Sampling recovery and sharp norm estimates of projection operators

Kateryna Pozharska

Institute of Mathematics of NAS of Ukraine, Kyiv, Ukraine
Faculty of Mathematics, Chemnitz University of Technology, Germany

MCQMC, Waterloo

UNIVERSITY OF TECHNOLOGY
IN THE EUROPEAN CAPITAL OF CULTURE
CHEMNITZ

The plan

- relations between the Sampling, Gelfand and Kolmogorov numbers, optimality of information: **linear vs standard**;
- discretization results;
- bounds for the norms of projection operators.

Joint work with **David Krieg, Mario Ullrich and Tino Ullrich**

- *Sampling projections in the uniform norm*. arXiv: 2401.02220, 2024.
- *Sampling recovery in L_2 and other norms*. arXiv: 2305.07539, 2023.

Information $\longrightarrow$ Recovery algorithm $\longrightarrow$ Reconstructed function

$N(f)$ $\quad$ ψ $\quad$ $f \approx \psi \circ N(f)$

Information $N(f)$:

- **linear**
  - Fourier coefficients
  - **function values**
- nonlinear

Recovery algorithm ψ:

- **linear**
- nonlinear
  - **non-adaptive**
  - adaptive

$$\left\|f - \psi \circ N(f)\right\|_{\boldsymbol{G}} \to \min$$

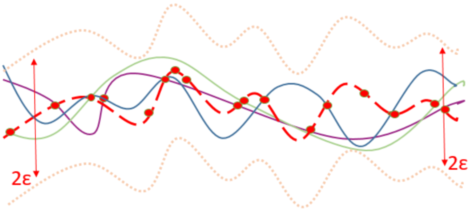
Recent overview:

D. Krieg, E. Novak, M. Ullrich, On the power of adaption and randomization. arXiv: 2406.07108, 2024.

Given: $f(x_1), ..., f(x_n)$ for functions $f \in F \subset G$.

Estimate: the **n-th linear sampling number** of F in G

$$g_n^{\mathrm{lin}}(F,G) := \inf_{\substack{x_1,\ldots,x_n \in D \\ \varphi_1,\ldots,\varphi_n \in G}} \sup_{f \in F} \left\| f - \sum_{i=1}^{n} f(x_i)\, \varphi_i \right\|_G .$$



Red line – target, other lines – approximants

Benchmark quantities

- **n-th Gelfand number** $c_n(F,G)$: **any** algorithm, **linear** information
- **n-th approximation number** $a_n(F,G)$: algorithm and information - **linear**

- **n-th Kolmogorov number**

$$d_n(F,G) := \inf_{\substack{\boldsymbol{V_n \subset G} \\ \mathbf{dim}(\boldsymbol{V_n})=\boldsymbol{n}}} \sup_{f\in F} \inf_{g\in V_n} \|f-g\|_G$$

General relations

Gelfand n.		approximation n.		sampling n.
$c_n(F,G)$	$\leq$	$a_n(F,G)$	$\leq$	$g_n^{\mathrm{lin}}(F,G)$
lin. inf.		lin. inf.		std. inf.
arb. alg.		lin. alg.		lin. alg.

General relations

Gelfand n.		approximation n.		sampling n.
$c_n(F,G)$	$\le$	$a_n(F,G)$	$\le$	$g_n^{\mathrm{lin}}(F,G)$
lin. inf. arb. alg.		lin. inf. lin. alg.		std. inf. lin. alg.

$$g_n^{\mathrm{lin}}(F,G) \le K(n) \cdot c_n(F,G) \qquad \longrightarrow \qquad \text{find the order of } K(n)$$

General relations

Gelfand n.		approximation n.		sampling n.
$c_n(F,G)$	$\leq$	$a_n(F,G)$	$\leq$	$g_n^{\mathrm{lin}}(F,G)$
lin. inf.		lin. inf.		std. inf.
arb. alg.		lin. alg.		lin. alg.

$g_n^{\mathrm{lin}}(F,G) \leq K(n) \cdot c_n(F,G)$ $\longrightarrow$ find the order of $K(n)$

For every **convex and balanced** set F, it holds

$$d_n(F, B(D)) \leq c_n(F, B(D))$$

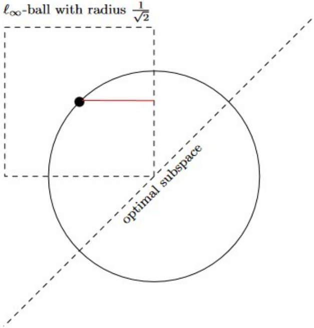
The goal: find $K(n)$ such that

$$g_n^{\mathrm{lin}}(F, B(D)) \leq K(n) \cdot d_n(F, B(D)) \leq K(n) \cdot c_n(F, B(D))$$

Kolmogorov vs Gelfand numbers

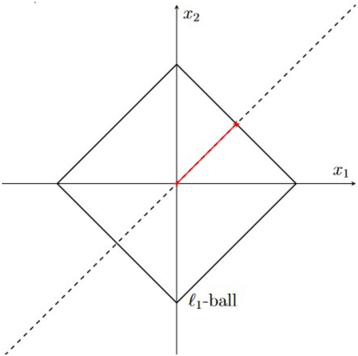
Kolmogorov numbers

$d_1(B_2, \ell_\infty)$



Gelfand numbers

$c_1(B_1, \ell_2)$



$$g_n^{\mathrm{lin}}(F, B(D)) \;\le\; K(n) \cdot d_n(F, B(D)) \;\le\; K(n) \cdot c_n(F, B(D))$$

How large are the gaps? What is the order of $\boldsymbol{K(n)}$?

Novak '88

$$g_{\boldsymbol{n}}^{\mathrm{lin}}(F, B(D)) \;\le\; \boldsymbol{(1+n)}\, d_n(F, B(D))$$

The factor $(1+n)$ cannot be improved even for Hilbert spaces.

Kashin/Konyagin/Temlyakov '21

$$g_{9n}^{\mathrm{lin}}(F, B(D)) \;\le\; 5\, d_n(F, B(D))$$

Novak ’88

$$g_{\boldsymbol{n}}^{\text{lin}}(F, B(D)) \,\leq\, \boldsymbol{(1+n)}\, d_n(F, B(D))$$

The factor $(1+n)$ cannot be improved even for Hilbert spaces.

With an oversampling bn, $b > 1$ – YES – to $\sqrt{n}$!

Kashin/Konyagin/Temlyakov ’21

$$g_{9^n}^{\text{lin}}(F, B(D)) \,\leq\, 5\, d_n(F, B(D))$$

Theorem [KPUU '23]

For **any set** D, **any** $F \subset B(D)$, and all $n \in \mathbb{N}$, it holds

$$g_{2n}^{\mathrm{lin}}(F, B(D)) \;\leq\; 58\,\sqrt{n}\; d_n(F, B(D)).$$

For **convex and balanced** F, it holds

$$g_{2n}^{\mathrm{lin}}(F, B(D)) \;\leq\; 58\,\sqrt{n}\; c_n(F, B(D)).$$

The factor $\sqrt{n}$ in general is sharp!

Note: Sharp bound without $\sqrt{n}$ for Hilbert spaces satisfying certain conditions.

Example. Sharpness of the results

Let $W_p^s := W_p^s([0,1])$, where $\|f\|_{W_p^s} := \|f\|_p + \|f^{(s)}\|_p \leq 1$, $s > 1$.

Both bounds are sharp:

$$g_n^{\mathrm{lin}}(W_1^s, B(D)) \asymp n^{-s+1}, \qquad c_n(W_1^s, B(D)) \asymp d_n(W_1^s, B(D)) \asymp n^{-s+1/2}$$

The first bound is sometimes better:

$$g_n^{\mathrm{lin}}(W_2^s, B(D)) \asymp c_n(W_2^s, B(D)) \asymp n^{-s+1/2}, \qquad d_n(W_2^s, B(D)) \asymp n^{-s}.$$

[Maiorov '76; Kashin '77; Pinkus '85; Novak '06]

Improvement in the Hilbert space setting

Let H be a RKHS on a set D with a **bounded** kernel

$$K(x,y) = \sum_{k=1}^{\infty} \sigma_k^2\, b_k(x)\overline{b_k(y)}, \tag{1}$$

where $\boldsymbol{\sigma_1 \geq \sigma_2 \geq \ldots \geq 0}$, $\boldsymbol{\{b_k\}_{k=1}^{\infty}}$ **is orthonormal in** $\boldsymbol{L_2(\mu)}$.

$$\boldsymbol{V_n = \mathrm{span}\{b_1, \ldots, b_n\}, \quad \Lambda_n^2 := \sup_{f \in V_n \setminus \{0\}} \|f\|_\infty^2 / \|f\|_2^2 = \Big\| \sum_{k=1}^{n} |b_k|^2 \Big\|_\infty}$$

Theorem [KPUU '23]

There is an absolute constant $c > 0$, such that the following holds for a RKHS $H \subset L_2(\mu)$ with the kernel (1). For all $n \in \mathbb{N}$

$$g_{cn}^{\mathrm{lin}}(H, B(D)) \leq \sqrt{\frac{866\,\Lambda_n^2}{n} \sum_{k > \lfloor n/2 \rfloor} \sigma_k^2} + \sqrt{2 \sum_{k > \lfloor n/4 \rfloor} \frac{\Lambda_{4k}^2 \sigma_k^2}{k}}.$$

Sharp bound in the Hilbert space setting

(independently obtained by [Geng/Wang '23] for the trigonometric system)

If $\Lambda_n \lesssim \sqrt{n}$, i.e.,
$$\Big\| \tfrac{1}{n} \textstyle\sum_{k=1}^{n} |b_k|^2 \Big\|_\infty \leq B, \quad B > 0,$$
then
$$g_{cn}^{\mathrm{lin}}(H, B(D))^2 \leq bB \cdot \sum_{k>n} \sigma_k^2, \qquad \sigma_k = d_k(H, L_2).$$

If μ is a **finite** measure, then

$$g_{cn}^{\mathrm{lin}}(H, B(D)) \leq \sqrt{bB \cdot \mu(D)} \cdot c_n(H, B(D))$$

by the bound $\sqrt{\sum_{k>n} \sigma_k^2} \leq \sqrt{\mu(D)} \cdot c_n(H, B(D))$ from
[Osipenko/Parfenov '95; Kuo/Wasilkowski/Wozniakowski '08; Cobos/Kühn/Sickel '16]

Linear sampling algorithms are (up to constants) **as powerful as** arbitrary non-linear algorithms using general linear measurements.

Hilbert spaces $\boldsymbol{H^w}$ on the d-torus $\mathbb{T}^d = [0, 2\pi]^d$

$$f \in L_1: \qquad \|f\|_{H^w}^2 = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} (w(\boldsymbol{k}))^2 |f_{\boldsymbol{k}}|^2 < \infty,$$

where $w(\boldsymbol{k}) > 0$, $\boldsymbol{k} \in \mathbb{Z}^d$, $\sum_{\boldsymbol{k}\in\mathbb{Z}^d} (w(\boldsymbol{k}))^{-2} < \infty$,

$f_{\boldsymbol{k}}$ are the Fourier coefficients w.r.t. the trigonometric system.

If

$$w(\boldsymbol{k}) = \prod_{j=1}^{d} (1 + |k_j|)^s\,, \quad s > 1/2,$$

the orders coincide with those for sparse grids (see $[\text{Temlyakov } '93]$)

$$g_n^{\text{lin}}(H^s_{\text{mix}}, B(\mathbb{T}^d)) \asymp_{s,d} n^{-s+1/2} (\log n)^{s(d-1)}\,.$$

Nikol'skii-Besov $\mathbf{B}^r_{p,\infty} \equiv \mathbf{H}^r_p$ and Sobolev $\mathbf{W}^r_p$ mixed smoothness classes on the torus $\mathbb{T}^d$

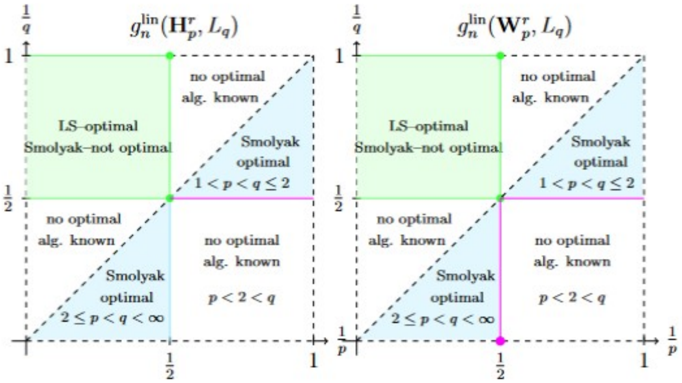
$$\sup_{f\in F} \|f - A_m f\|_q \lesssim m^{-(r-t)}(\log m)^{(d-1)(r-t+\alpha)},$$

$r > \max\{1/2, 1/p\}, \quad t := (1/p - 1/2)_+ + (1/2 - 1/q)_+,$

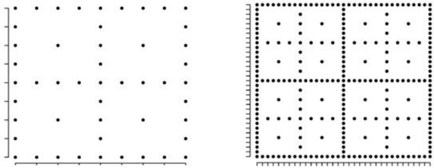
$$\alpha = \begin{cases} 0 & F = \mathbf{W}^r_p, 1 \le q < \infty; \\ 1/2 & F = \mathbf{H}^r_p, 1 \le q < \infty;\ F = \mathbf{W}^r_p, q = \infty; \\ 1 & F = \mathbf{H}^r_p, q = \infty. \end{cases}$$

A_m is a plain least-squares algorithm based on subsampled random points.

Nikol'skii-Besov $\mathbf{B}^r_{p,\infty} \equiv \mathbf{H}^r_p$ and Sobolev $\mathbf{W}^r_p$ mixed smoothness classes on the torus $\mathbb{T}^d$

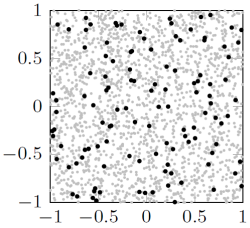


Smolyak points



Bounds for the respective worst case error using the Smolyak algorithm by [Dũng '91; Temlyakov '93; Dũng/T.Ullrich '15; / Byrenheid/T.Ullrich '17]

Subsampled random points



We also apply the results to:

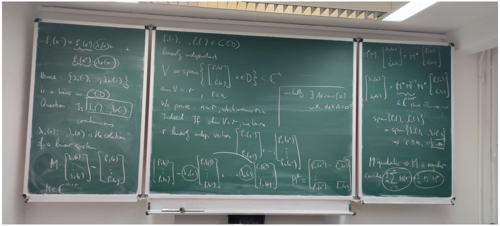
- Wiener-type and Korobov spaces on the torus $\mathbb{T}^d$
- Sobolev spaces on manifolds
- approximation in L_∞- and H^1-norm for a weigthed Sobolev space without uniformly bounded basis

Discretization of the uniform norm

Theorem [KPUU '23]

Let D be a set and $V_n \subset B(D)$, $\dim V_n = n$. Then there exist $2n$ points $x_1, \dots, x_{2n} \in D$ such that

$$\forall f \in V_n \qquad \|f\|_\infty \leq 58\sqrt{n}\Big(\frac{1}{2n}\sum_{k=1}^{2n}|f(x_k)|^2\Big)^{1/2} \leq 58\sqrt{n}\max_{k=1,\dots,2n}|f(x_k)|$$



Part of the proof.

Based on $[\text{Kiefer/Wolfowitz '60}]$ + constructive subsampling.

Algorithm for continuous on a compact domain functions

- Given n linearly independent complex-valued functions $f_1, \ldots, f_n \in C(D)$, for a fixed $x \in D$ solve

$$A := \max_{(x_k)_{k=1}^N, (\lambda_k)_{k=1}^N} \det(B^* B),$$

$$B := \begin{pmatrix} \sqrt{\lambda_1} f_1(x_1) & \sqrt{\lambda_1} f_2(x_1) & \cdots & \sqrt{\lambda_1} f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\lambda_N} f_1(x_N) & \sqrt{\lambda_N} f_2(x_N) & \cdots & \sqrt{\lambda_N} f_n(x_N) \end{pmatrix}$$

with $N = n^2 + 1$ nodes $(x_k)_{k=1}^N$ and weights $(\lambda_k)_{k=1}^N$.

- Write the singular value decomposition $A = VDV^*$, $A^{-1/2} = VD^{-1/2}V^*$.
- Put

$$G(x) := A^{-1/2} F(x), \qquad F := (f_1, \ldots, f_n)^\top .$$

The components of $G(x) = (g_1, \ldots, g_n)^\top$ form an ONB in $V_n = \operatorname{span}\{f_1, \ldots, f_n\}$ w.r.t. the atomic measure $\varrho = \sum_{i=1}^N \lambda_i \delta_{x_i}$, and it holds

$$\Lambda_n = \sup_{f \in V_n \setminus \{0\}} \frac{\|f\|_\infty}{\|f\|_{L_2(\varrho)}} = \sup_{x \in D} \left(\sum_{i=1}^n |g_i(x)|^2 \right)^{1/2} = \sqrt{n}.$$

Discretization of the uniform norm. The best-known results.

Kashin/Temlyakov '18; Dai/Prymak/Temlyakov/Tikhonov '19

For the case of trigonometric polynomials $t \in \mathcal{T}(Q), Q \subset \mathbb{Z}^d, |Q| = n$
$\exists C_1, C_2 > 0$ such that for a set of points X, $|X| \geq C_1 n$:

$$\|t\|_\infty \leq C_2 \sqrt{n} \max_{x \in X} |t(x)|.$$

Dai/Prymak/Temlyakov/Tikhonov '19; Kashin/Konyagin/Temlyakov '21

If any $f \in V_n \subset C(D)$ satisfies the Nikol'skii inequality

$$\|f\|_\infty \leq H(n) \|f\|_2,$$

then $\exists C_1, C_2 > 0$ such that there exists a set of points $X \subset D$, $|X| \leq C_1 n$:

$$\|f\|_\infty \leq C_2 H(n) \max_{x \in X} |f(x)|.$$

Discretization of the uniform norm. The best-known results.

Kashin/Temlyakov '18; Dai/Prymak/Temlyakov/Tikhonov '19; Kashin/Konyagin/Temlyakov '21

Let $V_n \subset C(D)$. There exists a set of points X, $|X| \le 9^n$ such that for all $f \in V_n$ it holds

$$\|f\|_\infty \le 2 \max_{x \in X} |f(x)|.$$

Novak '88

Let $V_n \subset B(D)$, $\varepsilon > 0$. There exists a set of points X, $|X| = n$ such that for all $f \in V_n$ it holds

$$\|f\|_\infty \le (n + \varepsilon) \max_{x \in X} |f(x)|.$$

Theorem [KPUU '23]

Let D be a set and $V_n \subset B(D)$, $\dim V_n = n$. Then there are $2n$ points $x_i \in D$ and functions $\varphi_i \in V_n$, such that $P\colon B(D) \to V_n$ with

$$Pf = \sum_{i=1}^{2n} f(x_i)\,\varphi_i \quad \text{is a projection with} \quad \|P\| \le 58\sqrt{n}.$$

Theorem is **sharp** in the following sense:

- Using $m = \mathcal{O}(n)$ samples, the norm bound $C\sqrt{n}$ cannot be replaced with a lower-order term.
- Using only $m = n$ samples, the norm bound $C\sqrt{n}$ has to be replaced by a linear term in n (*Novak's/Auerbach's lemma*).

+ Connection to the famous Kadets-Snobar theorem (1971).

Theorem [Bartel, Kämmerer, Schäfer, P., Ullrich ’24]

Let μ be a Borel probability measure on a compact topological space D, $V_n \subset C(D)$. Then there exists an $N \leq \dim(\operatorname{span}\{f \cdot \overline{g} \colon f, g \in V_n\}) \leq n^2$ and $N+1$ points $x_1, \ldots, x_{N+1} \in D$ together with non-negative weights satisfying $\mu_1 + \cdots + \mu_{N+1} = 1$ such that for all $f \in V_n$ it holds

$$\int_D |f(x)|^2 d\mu(x) = \sum_{i=1}^{N+1} \mu_i |f(x_i)|^2 \,.$$

$\boldsymbol{L_2}$-**Marcinkiewicz-Zygmund inequalities**, see:
[Mhaskar, Narcowich, Ward ’01], [Nuyens, Cools ’06], [Keiner, Kunis, Potts ’07], [Filbir, Mhaskar ’11], [Müller-Gronbach, Novak, Ritter ’12], [Kämmerer, Potts, Volkmer ’15], [Temlyakov ’18], [Trefethen ’19], [Gröchenig ’20], [Filbir, Hielscher, Jahn, T. Ullrich ’24], ...

+ Connection to spherical designs
+ Connection to Parseval-scalable frames ...

Theorem [Bartel, Kämmerer, Schäfer, P., Ullrich '24]

Let μ be a Borel probability measure on a compact topological space D, $V_n \subset C(D)$. Then there exists an $N \leq \dim(\mathrm{span}\{f \cdot \overline{g} \colon f, g \in V_n\}) \leq n^2$ and $N+1$ points $x_1, \dots, x_{N+1} \in D$ together with non-negative weights satisfying $\mu_1 + \cdots + \mu_{N+1} = 1$ such that for all $f \in V_n$ it holds

$$\int_D |f(x)|^2 d\mu(x) = \sum_{i=1}^{N+1} \mu_i |f(x_i)|^2 \,.$$

$\boldsymbol{L_2}$**-Marcinkiewicz-Zygmund inequalities**, see:
[Mhaskar, Narcowich, Ward '01], [Nuyens, Cools '06], [Keiner, Kunis, Potts '07], [Filbir, Mhaskar '11], [Müller-Gronbach, Novak, Ritter '12], [Kämmerer, Potts, Volkmer '15], [Temlyakov '18], [Trefethen '19], [Gröchenig '20], [Filbir, Hielscher, Jahn, T. Ullrich '24], ...

+ Connection to spherical designs
+ Connection to Parseval-scalable frames ...

Thank you for your attention!