the norm of the embedding T, is always one.

The dth tensor power $X^d = H^w_{\text{mix}}(\mathbb{T}^d)$ of X is a space of mixed order Sobolev functions on the d-torus. If $w_k \sim k^s$ for some $s \geq 0$, this is the space $H^s_{\text{mix}}(\mathbb{T}^d)$ of functions with dominating mixed smoothness s. If even $s \in \mathbb{N}_0$, this space consists of all real-valued functions on the d-torus, which have a weak (or distributional) derivative of order α in $L_2\left(\mathbb{T}^d\right)$ for any $\alpha \in \{0,1,\ldots,s\}^d$. Of course, the same holds for the dth tensor power $Y^d = H^{\tilde{w}}_{\text{mix}}\left(\mathbb{T}^d\right)$ of Y. The tensor power operator $T^d : X^d \to Y^d$ is the compact embedding of $H^{w}_{\text{mix}}\left(\mathbb{T}^d\right)$ into $H^{\tilde{w}}_{\text{mix}}\left(\mathbb{T}^d\right)$. Hence, the approximation numbers of this embedding are the nonincreasing rearrangement of the dth tensor power of σ.

If $(\tilde{w}_k/w_k)_{k\in\mathbb{N}}$ is of polynomial decay, Theorem 2 and Theorem 4 apply. We formulate the results for the embedding of $H^{s}_{\text{mix}}\left(\mathbb{T}^d\right)$ into $L_2\left(\mathbb{T}^d\right)$, where $H^{s}_{\text{mix}}\left(\mathbb{T}^d\right)$ will be equipped with different equivalent norms, indicated by the notation

$$
\begin{aligned}
&H^{s,\circ,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right), \quad \text{if} \quad w_k = \left(\sum_{l=0}^{s} \left|\gamma^{-1}\frac{2\pi k}{b-a}\right|^{2l}\right)^{1/2}, \\
&H^{s,*,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right), \quad \text{if} \quad w_k = \left(1 + \left|\gamma^{-1}\frac{2\pi k}{b-a}\right|^{2s}\right)^{1/2}, \\
&H^{s,+,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right), \quad \text{if} \quad w_k = \left(1 + \left|\gamma^{-1}\frac{2\pi k}{b-a}\right|^{2}\right)^{s/2}, \\
&H^{s,\#,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right), \quad \text{if} \quad w_k = \left(1 + \left|\gamma^{-1}\frac{2\pi k}{b-a}\right|\right)^{s},
\end{aligned}
\tag{4.18}
$$

for some $\gamma > 0$. The last three norms are due to Kühn, Sickel and Ullrich [KSU15], who study all these norms for $\gamma = 1$, $L = 1$ and $[a,b] = [0,2\pi]$. The last norm is also studied by Chernov and Dũng in [CD16] for $L = 2\pi$, $[a,b] = [-\pi,\pi]$ and arbitrary values of γ. If s is a natural number, the first two scalar products take the form

$$
\begin{aligned}
\langle f, g\rangle_{H^{s,\circ,\gamma}_{\text{mix}}} &= \sum_{\alpha\in\{0,\ldots,s\}^d} \gamma^{-2s|\alpha|} \left\langle D^\alpha f, D^\alpha g\right\rangle, \\
\langle f, g\rangle_{H^{s,*,\gamma}_{\text{mix}}} &= \sum_{\alpha\in\{0,s\}^d} \gamma^{-2s|\alpha|} \left\langle D^\alpha f, D^\alpha g\right\rangle.
\end{aligned}
\tag{4.19}
$$

This is why $H^{s,\circ,1}_{\text{mix}}\left(\mathbb{T}^d\right)$ and $H^{s,*,1}_{\text{mix}}\left(\mathbb{T}^d\right)$ might be considered the most natural choice. Note that the corresponding approximation numbers of the embedding T^d are independent of the normalization constant L, but they do depend on the length of the interval $[a,b]$.

Corollary 1. *The following limits exist and coincide:*

$$\left.\begin{array}{l}\lim\limits_{n\to\infty} a_n\left(H^{s,\circ,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right)\hookrightarrow L_2\left(\mathbb{T}^d\right)\right)\cdot n^s\left(\log n\right)^{-s(d-1)}\\ \lim\limits_{n\to\infty} a_n\left(H^{s,*,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right)\hookrightarrow L_2\left(\mathbb{T}^d\right)\right)\cdot n^s\left(\log n\right)^{-s(d-1)}\\ \lim\limits_{n\to\infty} a_n\left(H^{s,+,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right)\hookrightarrow L_2\left(\mathbb{T}^d\right)\right)\cdot n^s\left(\log n\right)^{-s(d-1)}\\ \lim\limits_{n\to\infty} a_n\left(H^{s,\#,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right)\hookrightarrow L_2\left(\mathbb{T}^d\right)\right)\cdot n^s\left(\log n\right)^{-s(d-1)}\end{array}\right\}=\left(\frac{\left(\gamma\,\frac{b-a}{\pi}\right)^d}{(d-1)!}\right)^s.$$

Of course, this coincides with the limits computed in [KSU15], if $\gamma^{-1}\frac{b-a}{\pi}=2$. The third limit (for $[a,b]=[-\pi,\pi]$, $L=2\pi$ and any $\gamma>0$) may not be written down explicitly in [CD16], but can be derived from their Theorem 4.6.

Corollary 2. *Let $\square\in\{\circ,*,+,\#\}$. For any $s>0$, $d\in\mathbb{N}$ and $n\in\left\{2,\dots,3^d\right\}$,*

$$\sigma_\square(2)\left(\frac{1}{n}\right)^{\beta_\square(d,n)}\leq a_n\left(H^{s,\square,\gamma}_{\text{mix}}\left(\mathbb{T}^d\right)\hookrightarrow L_2\left(\mathbb{T}^d\right)\right)\leq\left(\frac{\tilde{C}(\delta)}{n}\right)^{\alpha_\square(d,\delta)}.$$

The parameter $\delta\in(0,1]$ is arbitrary, $\tilde{C}(\delta)=\exp\left(\left(3/\eta\right)^{1+\delta}/\delta\right)$ for $\eta=\frac{2\pi}{\gamma(b-a)}$ and the values $\sigma_\square$, $\alpha_\square$ and $\beta_\square$ are listed below. The upper bound holds for all $n\in\mathbb{N}$.

$\square$	$\sigma_\square(2)$	$\alpha_\square(d,\delta)$	$\beta_\square(d,n)$
$\circ$	$\left(\sum_{l=0}^s\eta^{2l}\right)^{-\frac{1}{2}}$	$\frac{\frac{1}{2}\log\left(\sum_{l=0}^s\eta^{2l}\right)}{\log d+\frac{1+\delta}{2s}\cdot\log\left(\sum_{l=0}^s\eta^{2l}\right)}$	$\frac{\frac{1}{2}\log\left(\sum_{l=0}^s\eta^{2l}\right)}{\log\left(1+\frac{2}{\log_3 n}d\right)}$
$*$	$\left(1+\eta^{2s}\right)^{-\frac{1}{2}}$	$\frac{\frac{1}{2}\log\left(1+\eta^{2s}\right)}{\log d+\frac{1+\delta}{2s}\cdot\log(1+\eta^{2s})}$	$\frac{\frac{1}{2}\log\left(1+\eta^{2s}\right)}{\log\left(1+\frac{2}{\log_3 n}d\right)}$
$+$	$\left(1+\eta^2\right)^{-\frac{s}{2}}$	$\frac{\frac{s}{2}\log\left(1+\eta^2\right)}{\log d+\frac{1+\delta}{2}\cdot\log(1+\eta^2)}$	$\frac{\frac{s}{2}\log\left(1+\eta^2\right)}{\log\left(1+\frac{2}{\log_3 n}d\right)}$
$\#$	$(1+\eta)^{-s}$	$\frac{s\log(1+\eta)}{\log d+(1+\delta)\log(1+\eta)}$	$\frac{s\log(1+\eta)}{\log\left(1+\frac{2}{\log_3 n}d\right)}$

Let us consider the setting of [KSU15], where $\gamma=1$ and $b-a=2\pi$ and hence η is one. The exponents $\alpha_\#(d,\delta)=\frac{s}{\log_2 d+1+\delta}$ and $\alpha_+(d,\delta)=\frac{s}{2\log_2 d+1+\delta}$ in our upper bounds are slightly better than the exponents $\frac{s}{\log_2 d+2}$ and $\frac{s}{2\log_2 d+4}$ in Theorem 4.9, 4.10 and Theorem 4.17 of [KSU15], but almost the same. Also the lower bounds basically coincide. Regarding $H^{s,*,1}_{\text{mix}}\left(\mathbb{T}^d\right)$, Kühn, Sickel and Ullrich only studied the case $1/2\leq s\leq 1$ in Theorem 4.20. As we see now, there is a major difference between this natural norm and the last two norms: For large dimensions d, the preasymptotic behavior of the approximation numbers is roughly $n^{-t_{d,\square}}$, where

$$t_{d,\circ}=\frac{\log\left(s+1\right)}{2\log d},\quad t_{d,*}=\frac{1}{2\log_2 d},\quad t_{d,+}=\frac{s}{2\log_2 d},\quad t_{d,\#}=\frac{s}{\log_2 d}. \tag{4.20}$$

This means that the smoothness of the space only has a minor or even no impact on the preasymptotic decay of the approximation numbers, if $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ is equipped with one of the natural norms $\|\cdot\|_{H^{s,\circ,1}_{\text{mix}}}$ or $\|\cdot\|_{H^{s,*,1}_{\text{mix}}}$.

This changes, however, if the value of $\eta = \frac{2\pi}{\gamma(b-a)}$ changes. If η is larger than one, because we consider a shorter interval $[a,b]$ or because we put some weight $\gamma < \frac{2\pi}{b-a}$, also the exponents $t_{d,\circ}$ and $t_{d,*}$ get linear in s. For the other two families of norms, the smoothness does show and the value of η is less important.

There are no preasymptotic estimates in [CD16].

4.2 Approximation of Mixed Order Jacobi Functions on the Cube

The above results also apply to the approximation numbers of the embedding of mixed order Jacobi functions on the d-cube in the corresponding L_2-space as considered in [CD16, Section 5].

Let $\mathbb{I}$ be the 1-cube, a line segment, represented by $[-1,1]$. For fixed parameters $\alpha, \beta > -1$ with $a := \frac{\alpha+\beta+1}{2} > 0$, the weighted L_2-space $Y = L_2\left(\mathbb{I}, w\right)$ is the Hilbert space of measurable, real-valued functions on $\mathbb{I}$ with

$$\int_{\mathbb{I}} f(x)^2 w(x)\, \mathrm{d}x < \infty, \tag{4.21}$$

equipped with the scalar product

$$\langle f, g\rangle = \int_{\mathbb{I}} f(x)g(x)w(x)\, \mathrm{d}x \tag{4.22}$$

and the induced norm $\|\cdot\|$, where $w : \mathbb{I} \to \mathbb{R}$ is the Jacobi weight

$$w(x) = (1-x)^\alpha (1+x)^\beta. \tag{4.23}$$

This reduces to the classical space of square-integrable functions, if both parameters are zero. As α respectively β increases, the space grows, since we allow for stronger singularities on the right respectively left endpoint, and vice versa.

The family of Jacobi polynomials $(P_k)_{k\in\mathbb{N}_0}$ is an orthogonal basis of Y. These polynomials can be defined as the unique solutions of the differential equations

$$\mathcal{L}P_k = k(k+2a)P_k \tag{4.24}$$

for the second order differential operator

$$\mathcal{L} = -w(x)^{-1} \frac{d}{dx} \left(\left(1 - x^2\right) w(x) \frac{d}{dx} \right) \tag{4.25}$$

that satisfy

$$P_k(1) = \binom{k+\alpha}{k} \quad \text{and} \quad P_k(-1) = (-1)^k \binom{k+\beta}{k}. \tag{4.26}$$

We denote the kth Fourier coefficient of f with respect to the normalized Jacobi basis by f_k. The scalar product in Y hence admits the representation

$$\langle f, g \rangle = \sum_{k=0}^{\infty} f_k g_k. \tag{4.27}$$

For $s > 0$ let $X = K^s(\mathbb{I}, w)$ be the Hilbert space of functions $f \in Y$ with

$$\sum_{k=0}^{\infty} \left(1 + a^{-1} k\right)^{2s} f_k^2 < \infty, \tag{4.28}$$

equipped with the scalar product

$$\langle f, g \rangle_s = \sum_{k=0}^{\infty} \left(1 + a^{-1} k\right)^{2s} f_k g_k \tag{4.29}$$

and the induced norm $\|\cdot\|_s$. Obviously, $(P_k)_{k \in \mathbb{N}_0}$ is an orthogonal basis of X, too. In case s is an even integer, this is the space of all functions $f \in L_2(\mathbb{I}, w)$ such that $\mathcal{L}^j f \in L_2(\mathbb{I}, w)$ for $j = 1 \ldots \frac{s}{2}$ and the scalar product

$$\langle f, g \rangle_{s,*} = \sum_{j=0}^{s/2} \left\langle \mathcal{L}^j f, \mathcal{L}^j g \right\rangle \tag{4.30}$$

is equivalent to the one above. The parameter s can hence be interpreted as smoothness of the functions in $K^s(\mathbb{I}, w)$. The embedding T of X into Y is compact and its nth approximation number is given by

$$\sigma(n) = a_n(T) = \frac{\|P_{n-1}\|}{\|P_{n-1}\|_s} = \left(1 + a^{-1}(n-1)\right)^{-s}. \tag{4.31}$$

We can apply our theorems to study the approximation numbers of the dth tensor power T^d of T. This is the embedding of $X^d = K^s\left(\mathbb{I}^d, w_d\right)$ into $Y^d = L_2\left(\mathbb{I}^d, w_d\right)$, where Y^d is the weighted L_2-space on the d-cube with respect to the Jacobi weight $w_d = w \otimes \ldots \otimes w$ and X^d is the subspace of Jacobi functions of mixed order s. Like in the univariate case, X^d can be described via differentials of dominating mixed order s and less, if s is an even integer.

Corollary 3. *For any $d \in \mathbb{N}$ and $s > 0$, the following limit exists:*

$$\lim_{n\to\infty} a_n\left(K^s\left(\mathbb{I}^d, w_d\right) \hookrightarrow L_2\left(\mathbb{I}^d, w_d\right)\right) \cdot n^s \left(\log n\right)^{-s(d-1)} = \left(\frac{a^d}{(d-1)!}\right)^s.$$

This result could also be derived from Theorem 5.5 in [CD16]. In addition, we get the following preasymptotic estimates:

Corollary 4. *For any $\delta \in (0,1]$, $s > 0$, $d \in \mathbb{N}$ and $n \in \left\{2, \ldots, 2^d\right\}$,*

$$\left(\frac{a}{a+1}\right)^s \left(\frac{1}{n}\right)^{p_{s,a,d,n}} \leq a_n\left(K^s\left(\mathbb{I}^d, w_d\right) \hookrightarrow L_2\left(\mathbb{I}^d, w_d\right)\right) \leq \left(\frac{\exp\left(\frac{(2a)^{1+\delta}}{\delta}\right)}{n}\right)^{q_{s,a,d,\delta}}$$

with $\quad p_{s,a,d,n} = \dfrac{s \log \frac{a+1}{a}}{\log\left(1 + \frac{d}{\log_2 n}\right)} \quad$ *and* $\quad q_{s,a,d,\delta} = \dfrac{s \log \frac{a+1}{a}}{\log d + (1+\delta)\log \frac{a+1}{a}}.$

The upper bound even holds for all $n \in \mathbb{N}$.

This means that for large dimension d, a preasymptotic decay of approximate order $t_d = s \log \frac{a+1}{a} / \log d$ in n can be observed.

4.3 Approximation of Mixed Order Sobolev Functions on the Cube

Another example of a tensor power operator is given by the L_2-embedding of mixed order Sobolev functions on the d-cube. Let $\mathbb{I}$ be the 1-cube and $\mathbb{T}$ be the 1-torus. Both shall be represented by the interval $[a, b]$, where a and b are identified in the second case. For any $s \in \mathbb{N}_0$, the vector space

$$H^s\left(\mathbb{I}\right) = \left\{f \in L_2\left(\mathbb{I}\right) \mid f^{(l)} \in L_2\left(\mathbb{I}\right) \text{ for } 1 \leq l \leq s\right\}, \tag{4.32}$$

equipped with the scalar product

$$\langle f, g\rangle_s = \sum_{l=0}^{s} \int_a^b f^{(l)}(x) \cdot \overline{g^{(l)}(x)}\ \mathrm{d}x \tag{4.33}$$

and induced norm $\|\cdot\|_s$, is a Hilbert space, the Sobolev space of order s on $\mathbb{I}$. In case $s = 0$, it coincides with $L_2(\mathbb{I})$. The subset

$$H^s(\mathbb{T}) = \left\{ f \in H^s(\mathbb{I}) \mid f^{(l)}(a) = f^{(l)}(b) \ \text{ for } \ l = 0, 1, \ldots, s-1 \right\} \tag{4.34}$$

of periodic functions is a closed subspace with codimension s, the Sobolev space of order s on $\mathbb{T}$. By means of Parseval's identity and integration by parts, the above norm can be rearranged to

$$\|f\|_s^2 = \sum_{k \in \mathbb{Z}} \left| \hat{f}(k) \right|^2 \sum_{l=0}^{s} \left| \frac{2\pi k}{b-a} \right|^{2l} \quad \text{ for } f \in H^s(\mathbb{T}), \tag{4.35}$$

where

$$\hat{f}(k) = \sqrt{\frac{1}{b-a}} \int_a^b f(x) \cdot \exp\left(-2\pi i k \, \frac{x-a}{b-a} \right) \mathrm{d}x \tag{4.36}$$

is the kth Fourier coefficient of f. In the limiting case $s = \infty$, the Sobolev space $H^\infty(\mathbb{I})$ shall be defined as the Hilbert space

$$H^\infty(\mathbb{I}) = \left\{ f \in \mathcal{C}^\infty(\mathbb{I}) \mid \sum_{l=0}^{\infty} \left\| f^{(l)} \right\|_0^2 < \infty \right\}, \tag{4.37}$$

equipped with the scalar product (4.33) for $s = \infty$. It contains all polynomials and is hence infinite-dimensional. The space $H^\infty(\mathbb{T})$ shall be the closed subspace of periodic functions, i.e.

$$H^\infty(\mathbb{T}) = \left\{ f \in H^\infty(\mathbb{I}) \mid f^{(l)}(a) = f^{(l)}(b) \text{ for any } l \in \mathbb{N}_0 \right\}. \tag{4.38}$$

Note that (4.35) also holds for $s = \infty$. Hence,

$$H^\infty(\mathbb{T}) = \operatorname{span} \left\{ \exp\left(2\pi i k \, \frac{\cdot - a}{b-a} \right) \mid k \in \mathbb{Z} \text{ with } \left| \frac{2\pi k}{b-a} \right| < 1 \right\} \tag{4.39}$$

is finite-dimensional with dimension $2\lceil \frac{b-a}{2\pi} \rceil - 1$. In case $b - a \leq 2\pi$, it consists of constant functions only.

If s is positive, $H^s(\mathbb{I})$ is compactly embedded into $L_2(\mathbb{I})$. Let $\sigma^{(s)}(n)$ be the nth singular value of this embedding and let $\tilde{\sigma}^{(s)}(n)$ be the nth singular value of the embedding of the subspace $H^s(\mathbb{T})$ into $L_2(\mathbb{T})$. We want to study the approximation numbers of the compact embedding of the dth tensor power space

$H^s_{\text{mix}}\left(\mathbb{I}^d\right)$ into $L_2\left(\mathbb{I}^d\right)$. If s is finite, this is the space

$$H^s_{\text{mix}}\left(\mathbb{I}^d\right) = \left\{ f \in L_2\left(\mathbb{I}^d\right) \mid D^\alpha f \in L_2\left(\mathbb{I}^d\right) \text{ for each } \alpha \in \{0, \ldots, s\}^d \right\}, \tag{4.40}$$

equipped with the scalar product

$$\langle f, g \rangle_s = \sum_{\alpha \in \{0,\ldots,s\}^d} \int_{[a,b]^d} D^\alpha f(\boldsymbol{x}) \cdot \overline{D^\alpha g(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x}. \tag{4.41}$$

See Section 4.1 for a treatment of the L_2-approximation numbers of the dth tensor power $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ of the periodic space.

By means of Theorem 2 and Theorem 4, it is enough to study the singular values $\sigma^{(s)}$ of the embedding in the univariate case. As we have seen in Section 4.1,

$$\tilde{\sigma}^{(s)}(n) = \left(\sum_{l=0}^{s} \left| \frac{2\pi \lfloor n/2 \rfloor}{b-a} \right|^{2l} \right)^{-1/2} \quad \text{for } n \in \mathbb{N} \text{ and } s \in \mathbb{N} \tag{4.42}$$

and in particular,

$$\lim_{n\to\infty} \tilde{\sigma}^{(s)}(n)\, n^s = \left(\frac{b-a}{\pi} \right)^s. \tag{4.43}$$

The singular values for nonperiodic functions, on the other hand, are not known explicitly. However, $\sigma^{(s)}$ and $\tilde{\sigma}^{(s)}$ interrelate as follows.

Lemma 3. *For any $n \in \mathbb{N}$ and $s \in \mathbb{N}$, it holds that $\sigma^{(s)}(n+s) \leq \tilde{\sigma}^{(s)}(n) \leq \sigma^{(s)}(n)$.*

Proof. The second inequality is obvious, since $H^s(\mathbb{T})$ is a subspace of $H^s(\mathbb{I})$. The first inequality is true, since the codimension of this subspace is s. Let U be the orthogonal complement of of $H^s(\mathbb{T})$ in $H^s(\mathbb{I})$. By relation (4.5),

$$\begin{aligned} \sigma^{(s)}(n+s) &= \min_{\substack{V \subseteq H^s(\mathbb{I}) \\ \dim(V) \leq n+s-1}} \max_{\substack{f \in H^s(\mathbb{I}), f \perp V \\ \|f\|_s = 1}} \|f\|_0 \leq \min_{\substack{\tilde{V} \subseteq H^s(\mathbb{T}) \\ \dim(\tilde{V}) \leq n-1}} \max_{\substack{f \in H^s(\mathbb{I}), \|f\|_s = 1 \\ f \perp (\tilde{V} \oplus U)}} \|f\|_0 \\ &= \min_{\substack{\tilde{V} \subseteq H^s(\mathbb{T}) \\ \dim(\tilde{V}) \leq n-1}} \max_{\substack{f \in H^s(\mathbb{T}), f \perp \tilde{V} \\ \|f\|_s = 1}} \|f\|_0 = \tilde{\sigma}^{(s)}(n). \end{aligned} \tag{4.44}$$

Note that the same argument is not valid for $d > 1$. In this case, the codimension of $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ in $H^s_{\text{mix}}\left(\mathbb{I}^d\right)$ is not finite. $\square$

Lemma 3 implies that the asymptotic constants of the approximation numbers for the periodic and the nonperiodic functions coincide in the univariate case:

$$\begin{aligned}\lim_{n\to\infty} n^s\tilde{\sigma}^{(s)}(n) &\leq \lim_{n\to\infty} n^s\sigma^{(s)}(n) = \lim_{n\to\infty} (n+s)^s\sigma^{(s)}(n+s)\\ &= \lim_{n\to\infty} n^s\sigma^{(s)}(n+s) \leq \lim_{n\to\infty} n^s\tilde{\sigma}^{(s)}(n).\end{aligned} \tag{4.45}$$

Theorem 2 implies that they also coincide in the multivariate case.

Corollary 5. *For any $d \in \mathbb{N}$ and $s \in \mathbb{N}$, the following limit exists:*

$$\lim_{n\to\infty} a_n\left(H^s_{\text{mix}}\left(\mathbb{I}^d\right) \hookrightarrow L_2\left(\mathbb{I}^d\right)\right) \cdot n^s\left(\log n\right)^{-s(d-1)} = \left(\frac{(b-a)^d}{\pi^d\,(d-1)!}\right)^s.$$

As depicted in Section 3, the approximation numbers show a preasymptotic decay of approximate order $\frac{\log \sigma^{(s)}(2)^{-1}}{\log d}$. Lemma 3 gives no information on $\sigma^{(s)}(2)$. However, relation (4.5) implies that

$$\sigma^{(\infty)}(2) = \max_{f\perp 1,\, f\neq 0} \frac{\|f\|_0}{\|f\|_\infty} \geq \frac{\|2x-a-b\|_0}{\|2x-a-b\|_\infty} = \sqrt{\frac{(b-a)^2}{12+(b-a)^2}}. \tag{4.46}$$

If, for example, the length of the interval $\mathbb{I}$ is one, we obtain

$$\sigma^{(\infty)}(2) \geq 0.27735. \tag{4.47}$$

Since any lower bound on the approximation numbers for $s=\infty$ is a lower bound for $s \in \mathbb{N}$, Theorem 4 yields the following corollary.

Corollary 6. *For any $d \in \mathbb{N}$, any $s \in \mathbb{N}\cup\{\infty\}$ and $d < n \leq 2^d$,*

$$\begin{aligned}&a_n\left(H^s_{\text{mix}}\left([0,1]^d\right) \hookrightarrow L_2\left([0,1]^d\right)\right) \geq\ 0.27 \cdot n^{-c(d,n)},\\ &\textit{where}\qquad c(d,n) = \frac{1.2825}{\log\left(1+\frac{2d}{\log_2 n}\right)} \leq 1.17.\end{aligned}$$

On the other hand, any upper bound on the approximation numbers for $s=1$ is an upper bound for $s \geq 1$. The singular values $\sigma^{(s)}(n)$ for $s=1$ are known. Let T_s be the compact embedding of $H^s(\mathbb{I})$ into $L_2(\mathbb{I})$ and let $W_s = T_s^* T_s$. Then $\sigma^{(s)}(n)$ is the square-root of the nth largest eigenvalue of W_s. It is shown in [T96] that the family $(b_k)_{k\in\mathbb{N}_0}$ is a complete orthogonal system in $H^1(\mathbb{I})$, where the function $b_k : \mathbb{I} \to \mathbb{R}$ with

$$b_k(x) = \cos\left(k\pi \cdot \frac{x-a}{b-a}\right) \quad \text{for } k \in \mathbb{N}_0 \tag{4.48}$$

is an eigenfunction of W_1 with respective eigenvalue

$$\lambda_k = \left(1 + \left(\frac{k\pi}{b-a}\right)^2\right)^{-1}. \tag{4.49}$$

In case $\mathbb{I} = [0,1]$,

$$\sigma^{(1)}(2) = \left(\sqrt{1+\pi^2}\right)^{-1} \leq 0.30332 \tag{4.50}$$

and

$$\sigma^{(1)}(n) \leq 0.607 \cdot n^{-1} \tag{4.51}$$

for $n \geq 2$. Theorem 4 for $\delta = 0.65$ yields the following upper bound.

Corollary 7. *For any $d \in \mathbb{N}$, any $s \in \mathbb{N} \cup \{\infty\}$ and $n \in \mathbb{N}$,*

$$a_n \left(H^s_{\text{mix}}\left([0,1]^d\right) \hookrightarrow L_2\left([0,1]^d\right)\right) \leq \left(\frac{2}{n}\right)^{c(d)} \quad \textit{with} \quad c(d) = \frac{1.1929}{2+\log d}.$$

Apparently, the upper bound for $s = 1$ and the lower bound for $s = \infty$ are already close. The gap between the cases $s = 2$ and $s = \infty$ is even smaller.

Let c be the midpoint of $\mathbb{I}$ and let l be its radius. Moreover, let $\hat{\omega} = \sqrt{1+\omega^2}$ for $\omega \in \mathbb{R}$ and consider the countable sets

$$\begin{aligned} I_1 &= \left\{\omega \geq 0 \mid \hat{\omega}^3 \cosh(\hat{\omega} l) \sin(\omega l) + \omega^3 \sinh(\hat{\omega} l) \cos(\omega l) = 0\right\}, \\ I_2 &= \left\{\omega > 0 \mid \hat{\omega}^3 \sinh(\hat{\omega} l) \cos(\omega l) - \omega^3 \cosh(\hat{\omega} l) \sin(\omega l) = 0\right\}. \end{aligned} \tag{4.52}$$

It can be shown (with some effort) that the family $(b_\omega)_{\omega \in I_1 \cup I_2}$ is a complete orthogonal system in $H^2(\mathbb{I})$, where the function $b_\omega : \mathbb{I} \to \mathbb{R}$ with

$$\begin{aligned} b_\omega(x) &= \omega^2 \cdot \frac{\cosh(\hat{\omega}(x-c))}{\cosh(\hat{\omega} l)} + \hat{\omega}^2 \cdot \frac{\cos(\omega(x-c))}{\cos(\omega l)}, \quad \text{if } \omega \in I_1, \\ b_\omega(x) &= \omega^2 \cdot \frac{\sinh(\hat{\omega}(x-c))}{\sinh(\hat{\omega} l)} + \hat{\omega}^2 \cdot \frac{\sin(\omega(x-c))}{\sin(\omega l)}, \quad \text{if } \omega \in I_2, \end{aligned} \tag{4.53}$$

is an eigenfunction of W_2 with respective eigenvalue

$$\lambda_\omega = \left(1 + \omega^2 + \omega^4\right)^{-1}. \tag{4.54}$$

In particular,

$$\sigma^{(2)}(2) = \left(\sqrt{1+\omega_0^2+\omega_0^4}\right)^{-1}, \tag{4.55}$$

where ω_0 is the smallest nonzero element of $I_1 \cup I_2$. If, for example, the interval $\mathbb{I}$ has unit length, we obtain

$$\sigma^{(2)}(2) \leq 0.27795 \tag{4.56}$$

and like before,

$$\sigma^{(2)}(n) \leq 0.607 \cdot n^{-1} \tag{4.57}$$

for $n \geq 2$. Theorem 4 for $\delta = 0.65$ yields the following upper bound.

Corollary 8. *For any $d \in \mathbb{N}$, any $s \in \mathbb{N} \cup \{\infty\}$ with $s \geq 2$ and $n \in \mathbb{N}$,*

$$a_n \left(H^s_{\text{mix}} \left([0,1]^d\right) \hookrightarrow L_2 \left([0,1]^d\right)\right) \leq \left(\frac{2}{n}\right)^{c(d)} \qquad \textit{with} \quad c(d) = \frac{1.2803}{2 + \log d}.$$

In short, the preasymptotic rate of the L_2-approximation numbers of mixed order s Sobolev functions on the unit cube is $\frac{1.1929}{\log d}$ for $s = 1$, and in between $\frac{1.2803}{\log d}$ and $\frac{1.2825}{\log d}$ for any other $s \in \mathbb{N} \cup \{\infty\}$.

5 Tractability through Decreasing Complexity of the Univariate Problem

For every $d \in \mathbb{N}$, let X_d and Y_d be normed spaces and let F_d be a subset of X_d. We want to approximate the operator $T_d : F_d \to Y_d$ by an algorithm $A_n : F_d \to Y_d$ that uses at most n linear and continuous functionals on X_d. The nth minimal worst case error

$$e(n,d) = \inf_{A_n} \sup_{f \in F_d} \|T_d f - A_n f\|_{Y_d} \tag{5.1}$$

measures the worst case error of the best such algorithm A_n. If F_d is the unit ball of a pre-Hilbert space and T_d is linear, it is known to coincide with the $(n+1)$th approximation number of T_d. Conversely, the information complexity

$$n(\varepsilon, d) = \min \{n \in \mathbb{N}_0 \mid e(n,d) < \varepsilon\} \tag{5.2}$$

is the minimal number of linear and continuous functionals that is needed to achieve an error less than ε. The problem $\{T_d\}$ is called polynomially tractable, if there are nonnegative numbers C, p and q such that

$$n(\varepsilon, d) \leq C\, \varepsilon^{-q}\, d^p \qquad \text{for all } d \in \mathbb{N} \text{ and } \varepsilon > 0. \tag{5.3}$$

It is called strongly polynomially tractable, if (5.3) holds with p equal to zero. See [NW08] for a detailed treatment of these and other concepts of tractability.

In the following, X_d and Y_d will be Hilbert spaces and T_d will be a linear and compact norm-one operator with approximation numbers of polynomial decay. For example, one can think of T_d as the embedding of the Sobolev space $H^{s_d}(G)$ into $H^{r_d}(G)$ for some $r_d < s_d$ and a compact manifold G. Let T_d^d be the dth tensor power of T_d. In the chosen example, this is the embedding of $H^{s_d}_{\text{mix}}\left(G^d\right)$ into $H^{r_d}_{\text{mix}}\left(G^d\right)$. We will refer to $\{T_d\}$ as the univariate and to $\left\{T_d^d\right\}$ as the multivariate problem. It is proven in [NW08, Theorem 5.5] that the multivariate problem is not polynomially tractable, if T_d is the same operator for every $d \in \mathbb{N}$. This corresponds to the case, where the complexity of the univariate problem is constant in d. Can we achieve polynomial tractability of the multivariate problem, if the complexity of the univariate problem decreases, as d increases? If yes, to which extent do we have to simplify the univariate problem? The answer is given by the following theorem.

Theorem 5. *For every natural number d, let T_d be a compact norm-one operator between Hilbert spaces and let T_d^d be its dth tensor power. Assume that $a_n\left(T_d\right)$ is nonincreasing in d and $a_n\left(T_1\right)$ decays polynomially in n. The problem $\left\{T_d^d\right\}$ is strongly polynomially tractable, iff it is polynomially tractable, iff $a_2\left(T_d\right)$ decays polynomially in d.*

Proof. Clearly, strong polynomial tractability implies polynomial tractability.

Let $\left\{T_d^d\right\}$ be polynomially tractable and choose nonnegative numbers C, p and q such that

$$n(\varepsilon, d) = \#\left\{n \in \mathbb{N} \mid a_n(T_d^d) \geq \varepsilon\right\} \leq C\, \varepsilon^{-q}\, d^p \tag{5.4}$$

for all $\varepsilon > 0$ and $d \in \mathbb{N}$. In particular, there is an $r \in \mathbb{N}$ with

$$n\left(d^{-1}, d\right) \leq d^r - 1 \tag{5.5}$$

for every $d \geq 2$. If d is large enough, we can apply Part (ii) of Theorem 4 for $n = d^r$ and the estimate

$$\beta\left(d, d^r\right) = \frac{\log a_2(T_d)^{-1}}{\log\left(1 + \frac{v \cdot d}{r \log_{1+v} d}\right)} \leq \frac{2 \log a_2(T_d)^{-1}}{\log d} \tag{5.6}$$

to obtain

$$d^{-1} > a_{d^r}(T_d^d) \geq a_2(T_d) \cdot d^{-r\beta(d,d^r)} \geq a_2(T_d)^{2r+1}. \tag{5.7}$$

Consequently, $a_2(T_d)$ decays polynomially in d.

Now let $a_2(T_d)$ be of polynomial decay. Then there are constants $p > 0$ and $d_0 \in \mathbb{N}$ such that $a_2(T_d)$ is bounded above by d^{-p} for any $d \geq d_0$. On the other hand, there are positive constants C and s such that

$$a_n(T_d) \leq a_n(T_1) \leq C\, n^{-s}. \tag{5.8}$$

We apply Part (i) of Theorem 4 and the estimate

$$\alpha\,(d,1) = \frac{\log a_2(T_d)^{-1}}{\log d + \frac{2}{s}\log a_2(T_d)^{-1}} \geq \frac{p}{1+\frac{2p}{s}} = r > 0 \tag{5.9}$$

to obtain

$$a_n(T_d^d) \leq \left(\frac{\exp\left(C^{2/s}\right)}{n}\right)^r \tag{5.10}$$

for any $n \in \mathbb{N}$ and $d \geq d_0$. Consequently,

$$n(\varepsilon, d) = \#\left\{n \in \mathbb{N} \mid a_n(T_d^d) \geq \varepsilon\right\} \leq \exp\left(C^{2/s}\right) \cdot \varepsilon^{-1/r} \tag{5.11}$$

for any $d \geq d_0$ and $\varepsilon > 0$ and $\left\{T_d^d\right\}$ is strongly polynomially tractable. $\square$

Let us consider the spaces $H^s_{\text{mix}}\left(\mathbb{I}^d\right)$ and $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ as defined in Section 4.3. The L_2-approximation in these spaces is not polynomially tractable. Can we achieve polynomial tractability by increasing the smoothness with the dimension?

Corollary 9. *The problem $\left\{H^{s_d}_{\text{mix}}\left(\mathbb{I}^d\right) \hookrightarrow L_2\left(\mathbb{I}^d\right)\right\}$ is not polynomially tractable for any choice of natural numbers s_d. The problem $\left\{H^{s_d}_{\text{mix}}\left(\mathbb{T}^d\right) \hookrightarrow L_2\left(\mathbb{T}^d\right)\right\}$ is strongly polynomially tractable, iff it is polynomially tractable, iff $b-a < 2\pi$ and s_d grows at least logarithmically in d or $b-a = 2\pi$ and s_d grows at least polynomially in d.*

With regard to tractability, the L_2-approximation of mixed order Sobolev functions is hence much harder for nonperiodic than for periodic functions. The negative tractability result for nonperiodic functions can be explained by the difficulty of approximating d-variate polynomials with degree one or less in each variable and H^1_{mix}-norm less than one. The corresponding set of functions is contained in the unit ball of the nonperiodic space H^s_{mix} for every $s \in \mathbb{N} \cup \{\infty\}$.

Note that Corollary 9 for cubes of unit length is in accordance with [PW10], where Papageorgiou and Woźniakowski prove the corresponding statement for the L_2-approximation in Sobolev spaces of mixed smoothness $(s_1, \ldots, s_d)$ on the unit cube. The smoothness of such functions increases from variable to variable, but the smoothness with respect to a fixed variable does not increase with the dimension. There, the authors raise the question for a characterization of spaces and their norms for which increasing smoothness yields polynomial tractability. Theorem 5 says that in the setting of uniformly increasing mixed smoothness, polynomial tractability is achieved, if and only if it leads to a polynomial decay of the second singular value of the univariate problem. It would be interesting to verify whether the same holds in the case of variable-wise increasing smoothness and to compute the exponents of strong polynomial tractability.

The reason for the great sensibility of the tractability results for the periodic spaces to the length of the interval can be seen in the difficulty of approximating trigonometric polynomials with frequencies in $\frac{2\pi}{b-a}\{-1,0,1\}^d$ that are contained in the unit ball of $H^{\infty}_{\text{mix}}\left(\mathbb{T}^d\right)$. The corresponding set of functions is nontrivial, if and only if $\frac{2\pi}{b-a}$ is smaller than one.

It may yet seem unnatural that the approximation numbers are so sensible to the representation $[\boldsymbol{a},\boldsymbol{b}]$ of the d-torus or the d-cube. This can only happen, since the above and common scalar products

$$\langle f,g\rangle = \sum_{\alpha\in\{0,\dots,s\}^d} \langle D^\alpha f, D^\alpha g\rangle_{L_2} \tag{5.12}$$

do not define a homogeneous family of norms on $H^s_{\text{mix}}([\boldsymbol{a},\boldsymbol{b}])$. To see that, let T be the embedding of $H^s_{\text{mix}}([\boldsymbol{a},\boldsymbol{b}])$ into $L_2([\boldsymbol{a},\boldsymbol{b}])$ and let T_0 be the embedding in the case $[\boldsymbol{a},\boldsymbol{b}]=[0,1]^d$. The dilation operation $Mf=f(\boldsymbol{a}+(\boldsymbol{b}-\boldsymbol{a})\cdot)$ defines a linear homeomorphism both from $L_2([\boldsymbol{a},\boldsymbol{b}])$ into $L_2\left([0,1]^d\right)$ and from $H^s_{\text{mix}}([\boldsymbol{a},\boldsymbol{b}])$ into $H^s_{\text{mix}}\left([0,1]^d\right)$ and

$$T_0 = MTM^{-1}. \tag{5.13}$$

The L_2-spaces satisfy the homogeneity relation

$$\|Mf\|_{L_2\left([0,1]^d\right)} = \lambda^d([\boldsymbol{a},\boldsymbol{b}])\cdot\|f\|_{L_2([\boldsymbol{a},\boldsymbol{b}])} \quad \text{for} \quad f\in L_2([\boldsymbol{a},\boldsymbol{b}])\,. \tag{5.14}$$

If the chosen family of norms on $H^s_{\text{mix}}\left(\mathbb{T}^d\right)$ is also homogeneous, i.e.

$$\|Mf\|_{H^s_{\text{mix}}\left([0,1]^d\right)} = \lambda^d([\boldsymbol{a},\boldsymbol{b}])\cdot\|f\|_{H^s_{\text{mix}}([\boldsymbol{a},\boldsymbol{b}])} \quad \text{for} \quad f\in H^s_{\text{mix}}([\boldsymbol{a},\boldsymbol{b}])\,, \tag{5.15}$$

the approximation numbers of T and T_0 clearly must coincide. The above scalar products do not yield a homogeneous family of norms. An example of an equivalent and homogeneous family of norms on $H^s_{\mathrm{mix}}([\boldsymbol{a},\boldsymbol{b}])$ is defined by the scalar products

$$\langle f, g\rangle = \sum_{\alpha\in\{0,\dots,s\}^d} (\boldsymbol{b}-\boldsymbol{a})^{2\alpha} \langle D^\alpha f, D^\alpha g\rangle_{L_2}. \tag{5.16}$$

Hence, the approximation numbers and tractability results with respect to this scalar product do not depend on $\boldsymbol{a}$ and $\boldsymbol{b}$ at all. They coincide with the approximation numbers with respect to the previous scalar product on $H^s_{\mathrm{mix}}\left([0,1]^d\right)$.

References

[B60] K. I. Babenko: *About the approximation of periodic functions of many variable trigonometric polynomials.* Dokl. Akad. Nauk SSR **32**, 247–250, 1960.

[CD16] A. Chernov, D. Dũng: *New explicit-in-dimension estimates for the cardinality of high-dimensional hyperbolic crosses and approximation of functions having mixed smoothness.* J. Complexity **32**, 92–121, 2016.

[CW16] J. Chen, H. Wang: *Preasymptotics and asymptotics of approximation numbers of anisotropic Sobolev embeddings.* J. Complexity, 2016. http://dx.doi.org/10.1016/j.jco.2016.10.005

[CW17] J. Chen, H. Wang: *Approximation numbers of Sobolev and Gevrey type embeddings on the sphere and on the ball – Preasymptotics, asymptotics, and tractability.* ArXiv e-prints, 2017. arXiv:1701.03545 [math.CA]

[DTU16] D. Dũng, V.N. Temlyakov, T. Ullrich: *Hyperbolic cross approximation.* ArXiv e-prints, 2015. arXiv:1601.03978 [math.NA]

[GW11] M. Gnewuch, H. Woźniakowski: *Quasi-polynomial tractability.* J. Complexity **27**, 312–330, 2011.

[K84] H. König: *On the tensor stability of s-number ideals.* Math. Ann. **269**, 77–93, 1984.

[KSU15] T. Kühn, W. Sickel, T. Ullrich: *Approximation of mixed order Sobolev functions on the d-torus – asymptotics, preasymptotics and d-dependence.* Constructive Approximation **42**, 353–398, 2015.

[KMU16] T. Kühn, S. Mayer, T. Ullrich: *Counting via entropy: new preasymptotics for the approximation numbers of Sobolev embeddings.* SIAM J. Numerical Analysis **54**(6), 3625–3647, 2016.

[M62] B.S. Mityagin: *Approximation of functions in L^p and C on the torus.* Math. Notes **58**, 397–414, 1962.

[N74] N.S. Nikol'skaya: *Approximation of differentiable functions of several variables by Fourier sums in the L_p-metric.* Sibirsk. Mat. Zh. **15**, 395–412, 1974; English transl. in Siberian Math. J. **15**, 1974.

[NW08] E. Novak, H. Woźniakowski: *Tractability of Multivariate Problems. Volume I: Linear Information.* EMS, Zürich, 2008.

[PW10] A. Papageorgiou, H. Woźniakowski: *Tractability through increasing smoothness.* J. Complexity **26**, 409–421, 2010.

[P82] A. Pietsch: *Tensor products of sequences, functions, and operators.* Arch. Math. **38**, 335–344, 1982.

[T86] V.N. Temlyakov: *Approximation of functions with bounded mixed derivative.* Trudy MIAN **178**, 1–112, 1986; English transl. in Proc. Steklov Inst. Math. **1**, 1989.

[T96] C. Thomas-Agnan: *Computing a family of reproducing kernels for statistical applications.* Numerical Algorithms **13**, 21–32, 1996.