Gelfand and sampling numbers for weighted mixed Wiener classes in L_2

Moritz Moeller

Mathematical Signal and Image Analysis

21.03.2023

UNIVERSITY OF TECHNOLOGY
IN THE EUROPEAN CAPITAL OF CULTURE
CHEMNITZ

Outline

1. Introduction

2. Gelfand numbers

3. Best trigonometric m-term approximation

4. Sampling numbers

weighted mixed Wiener spaces

For $\alpha > 0$ and $0 < p < \infty$ we define the weighted 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 weighted 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 und 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.

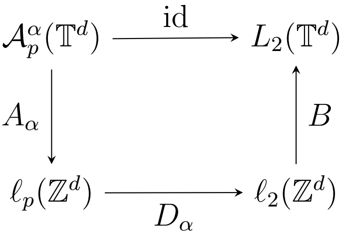
Theorem 2

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 (for $p = 1$):



diagonal operator

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

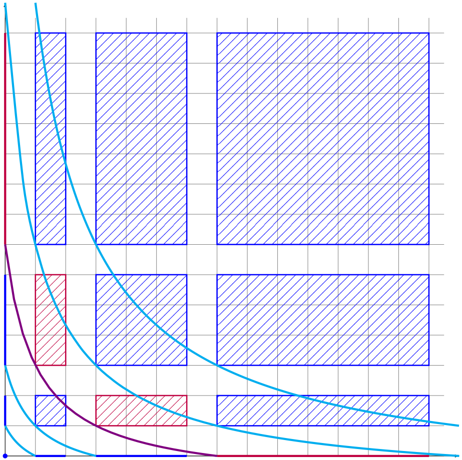
From this it follows

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

with

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

hyperbolic cross



A hyperbolic cross is a set of the form $\left\{\mathbf{n} \in \mathbb{N}_0^d \middle| \prod_{j=1}^d (n_j+1) \leq c\right\}$. 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

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,

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

where $L = \lfloor m - (d-1)\log_2 m \rfloor$, $M = \lfloor m(1+\frac{\lambda}{\alpha}) - (d-1)\log_2 m \rfloor$ for $2^m = n$.
For $j = 0 \ldots L$ chose $n_j = 2C_j$. Then $S_1 = 0$ and

$$\sum_{j=0}^{L} n_j \asymp \sum_{j=0}^{L} 2^j j^{d-1} \leq \sum_{j=0}^{L} 2^j m^{d-1} \asymp 2^L m^{(d-1)} = 2^m.$$

Definition 3

The space ℓ_p^N *is defined for* $0 < p < \infty$ *as*

$$\ell_p^N := \left\{ (x_n)_{n=1}^N \,\middle|\, \|\mathbf{x}|\ell_p^N\| := \Big(\sum_{n=1}^N |x_n|^p \Big)^{\frac{1}{p}} < \infty \right\},$$

with the usual modifications for $p = \infty$.

Proposition 4 ([3, Vybiral 2008] and [4, Foucart, Pajor, Rauhut and Ullrich 2010])

Let $0 < p \leq 1$ *then it holds*

$$c_m(\mathrm{id} : \ell_p^N \to \ell_2^N) \asymp \Big(\frac{1 + \log(\frac{N}{m})}{m} \Big)^{\frac{1}{p} - \frac{1}{2}},$$

for $m < N$.

Chose now $n_j = 2^j 2^{(L-j)\eta} j^{d-1}$ for $j = L+1, \dots, M$ and $\eta > 1$, then S_2 can be estimated as follows

$$
\begin{aligned}
S_2 = \sum_{j=L+1}^{M} c_{n_j}(\Delta_j : \ell_p \to \ell_2) &\asymp \sum_{j=L+1}^{M} c_{n_j}(\mathrm{id} : \ell_p^{C_j} \to \ell_2^{C_j}) 2^{-j\alpha} \\
&\asymp \sum_{j=L+1}^{M} \left(\frac{\log\left(\frac{2^j j^{d-1}}{2^{j-1} 2^{(L-j)\eta} j^{d-1}}\right)}{2^{j-1} 2^{(L-j)\eta} j^{d-1}} \right)^{\lambda} 2^{-j\alpha} \\
&\asymp 2^{-L(\lambda+\alpha)} \sum_{j=L+1}^{M} \left[(j-L)\eta 2^{-(j-L)[\boldsymbol{\alpha-(\eta-1)\lambda}]} j^{-\lambda(d-1)} \right] \\
&\asymp 2^{-(m-(d-1)\log_2 m)(\lambda+\alpha)} (m-(d-1)\log_2 m)^{-\lambda(d-1)} \\
&\asymp 2^{-m(\lambda+\alpha)} m^{(d-1)\alpha}.
\end{aligned} \tag{8}
$$

It holds that

$$\sum_{j=L+1}^{M} n_j \asymp 2^{L\eta} \sum_{j=L+1}^{M} 2^{j(1-\eta)} j^{d-1} \asymp 2^L L^{d-1} \leq 2^L m^{d-1} = 2^m. \tag{9}$$

S_3 can be controlled by its norm

$$\begin{aligned} S_3 = c_1\Big(\sum_{j=M+1}^{\infty} \Delta_j : \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)\Big) &\leq \Big\| \sum_{j=M+1}^{\infty} \Delta_j : \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d) \Big\| \\ &\lesssim 2^{-M\alpha} \\ &= 2^{-(m(1+\frac{\lambda}{\alpha})-(d-1)\log_2 m)\alpha} \\ &= 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha}. \end{aligned} \tag{10}$$

best m-term approximation

Theorem 5

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

$$n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha} \lesssim \sigma_n\left(\mathcal{A}_p^\alpha(\mathbb{T}^d)\right)_q \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+\mu} \tag{11}$$

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 6 ([5, Rauhut, 2008])

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

$$\min \|\mathbf{c}\|_1 \quad \text{subject to} \, \|F_{\mathbf{X}}\mathbf{c} - y\|_2 \leq \nu \tag{12}$$

then the bound

$$\|\mathbf{c} - \mathbf{c}^*\|_2 \leq C_1 \frac{\nu}{\sqrt{N}} \tag{13}$$

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

linear sampling numbers

Proposition 7 (see [1, Nguyen, Nguyen und Sickel, 2022])

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

$$d_n\big(id : \mathcal{A}_1^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)\big) \asymp n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{14}$$

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

$$\varrho_n^{\mathrm{lin}}\big(\mathcal{A}_1^\alpha(\mathbb{T}^d)\big)_2 \gtrsim n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{15}$$

non-linear sampling numbers

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

$$\varrho_n\big(\mathcal{A}_1^\alpha(\mathbb{T}^d)\big)_2 \gtrsim n^{-(\alpha+\frac{1}{2})}\log(n)^{\alpha(d-1)}. \tag{16}$$

Proposition 1 together with Theorem 5 now yields

$$\varrho_n\big(\mathcal{A}_1^\alpha(\mathbb{T}^d)\big)_2 \lesssim n^{-(\alpha+\frac{1}{2})}\log(n)^{\alpha(d-1)+3(\alpha+\frac{1}{2})+\frac{1}{2}} \tag{17}$$

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

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