Efficient recovery of non-periodic multivariate functions from few scattered samples

Nicolas Nagel

Joint work with Felix Bartel, Kai Lüttgen and Tino Ullrich

UNIVERSITY OF TECHNOLOGY
IN THE EUROPEAN CAPITAL OF CULTURE
CHEMNITZ

SampTA 2023 - Yale University
07/11/2023

Setting

- Given: $f : [-1,1]^d \to \mathbb{C}$ from some function space (Sobolev space)
- Goal: approximation of f via samples $f(\mathbf{x}^1), ..., f(\mathbf{x}^n)$
- Requirements:
  - Universality: $\mathbf{x}^1, ..., \mathbf{x}^n$ should work for the whole function space
  - Bound on the approximation error (depending on the dimension d, the number of samples n and the smoothness s of f)

It has been observed: if $\mathbf{x}^1, ..., \mathbf{x}^n$ follow a **Chebyshev distribution** and one uses **Chebyshev polynomials** one obtains near optimal approximations.

We give a theoretical explanation of this phenomenon.

Periodic functions

- $f(\mathbf{x}) \approx \sum_{\mathbf{k}\in\Lambda} c_{\mathbf{k}} \exp(\pi i \mathbf{k} \cdot \mathbf{x})$
  - $\Lambda \subset \mathbb{Z}^d$ finite set of frequencies
  - $c_{\mathbf{k}}$ determined from the samples $f(\mathbf{x}^i)$
- Questions:
  - Which frequencies Λ to use?
  - How to choose the sample nodes?
  - How to determine the coefficients $c_{\mathbf{k}}$?

Which frequencies Λ to use?

- If f is s-smooth with $s > 1/2$, i.e. $f \in H^s_{\text{mix}}(\mathbb{T}^d)$, then the Fourier coefficients

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \hat{f}_{\mathbf{k}} \exp(\pi i \mathbf{k}\cdot\mathbf{x})$$

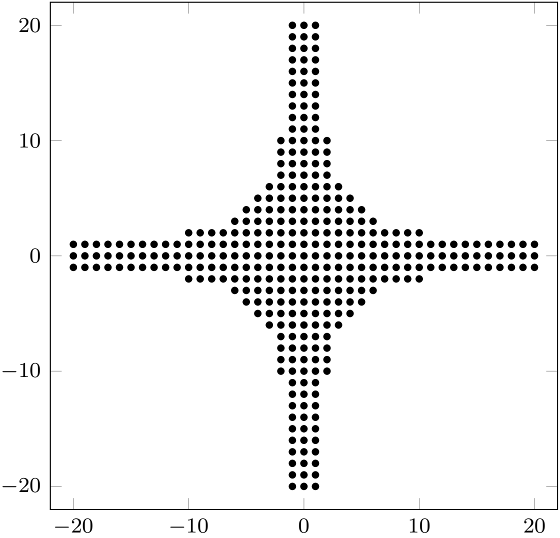
decay like

$$\underbrace{\sum_{\mathbf{k}\in\mathbb{Z}^d} |\hat{f}_{\mathbf{k}}|^2 \prod_{\ell=1}^{d} (1+|k_\ell|^2)^s}_{\asymp \|f\|^2_{H^s_{\text{mix}}}} < \infty$$

- Thus: frequencies near 0 are the most important ones
- Use **hyperbolic cross**

$$\Lambda = \{\mathbf{k}\in\mathbb{Z}^d : \prod_{\ell=1}^{d} \max\{1, |k_\ell|\} \le R\}$$

Hyperbolic cross $d = 2, R = 20, \#\Lambda = 345$



How to determine the coefficients $c_{\mathbf{k}}$?

- $m = \#\Lambda$ number of frequencies, N number of sample nodes,

$$\begin{bmatrix} \eta_{\mathbf{k}_1}(\mathbf{x}^1) & \cdots & \eta_{\mathbf{k}_m}(\mathbf{x}^1) \\ \vdots & & \vdots \\ \eta_{\mathbf{k}_1}(\mathbf{x}^N) & \cdots & \eta_{\mathbf{k}_m}(\mathbf{x}^N) \end{bmatrix} \begin{bmatrix} c_{\mathbf{k}_1} \\ \vdots \\ c_{\mathbf{k}_m} \end{bmatrix} \approx \begin{bmatrix} f(\mathbf{x}^1) \\ \vdots \\ f(\mathbf{x}^N) \end{bmatrix}$$

  with $\eta_{\mathbf{k}}(\mathbf{x}) = \prod_{\ell=1}^{d} e^{\pi i k_\ell x_\ell}$
- Approximate $f(\mathbf{x}) \approx \sum_{\mathbf{k}\in\Lambda} c_{\mathbf{k}} \eta_{\mathbf{k}}(\mathbf{x})$

How to choose the sample nodes?

- Symmetry of $\mathbb{T}^d$: $\mathbf{x}^1, ..., \mathbf{x}^n$ should be uniformly distributed
- Choose $\mathbf{x}^i, i = 1, ..., N$ uniformly at random
- To get a good approximation: need a logarithmic oversampling of $N = O(m \log m)$ due to randomness

Theorem (Krieg, M. Ullrich '21)

For $s > 1/2$ and using N samples, the above described procedure yields (with high probability) an approximation $\tilde{f}$ with

$$\|f - \tilde{f}\|_{L_2} \lesssim N^{-s} (\log N)^{ds} \|f\|_{H^s_{mix}(\mathbb{T}^d)}.$$

Subsampling

- Wish: $n = O(m)$ samples should be enough
- Indeed: Possible by the Kadison-Singer problem via Weaver's KS_2-conjecture

Kadison-Singer problem:

- Posed by Richard Kadison and Isadore Singer in 1959
- Rather abstract problem from functional analysis (motivated by quantum physics)
- Many equivalent formulations:

$$\begin{aligned}\text{Kadison-Singer} &\Leftrightarrow \text{Anderson's paving conjecture}\\ &\Leftrightarrow \text{Weaver's } KS_2\text{-conjecture}\\ &\Leftrightarrow \text{Feichtinger conjecture}\\ &\Leftrightarrow \text{Bourgain-Tzafriri conjecture}\end{aligned}$$

- Solved by Marcus, Spielman, Srivastava 2015

Consider $\mathbb{C}^m$ as a Hilbert space

- A **frame** is a sequence $\mathbf{u}^1, ..., \mathbf{u}^N \in \mathbb{C}^m$ such that

$$A\|\mathbf{w}\|_2^2 \leq \sum_{i=1}^{N} |\langle \mathbf{w}, \mathbf{u}^i \rangle|^2 \leq B\|\mathbf{w}\|_2^2$$

  for all $\mathbf{w} \in \mathbb{C}^m$, where A and B are constants (the **frame bounds**)
- B/A the **condition** of the frame

Theorem (Weaver's KS_2-conjecture; Marcus, Spielman, Srivastava '15)

If $\mathbf{u}^1, ..., \mathbf{u}^N \in \mathbb{C}^m$ *with* $\|\mathbf{u}^i\|_2 = 1$ *for all* i *and*

$$\sum_{i=1}^{N} |\langle \mathbf{w}, \mathbf{u}^i \rangle|^2 = 18\|\mathbf{w}\|_2^2$$

for all $\mathbf{w} \in \mathbb{C}^m$*, then one can partition* $S_1 \dot\cup S_2 = [N]$ *such that*

$$2\|\mathbf{w}\|_2^2 \leq \sum_{i \in S_j} |\langle \mathbf{w}, \mathbf{u}^i \rangle|^2 \leq 16\|\mathbf{w}\|_2^2$$

for all $\mathbf{w} \in \mathbb{C}^m$ *and* $j = 1, 2$*.*

Need: extract subframes from large frames with guaranties on their condition

- NITZAN, OLEVSKII, ULANOVSKII 2014: 1-tight frame $\mathbf{u}^1, ..., \mathbf{u}^N \in \mathbb{C}^m$ with $\|\mathbf{u}^i\|_2^2 = m/N$, there is a $J \subseteq [N]$ with $\#J = O(m)$ and resulting frame bounds $c\frac{m}{N}$ and $C\frac{m}{N}$
- TEMLYAKOV/N, SCHÄFER, T. ULLRICH 2020: Non-tight frames and only upper bound on the norms
- DOLBEAULT, KRIEG, M. ULLRICH 2022: Infinite-dimensional version in ℓ_2

Apply to $\mathbf{u}^i = [\eta_{\mathbf{k}}(\mathbf{x}^i)]_{\mathbf{k}\in\Lambda}$ to get well-conditioned subframe on $J \subseteq [N]$ of size $n = O(m)$ (down from $N = O(m \log m)$) with almost asymptotically equal approximation properties

Theorem (N, Schäfer, T. Ullrich '22)

For $s > 1/2$ and using n samples, the algorithm together with the subsampling step yields (with high probability) an approximation $\tilde{f}$ with

$$\|f - \tilde{f}\|_{L_2} \lesssim n^{-s} (\log n)^{(d-1)s+1/2} \|f\|_{H^s_{mix}(\mathbb{T}^d)}$$

- Problems:
  - Non-algorithmic: Kadison-Singer only gives existence of a subframe
  - The oversampling factor $n = bm$ might be huge (e.g. $b = 6000$)
- Polynomial time algorithm with small oversampling factor $b = 1 + \varepsilon$: Based on Batson, Spielman, Srivastava 2009 and developed further by Bartel, Schäfer, T. Ullrich 2023

Theorem (Batson, Spielman, Srivastava '09/Bartel, Schäfer, T. Ullrich '23)

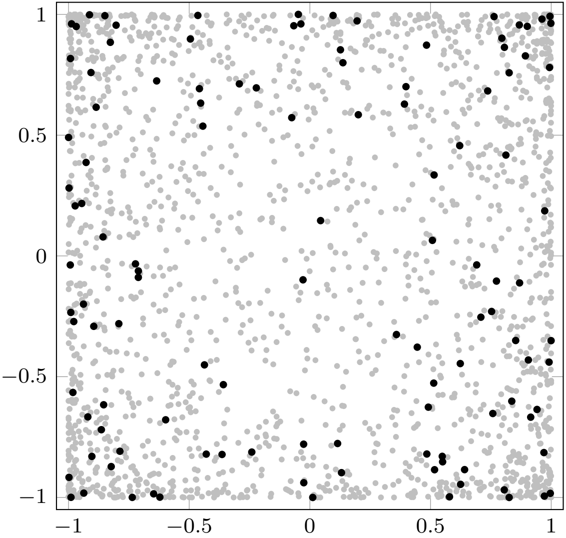
Let $\mathbf{u}^1, \ldots, \mathbf{u}^N \in \mathbb{C}^m$ *(arbitrary), choose* $b > 1 + \frac{1}{m}$ *and assume* $N \geq bm$. *There is a polynomial time algorithm to construct a* $J \subseteq [N]$ *with* $\#J \leq \lceil bm \rceil$ *and*

$$\frac{1}{N} \sum_{i=1}^{N} |\langle \mathbf{w}, \mathbf{u}^i \rangle|^2 \leq 89 \frac{(b+1)^2}{(b-1)^3} \cdot \frac{1}{m} \sum_{j \in J} |\langle \mathbf{w}, \mathbf{u}^j \rangle|^2.$$

Note: only lower bound, algorithm with guaranty on the upper bound unknown
Still: resulting sample nodes $\mathbf{x}^1, \ldots, \mathbf{x}^n$ random, their quality is measured by the condition of the matrix

$$\begin{bmatrix} \eta_{\mathbf{k}_1}(\mathbf{x}^1) & \cdots & \eta_{\mathbf{k}_m}(\mathbf{x}^1) \\ \vdots & & \vdots \\ \eta_{\mathbf{k}_1}(\mathbf{x}^n) & \cdots & \eta_{\mathbf{k}_m}(\mathbf{x}^n) \end{bmatrix}$$

Subsampling ($d = 2, R = 20, m = 107, N = 2000, n = 117$)



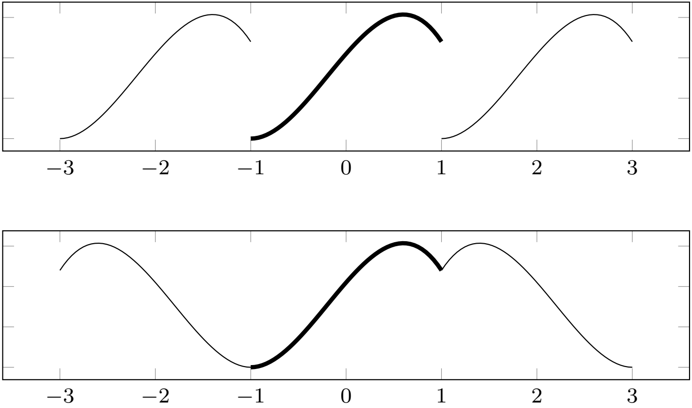
Final algorithm (periodic):

- Choose size of hyperbolic cross $\Lambda = \{\mathbf{k} \in \mathbb{Z}^d : \prod_{\ell=1}^d \max\{1, |k_\ell|\} \leq R\}$, $m = \#\Lambda$
- Set $N = \lceil 4m \log m \rceil$
- Choose N nodes $\mathbf{x}^i \in [-1, 1]^d$ uniformly at random
- Subsampling gives nodes $\{\mathbf{x}^j : j \in J\}$ with $n = \#J \leq \lceil 1.1m \rceil$ using the basis functions $\eta_{\mathbf{k}}(\mathbf{x}) = \prod_{\ell=1}^d e^{\pi i k_\ell x_\ell}$ (works for all of $H^s_{\text{mix}}(\mathbb{T}^d)$)
- Determine the coefficients $c_{\mathbf{k}}$ from the $\mathbf{x}^j, j \in J$ via the least-squares system with the basis functions $\eta_{\mathbf{k}}(\mathbf{x})$

Non-periodic functions

Applying the above procedure to a more general function $f : [-1, 1]^d \to \mathbb{C}$ treats f like a periodized version on $\mathbb{R}^d$

$\Rightarrow$ may introduce non-regularities



To apply the algorithm for periodic functions, we need to periodize f in a way that preserves regularity

- Periodic extension: May introduce discontinuities
- Tent transform: Preserves continuity, might destroy smoothness (kinks)

We will use a **cosine composition** $T_{\cos}$ defined by

$$(T_{\cos} f)(x_1, ..., x_d) = f(\cos \pi x_1, ..., \cos \pi x_d)$$

Theorem (Bartel, Lüttgen, N, T. Ullrich)

The operator $T_{\cos}$ is continuous as

$$T_{\cos} : H^s_{mix}([-1,1]^d) \to H^s_{mix}(\mathbb{T}^d)$$

for $s > 1/2$.

More general versions over Besov spaces are possible

Strategy: Approximate $T_{\cos} f$ with the Fourier basis and undo the periodization

	$f(\cos \pi x_1, \ldots, \cos \pi x_d)$	$f(x_1, \ldots, x_d)$
Sample nodes	$\mathbf{x}^i \sim \mathcal{U}[-1,1]^d$	$\mathbf{x}^i = \cos(\pi \mathbf{U}^i), \mathbf{U}^i \sim \mathcal{U}[-1,1]^d$ $\mathrm{d}\varrho(\mathbf{x}) = \prod_{\ell=1}^d \left(\pi\sqrt{1-x_\ell^2}\right)^{-1} \mathrm{d}\mathbf{x}$
Basis functions	$\prod_{\ell=1}^d \cos(\pi k_\ell x_\ell)$	$\prod_{\ell=1}^d \cos(k_\ell \arccos x_\ell)$

- $\mathbf{x}^i$ Chebyshev distributed
- $\eta_{\mathbf{k}}(\mathbf{x}) = \prod_{\ell=1}^d T_{k_\ell}(x_\ell), \mathbf{k} \in \mathbb{N}_0^d$ with $T_k(x) = \sqrt{2}^{\min\{1,k\}} \cos(k \arccos x)$ ($L_2(\varrho)$-normalized Chebyshev polynomials)

Final algorithm (non-periodic):

- Choose size of hyperbolic cross $\Lambda = \{\mathbf{k} \in \mathbb{N}_0^d : \prod_{\ell=1}^d \max\{1, k_\ell\} \leq R\}$ (in the positive orthant), $m = \#\Lambda$
- Set $N = \lceil 4m \log m \rceil$
- Choose N nodes $\mathbf{x}^i \in [-1,1]^d$ Chebyshev distributed
- Subsampling gives nodes $\{\mathbf{x}^j : j \in J\}$ with $n = \#J \leq \lceil 1.1m \rceil$ using the basis functions $\eta_{\mathbf{k}}(\mathbf{x}) = \prod_{\ell=1}^d T_{k_\ell}(x_\ell)$ (works for all of $H^s_{\text{mix}}([-1,1]^d)$)
- Determine the coefficients $c_{\mathbf{k}}$ from the $\mathbf{x}^j, j \in J$ via the least-squares system with the basis functions $\eta_{\mathbf{k}}(\mathbf{x})$

Theorem (Bartel, Lüttgen, N, T. Ullrich)

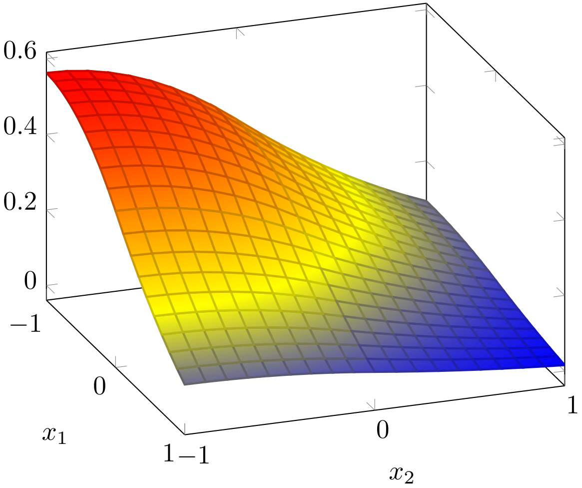
For $s > 1/2$ and using n samples, the above algorithm yields (with high probability) an approximation $\tilde{f}$ with

$$\|f - \tilde{f}\|_{L_2(\varrho)} \lesssim n^{-s} (\log n)^{(d-1)s+1/2} \|f\|_{H^s_{mix}([-1,1]^d)}.$$

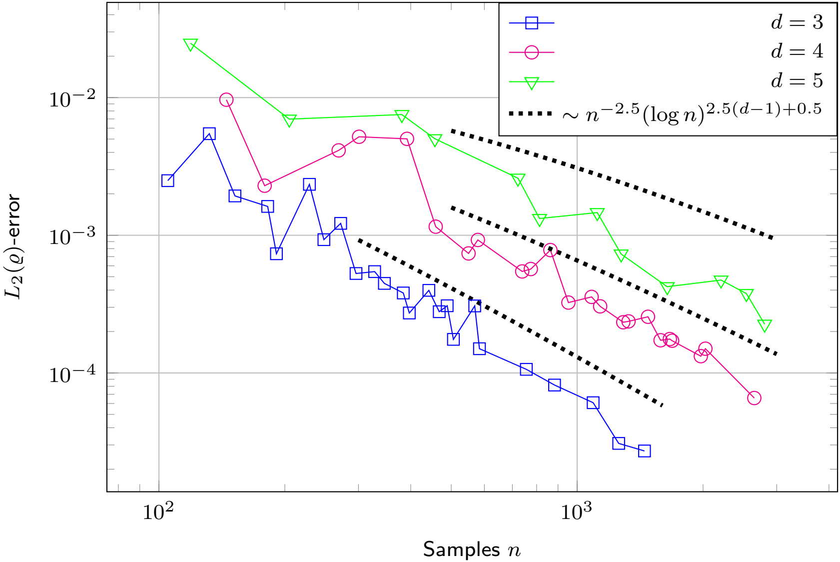
- $L_2(\varrho)$-norm stronger than usual (Lebesgue) L_2-norm
- $\sqrt{\log n}$ could be removed, but no know (efficient) subsamplings algorithm

Numerical experiment

Test function: Tensored cutout of a quadratic B-spline (smoothness $s = 2.5$)



Approximation error for an $f \in H^{2.5-\varepsilon}_{\text{mix}}([-1,1]^d)$



The end

- "Deterministic Kadison-Singer"? (remove the $\sqrt{\log}$-factor)
- Deterministic constructions for good samples nodes?

Thank you for your attention!