UNIVERSITY OF TECHNOLOGY
IN THE EUROPEAN CAPITAL OF CULTURE
CHEMNITZ

Learning anisotropy parameters for ANOVA approximation

F. Bartel and P. Schröter

Goal

Task

Unknown function: $f\colon \mathbb{T}^d \to \mathbb{C}$

Given:

- $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_M\} \subset \mathbb{T}^d$
- $\mathbf{y} = (f(\mathbf{x}_j))_{j=1}^M \in \mathbb{C}^M$

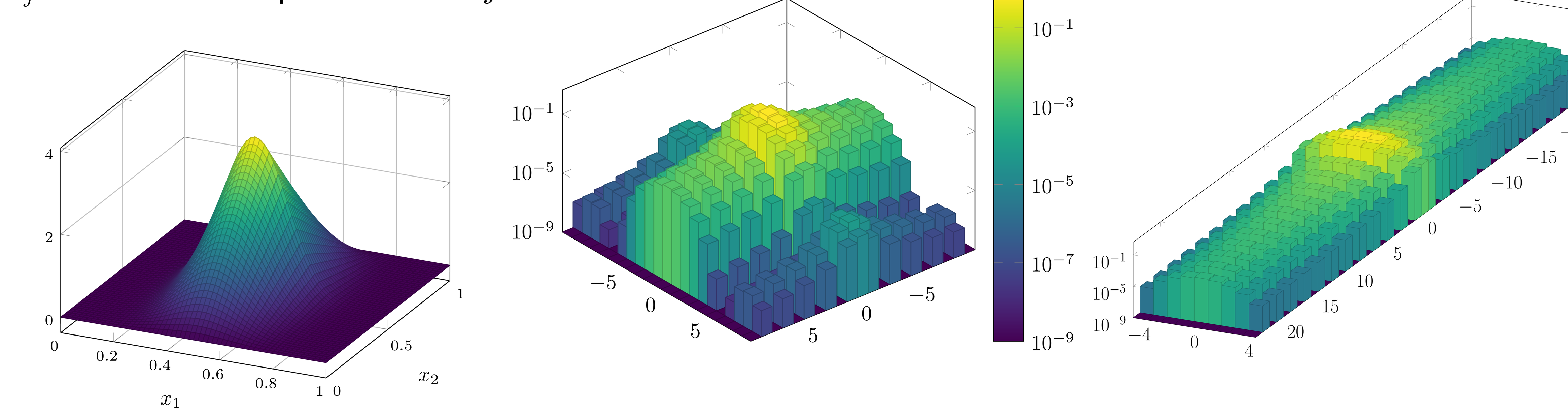
Goal: find $\tilde f \approx f$

Ansatz: determine

- $B \in \mathbb{N}, \mathcal{I} \subset \mathbb{Z}^d, |\mathcal{I}| = B$
- $\hat f_{\mathbf{k}}, \mathbf{k} \in \mathcal{I}$

such that $\tilde f := \sum_{\mathbf{k}\in\mathcal{I}} \hat f_{\mathbf{k}} \mathrm{e}^{2\pi\mathrm{i}\langle \mathbf{k},\cdot\rangle} \approx f$

$f(x_1,x_2) = b_6(x_1)b_2(x_2)$ with b_j the cardinal B-Spline of order j | $\hat f_{\mathbf{k}}$, $\mathbf{k}\in[-10,10]^2$ | $\hat f_{\mathbf{k}}$, $\mathbf{k}\in[-4,4]\times[-24,24]$



Motivation

Lemma

- $\mathbf{N} = (N_1,\dots,N_d) \in (2\mathbb{N})^d$
- $\mathcal{I}_{\mathbf{N}} := \bigtimes_{j=1}^d [-N_j/2, N_j/2) \cap \mathbb{Z}^d$, i.e. width N_j in the dimension j
- f in the anisotropic Sobolev spaces $H^{s_1,\dots,s_d}$ with $H^{s_1,\dots,s_d} := \{f \in L_2 \mid D^{\boldsymbol\alpha} f \in L_2 \text{ for } \boldsymbol\alpha \in \mathbb{N}_0^d \text{ with } \sum_{j=1}^d \alpha_j/s_j \le 1\}$

Then

$$\sup_{\|f\|_{H^{s_1,\dots,s_d}}\le 1} \|f - P_{\mathcal{I}_{\mathbf{N}}} f\|_{L_2}^2 = \left(\max\left\{1, \left(\frac{N_1}{2}\right)^{s_1}, \dots, \left(\frac{N_d}{2}\right)^{s_d}\right\}\right)^{-2}.$$

Example

frequency cubes $\mathcal{I}_{\mathbf{N}}$

with $N_j = \sqrt[d]{B}, |\mathcal{I}_{\mathbf{N}}| = B$

$$\boxed{\sup_{\|f\|_{H^{s_1,\dots,s_d}}\le 1} \|f - P_{\mathcal{I}_{\mathbf{N}}} f\|_{L_2}^2 \sim B^{-\frac{2\min\{s_1,\dots,s_d\}}{d}}}$$

optimal frequency boxes $\mathcal{I}_{\mathbf{N}^*}$

with $N_j^* = B^{\frac{1/s_j}{1/s_1+\dots+1/s_d}}, |\mathcal{I}_{\mathbf{N}^*}| = B$

$$\boxed{\sup_{\|f\|_{H^{s_1,\dots,s_d}}\le 1} \|f - P_{\mathcal{I}_{\mathbf{N}^*}} f\|_{L_2}^2 \sim B^{-\frac{2}{1/s_1+\dots+1/s_d}}}$$

For $f(x_1,x_2) = b_6(x_1)b_2(x_2)$ with the smoothness $s_1 = \frac{11}{2}$ and $s_2 = \frac{3}{2}$ it follows

$$\|f - P_{\mathcal{I}_{\mathbf{N}}} f\|_{L_2}^2 \lesssim B^{-\frac{3}{2}} \qquad \|f - P_{\mathcal{I}_{\mathbf{N}^*}} f\|_{L_2}^2 \lesssim B^{-\frac{33}{14}} \lesssim B^{-2.357}.$$

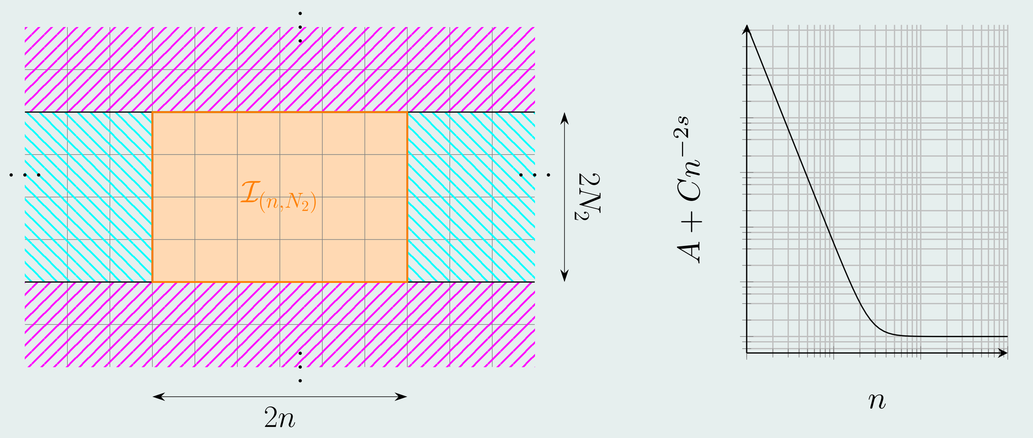
Ansatz

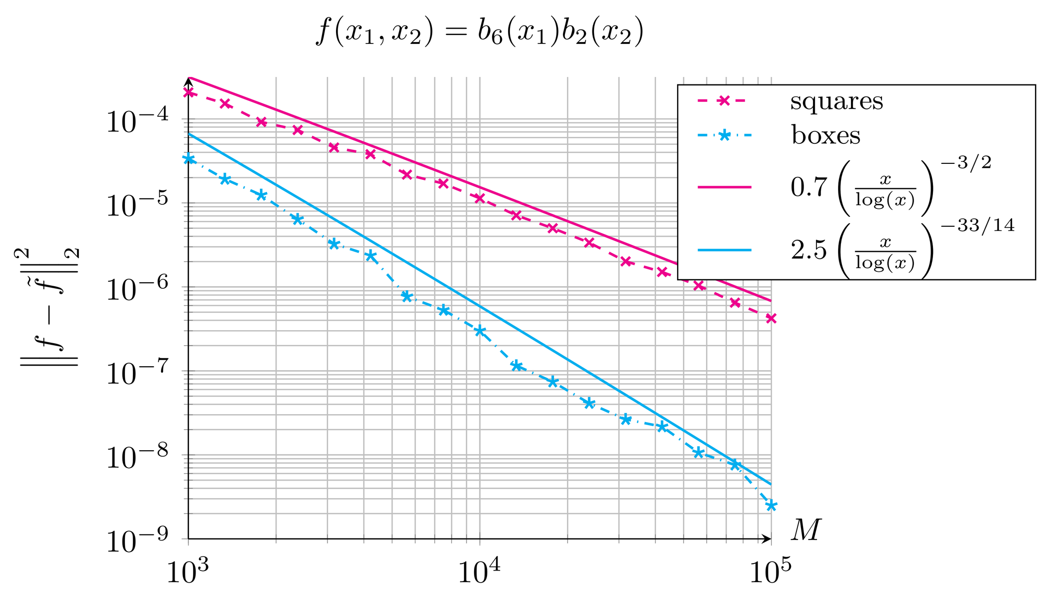
How to find the smoothness $s_1,\dots,s_d$?

Example: $d=2$, find smoothness s_1

$$\left\|f - P_{\mathcal{I}_{(n,N_2)}} f\right\|_{L_2}^2 = \left\|\sum_{\mathbf{k}\notin\mathcal{I}_{(n,N_2)}} c_{\mathbf{k}}(f)\right\|_{L_2}^2 = \sum_{\mathbf{k}\notin\mathcal{I}_{(n,N_2)}} |c_{\mathbf{k}}(f)|^2$$

$$= \sum_{\substack{k_1\in\mathbb{Z}\\ |k_2|>N_2}} |c_{\mathbf{k}}(f)|^2 + \sum_{\substack{|k_1|>n\\ |k_2|\le N_2}} |c_{\mathbf{k}}(f)|^2 \lesssim \underbrace{A + C_1 n^{-2s_1}}_{\star}$$





ANOVA

ANOVA approximation

$$f = \sum_{\mathbf{u}\in\mathcal{P}([d])} f_{\mathbf{u}} \quad\rightsquigarrow\quad f \approx \sum_{\mathbf{u}\in U} f_{\mathbf{u}} \quad\rightsquigarrow\quad f \approx \sum_{\mathbf{u}\in U} f_{\mathbf{u}} \quad\rightsquigarrow\quad f \approx \sum_{\mathbf{u}\in U} f_{\mathbf{u}}$$

$$f_{\mathbf{u}} = \sum_{\substack{\mathbf{k}\in\mathbb{Z}^d\\ \operatorname{supp}\mathbf{k}=\mathbf{u}}} c_{\mathbf{k}}(f)\mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle} \qquad f_{\mathbf{u}} = \sum_{\substack{\mathbf{k}\in\mathbb{Z}^d\\ \operatorname{supp}\mathbf{k}=\mathbf{u}}} c_{\mathbf{k}}(f)\mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle} \qquad f_{\mathbf{u}} \approx \sum_{\mathbf{k}\in\mathcal{I}_{\mathbf{N}_{\mathbf{u}}}} c_{\mathbf{k}}(f)\mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle} \qquad f_{\mathbf{u}} \approx \sum_{\mathbf{k}\in\mathcal{I}_{\mathbf{N}_{\mathbf{u}}}} \hat f_{\mathbf{k}}\mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}$$

$$c_{\mathbf{k}}(f) = \langle f, \mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}\rangle_{L_2} \qquad c_{\mathbf{k}}(f) = \langle f, \mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}\rangle_{L_2} \qquad c_{\mathbf{k}}(f) = \langle f, \mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}\rangle_{L_2} \qquad \hat f_{\mathbf{k}} \approx c_{\mathbf{k}}(f) = \langle f, \mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}\rangle_{L_2}$$

ANOVA decomposition

We define the ANOVA terms [2]

$$f_{\mathbf{u}} := \sum_{\substack{\mathbf{k}\in\mathbb{Z}^d\\ \operatorname{supp}\mathbf{k}=\mathbf{u}}} c_{\mathbf{k}}(f)\mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}$$

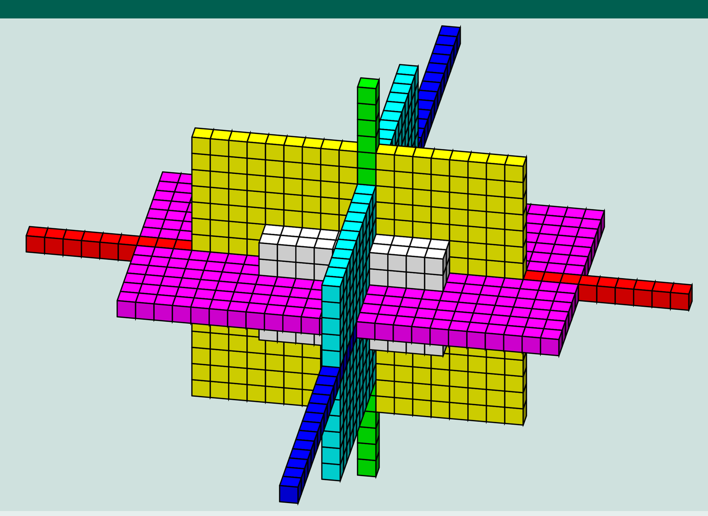
for $\mathbf{u} \subset \{1,\dots,d\} \Rightarrow \sum_{\mathbf{u}\in\mathcal{P}([d])} f_{\mathbf{u}}$.

$f \in H^{s_1,\dots,s_d} \Rightarrow f_{\mathbf{u}} \in H^{s_{u_1},\dots,s_{u_{|\mathbf{u}|}}}$

ANOVA truncation

$$\begin{aligned} f \approx\ & f_\emptyset \\ &+ f_{\{1\}} + f_{\{2\}} + \dots + f_{\{d\}} \\ &+ f_{\{1,2\}} + f_{\{1,3\}} + \dots + f_{\{d-1,d\}} \\ &+ f_{\{1,2,3\}} + f_{\{1,2,4\}} + \dots + f_{\{d-2,d-1,d\}} \\ &\cancel{+ f_{\{1,2,3,4\}} + \dots + f_{\{d-3,d-2,d-1,d\}}} \\ &\vdots \\ &\cancel{+ f_{\{1,2,\dots,d\}}} \end{aligned}$$

Frequency truncation [3]



Least squares $(\star\star)$

$$\min \left\|(\tilde f(\mathbf{x}_j) - f(\mathbf{x}_j))_{j=1}^d\right\|_2^2 = \left\| \underbrace{\begin{pmatrix} \phi_{\mathbf{k}_1}(\mathbf{x}_1) & \cdots & \phi_{\mathbf{k}_B}(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ \phi_{\mathbf{k}_1}(\mathbf{x}_M) & \cdots & \phi_{\mathbf{k}_B}(\mathbf{x}_M) \end{pmatrix}}_{=:\mathbf{A}\in\mathbb{C}^{M\times B},\ M\ge B} \underbrace{\begin{pmatrix} \hat f_{\mathbf{k}_1} \\ \vdots \\ \hat f_{\mathbf{k}_B} \end{pmatrix}}_{\in\mathbb{C}^B} - \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix}}_{\in\mathbb{C}^M} \right\|_2^2$$

with $\mathcal{I} = \{\mathbf{k}_1,\dots,\mathbf{k}_B\}$, $\phi_{\mathbf{k}} = \mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{k},\cdot\rangle}$

$B \sim M\log(M) \Rightarrow \mathbf{A}$ has a good condition number [1]

Results

Theorem (Bartel and Schröter '24)

- $B \in \mathbb{N}$
- $U \subseteq \mathcal{P}(\{1,\dots,d\})$
- $C_{\mathbf{u},j} > 0$ and $s_{\mathbf{u},j} > 0$ with $j \in \mathbf{u}$ from $\star$ for $\mathbf{u}\in U$

Then

$$\begin{array}{c}\min_{\mathbf{N}} \|f - P_{\mathcal{I}} f\|_{L_2}^2 \\ \text{s.t. } |\mathcal{I}_{\mathbf{N}}| = B\end{array} \iff \begin{array}{c}\min_{\mathbf{N}} \sum_{\mathbf{u}\in U}\sum_{j\in\mathbf{u}} C_{\mathbf{u},j} N_{\mathbf{u},j}^{-2s_{\mathbf{u},j}} \\ \text{s.t. } \sum_{\mathbf{u}\in U}\prod_{j\in\mathbf{u}} N_{\mathbf{u},j} = B.\end{array} \qquad (\star\star\star)$$

We can solve optimisation problem $(\star\star\star)$.

Numerics

Algorithm:

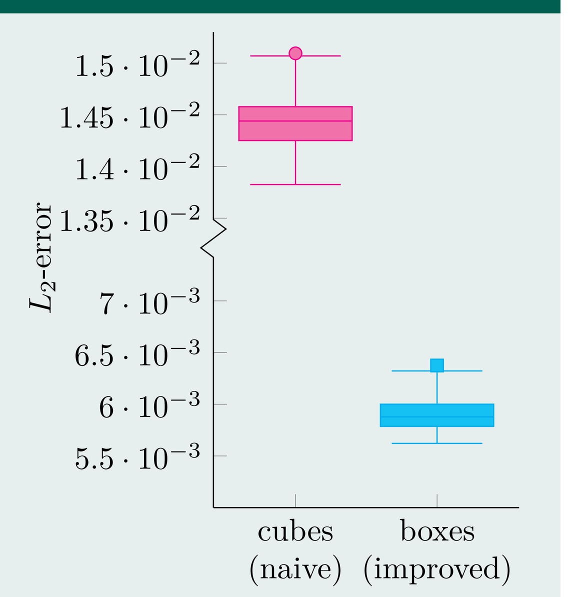
1. set initial $\mathcal{I}$
2. get $\hat f$ using $(\star\star)$
3. solve $(\star\star\star)$ to update $\mathcal{I}$
4. get $\hat f$ using $(\star\star)$
5. repeat 3.& 4. to improve $\mathcal{I}$

Example:

$f\colon \mathbb{T}^9 \to \mathbb{C}$

$$\begin{aligned} f(\mathbf{x}) = &\ b_2(x_1)b_4(x_3)b_6(x_8) \\ &+ b_2(x_2)b_4(x_5)b_6(x_6) \\ &+ b_2(x_4)b_4(x_7)b_6(x_9) \end{aligned}$$

$M = 10000$ points
$B = 6730$ frequencies



[1] L. Kämmerer, T. Ullrich, T. Volkmer, **Worst-case recovery guarantees for least squares approximation using random samples**, *Constr. Approx.*, 54:295–352, (2021).

[2] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski and H. Woźniakowski, **On decompositions of multivariate functions**, *Math. Comput.*, (2010).

[3] M. Schmischke, **Interpretable Approximation of High-Dimensional Data based on the ANOVA Decomposition**, Thesis, *Universitätsverlag Chemnitz*, (2022).

Felix Bartel and Pascal Schröter
{felix.bartel, pascal.schroeter}@mathematik.tu-chemnitz.de
https://tu-chemnitz.de~{febar, pasca}

Preprint:
Learning Anisotropy parameters in ANOVA approximation
Software:
https://github.com/NFFT/ANOVAapprox.jl/tree/bandwidthDetection