Tractable s-numbers of mixed Wiener spaces

Moritz Moeller

Siegmundsburg Seminar on Analysis & Theoretical Numerics

30.08.2023

UNIVERSITY OF TECHNOLOGY
IN THE EUROPEAN CAPITAL OF CULTURE
CHEMNITZ

Outline

mixed Wiener spaces

For $\alpha > 0$ and $0 < p < \infty$ we define the mixed Wiener space $\mathcal{A}_p^\alpha(\mathbb{T}^d) \subset L_1(\mathbb{T}^d)$ via its norm

$$\|f\|_{\mathcal{A}_p^\alpha(\mathbb{T}^d)} = \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d} \prod_{i=1}^d (1+|k_i|)^{\alpha p} |\hat{f}(\mathbf{k})|^p \Big)^{\frac{1}{p}}.$$

They have a useful embedding into the sequence spaces

$$A_\alpha f = \Big(\prod_{i=1}^d (1+|k_i|)^\alpha \hat{f}(\mathbf{k}) \Big)_{\mathbf{k}\in\mathbb{Z}^d}, \quad \|A_\alpha : \mathcal{A}_p^\alpha(\mathbb{T}^d) \to \ell_p(\mathbb{Z}^d)\| = 1.$$

quasi s-Numbers

For $n \in \mathbb{N}_0$, and $X(\Omega), Y$ quasi-Banach function spaces with a continuous linear embedding $T : X \to Y$ the following (quasi) s-Numbers are defined:

- Sampling numbers (linear and non-linear)

$$\varrho_n(X)_Y = \inf_{t_1 \ldots t_n \in \Omega} \inf_{R:\mathbb{C}^n \to Y} \sup_{\|f\|_X \leq 1} \big\| f - R\big(f(t_1) \ldots f(t_n)\big) \big\|_Y \tag{1}$$

- Gelfand numbers

$$c_n(T : X \to Y) = \inf \left\{ \sup_{f \in B_X \cap M} \|Tf\|_Y : M \subset X \text{ linear subspace with } \operatorname{codim} M < n \right\} \tag{2}$$

- best trigonometric m-term approximation

$$\sigma_n(X)_Y := \sup_{\|f\|_X \leq 1} \inf_{s \in \Sigma_n} \|f - s\|_Y \tag{3}$$

Motivation

- Nguyen Nguyen and Sickel recently studied some s-numbers of mixed Wiener classes in [1], however they studied neither Gelfand numbers, sampling numbers nor best m-term approximation
- new results concerning sampling numbers

Proposition 1 ([2, Jahn, Ullrich and Voigtlaender 2023])

Let $n, d \in \mathbb{N}$ then it holds for a quasi-normed function space with continuous embedding into L_∞

$$\varrho_{n \log(n)^3}(\mathcal{F})_2 \lesssim \sigma_n(\mathcal{F})_\infty. \tag{4}$$

Relations between s-numbers

- Gelfand numbers form a lower bound for the non-linear sampling numbers, in particular it holds
$$\varrho_n\big(\mathcal{A}_p^\alpha(\mathbb{T}^d)\big)_2 \gtrsim c_n(id : \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2)$$
- Kolmogorov numbers form a lower bound for the linear sampling numbers, in particular it holds
$$\varrho_n^{\mathrm{lin}}\big(\mathcal{A}_p^\alpha(\mathbb{T}^d)\big)_2 \gtrsim d_n(id : \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2)$$

In total the Gelfand and sampling numbers give upper and lower bounds for the non-linear sampling numbers.
One important property of (quasi) s-numbers is, that for two operators S, R it holds

$$s_n(R \circ S) \leq s_n(R)s_1(S) = s_n(R)\|S\|$$

Theorem 2

For $n, d \in \mathbb{N}$, $0 < p \leq 2$ *and* $\alpha > \left(\frac{p-1}{p}\right)_+$ *it holds*

$$c_n(\mathrm{id} : \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) \asymp n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha} \tag{5}$$

where $\lambda = \frac{1}{p} - \frac{1}{2}$.

Idea of proof: Rewrite

$$\mathrm{id} : \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)$$

as

$$D_\alpha(x_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}} = \left(\prod_{i=1}^d (1 + |k_i|)^{-\alpha} x_{\mathbf{k}}\right)_{\mathbf{k}\in\mathbb{Z}}.$$

diagonal operator

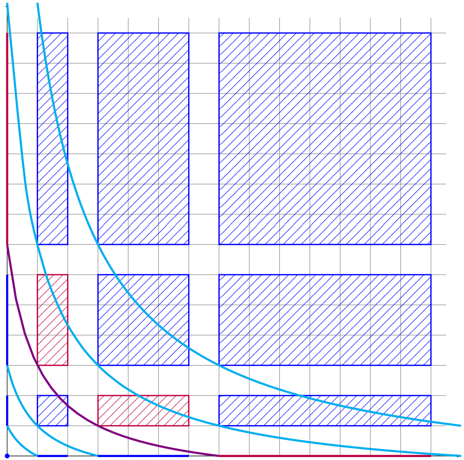
$$\begin{array}{ccc}
\mathcal{A}_p^\alpha(\mathbb{T}^d) & \xrightarrow{\quad \mathrm{id} \quad} & L_2(\mathbb{T}^d) \\
\Big\downarrow {\scriptstyle A_\alpha} & & \Big\uparrow {\scriptstyle B} \\
\ell_p(\mathbb{Z}^d) & \xrightarrow[\quad D_\alpha \quad]{} & \ell_2(\mathbb{Z}^d)
\end{array}$$

where

$$B(x_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}} = \frac{1}{\sqrt{2\pi}^d} \sum_{\mathbf{k}\in\mathbb{Z}} x_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}}, \quad \|B\| = 1$$

$$c_n(\mathrm{id} : \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) = c_n(D_\alpha : \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)), \tag{6}$$

hyperbolic cross



A hyperbolic cross is a set of the form $\left\{\mathbf{n} \in \mathbb{N}_0^d \middle| \prod_{j=1}^d (n_j + 1) \le c\right\}$. Decompose now $\mathbb{N}_0^d$ in dyadic blocks, where blocks on the same hyperbolic layer have

- the same number of points
- the same maximal weight

Now use this decomposition on D_α. The number of points per layer can now be computed as

$$C_j := \#\square_j = 2^j \binom{j+d-1}{j} \asymp 2^j j^{d-1}.$$

decomposition of the diagonal operator

We can now decompose the diagonal operator in the same way. To that end, now call D_α restricted to the k-th layer Δ_k then it holds,

$$c_n(D_\alpha) \le \Big(\sum_{j=0}^{L} c_{n_j}(\Delta_j) + \sum_{j=L+1}^{M} c_{n_j}(\Delta_j) + c_1\big(\sum_{j=M+1}^{\infty} \Delta_j\big)\Big) =: S_1 + S_2 + S_3, \tag{7}$$

where we do an exact approximation on S_1 a smart approximation on S_2 and no approximation at all on S_3.

best m-term approximation

Theorem 3

For $n, d \in \mathbb{N}$ *with* $0 < p \leq q$ *and* $2 \leq q \leq \infty$ *as well as* $\alpha > \left(\frac{p-1}{p}\right)_+$ *it holds*

$$n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha} \lesssim \sigma_n\left(\mathcal{A}_p^\alpha(\mathbb{T}^d)\right)_q \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+\mu} \tag{8}$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2},$$

and $\mu = \frac{1}{2}$ *if both* $q = \infty$ *and* $d > 1$ *otherwise* $\mu = 0$.

basis pursuit denoising

- sparse trigonometric polynomials can be approximated well via l_1 minimisation, even when the data is noisy
- the previous result for best trigonometric m-term approximation ensures that for every function f from a weighted mixed Wiener class there is such a sparse trigonometric polynomial that approximates f well

Proposition 4 ([5, Rauhut, 2008])

Let $\mathbf{c}^*$ be the solution of the minimisation problem

$$\min \|\mathbf{c}\|_1 \quad \text{subject to} \, \|F_{\mathbf{X}}\mathbf{c} - y\|_2 \leq \nu$$

then the bound

$$\|\mathbf{c} - \mathbf{c}^*\|_2 \leq C_1 \frac{\nu}{\sqrt{N}}$$

holds with high probability, if at least $N \geq C_0 \, n \, d \log(n)^4 \log(\varepsilon^{-1})$ samples were used.

Basis pursuit denoising - numerical experiments

Figure 3: Approximation in 3 Dimensions

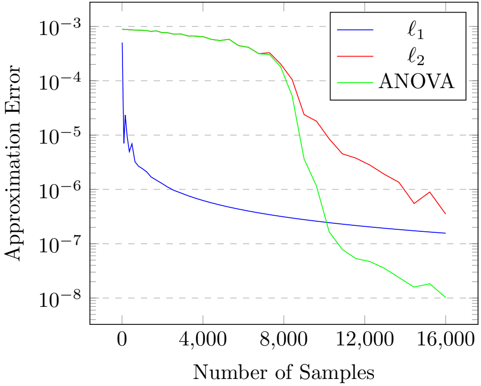
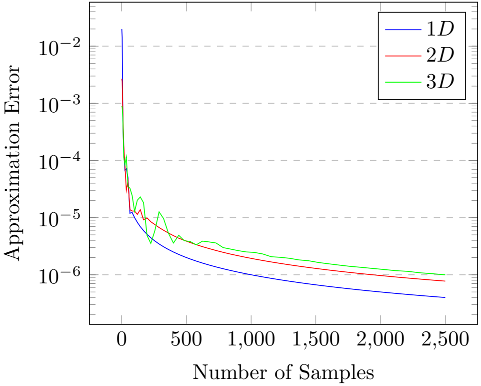


Figure 4: Dimension Comparison



moritz.moeller@mathematik.tu-chemnitz.de

linear sampling numbers

Proposition 5 (see [1, Nguyen, Nguyen and Sickel, 2022])

For the Kolmogorov numbers d_n it holds for $\alpha > 0$,

$$d_n\big(id : \mathcal{A}_1^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)\big) \asymp n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{9}$$

Since the Kolmogorov numbers form a lower bound for the linear sampling numbers this immediately gives the following result

$$\varrho_n^{\mathrm{lin}}\big(\mathcal{A}_1^\alpha(\mathbb{T}^d)\big)_2 \gtrsim n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{10}$$

non-linear sampling numbers

For the non-linear sampling numbers an analogous bound holds in terms of the Gelfand numbers

$$\varrho_n\big(\mathcal{A}_1^\alpha(\mathbb{T}^d)\big)_2 \gtrsim n^{-(\alpha+\frac{1}{2})}\log(n)^{\alpha(d-1)}. \tag{11}$$

Proposition 12 together with Theorem 3 now yields

$$\varrho_n\big(\mathcal{A}_1^\alpha(\mathbb{T}^d)\big)_2 \lesssim n^{-(\alpha+\frac{1}{2})}\log(n)^{\alpha(d-1)+3(\alpha+\frac{1}{2})+\frac{1}{2}} \tag{12}$$

There is a difference of $\frac{1}{2}$ in the main rate of the decay between the linear and non-linear sampling numbers in mixed Wiener classes measured in L_2.

tractable bound on Gelfand numbers

Theorem 2 states

$$c_n(\mathrm{id} : \mathcal{A}_p^\alpha(\mathbb{T}^d))_2 \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}$$

Where another 2^d term is hidden by the $\lesssim$. For $n = d^s$ this is a horrible bound.

Theorem 6

Let $m, d \in \mathbb{N}$ such that $p \leq 1$ and further $\alpha > 0$ then it holds

$$c_n(\mathcal{A}_p^\alpha(\mathbb{T}^d))_2 \leq C n^{-\lambda} d \log(n)^\lambda \tag{13}$$

where C is a constant dependent only on p.

For $s > \lambda^{-1}$ this is decaying.

tractable bound on the best m-term approximation

Again the original Theorem 3 states

$$\sigma_n\big(\mathcal{A}_p^\alpha(\mathbb{T}^d)\big)_q \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+\mu}$$

Where another 2^d term is hidden by the $\lesssim$. This is again not suitable in a setting where $n = d^s$.

Theorem 7

Let $m, d \in \mathbb{N}$ such that $p \leq 1$ and further $\alpha > 0$ then it holds

$$\sigma_n\big(\mathcal{A}_p^\alpha(\mathbb{T}^d)\big)_\infty \lesssim n^{-\lambda} d \log(n)^{\frac{1}{p}} \tag{14}$$

with no hidden d-dependent constants.

tractable bound on the best m-term approximation

Again the original Equation (12) states

$$\varrho_n\left(\mathcal{A}_1^\alpha(\mathbb{T}^d)\right)_2 \lesssim n^{-(\alpha+\frac{1}{2})}\log(n)^{\alpha(d-1)+3(\alpha+\frac{1}{2})+\frac{1}{2}}$$

Where again a 2^d term is hidden in the $\lesssim$.

Theorem 8

For $0 < p \leq 1$ and $\alpha > \left(\frac{p-1}{p}\right)_+$ it holds

$$\varrho_n\left(\mathcal{A}_p^\alpha(\mathbb{T}^d)\right)_2 \leq Cn^{-\lambda} d\log(n)^{\frac{1}{p}+3\lambda}. \tag{15}$$

where C is an absolute constant (and not dependent on d).

Thank you for your attention

References I

- Van Dung Nguyen, Van Kien Nguyen, and Winfried Sickel
  s-Numbers of Embeddings of Weighted Wiener Algebras
  Journal of Approximation Theory, Volume 279, July 2022, 105745
- Thomas Jahn, Tino Ullrich, Felix Voigtlaender
  Sampling numbers of smoothness classes via ℓ_1-minimization
  preprint, arxiv.2212.00445 2023
- Jan Vybiral
  Widths of embeddings in function spaces
  Journal of Complexity,Volume 24, Issue 4, August 2008, Pages 545-570
- Simon Foucart, Alain Pajor, Holger Rauhut, Tino Ullrich
  The Gelfand widths of ℓ_p-balls for $0 < p \leq 1$
  Journal of Complexity, Volume 26, Issue 6, December 2010, Pages 629-640

References II

Holger Rauhut
Stability Results for Random Sampling of Sparse Trigonometric Polynomials
IEEE Transactions on Information Theory, Volume: 54, Issue: 12, December 2008