On the reconstruction of functions from values at subsampled quadrature points

Felix Bartel   Lutz Kämmerer   Daniel Potts   Tino Ullrich

Approximation and geometry in high dimensions 2022

12th of October 2022

Mathematik!
TU Chemnitz

Function recovery from samples

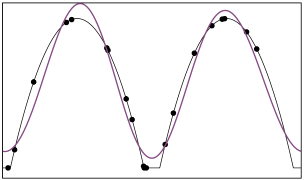
Given:

- $L_2 = L_2(D, \nu)$
- reproducing kernel Hilbert space $H(K) \hookrightarrow L_2$

Goal:

- find good points $\boldsymbol{X} = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^n\} \subset D$ and a sampling recovery operator $S_{\boldsymbol{X}} \colon H(K) \to L_2$ with small **worst-case error**

$$\sup_{\|f\|_{H(K)} \le 1} \|f - S_{\boldsymbol{X}} f\|_{L_2}$$



$f(\boldsymbol{x})$ ——

$(\boldsymbol{x}^i, f(\boldsymbol{x}^i))$ •

$(S_{\boldsymbol{X}} f)(\boldsymbol{x})$ ——

Reproducing kernel Hilbert space

- assume finite trace $\int_D K(\boldsymbol{x},\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{x}) < \infty$
- embedding $\mathrm{Id}_{K,\nu}\colon H(K) \hookrightarrow L_2$ has the representation

$$\mathrm{Id}_{K,\nu}(f) = \sum_{k=0}^{\infty} \sigma_k \langle f, e_k\rangle_{H(K)} \eta_k$$

with $e_k = \sigma_k \eta_k$ and

  - **singular values** $\sigma_0 \geq \sigma_1 \geq \cdots \geq 0$
  - **right singular functions** $e_0, e_1, \dots$ forming an ONS in $H(K)$
  - **left singular functions** $\eta_0, \eta_1, \dots$ forming an ONS in L_2

Reproducing kernel Hilbert space

- assume finite trace $\int_D K(\boldsymbol{x}, \boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{x}) < \infty$
- embedding $\mathrm{Id}_{K,\nu} \colon H(K) \hookrightarrow L_2$ has the representation

$$\mathrm{Id}_{K,\nu}(f) = \sum_{k=0}^{\infty} \sigma_k \langle f, e_k \rangle_{H(K)} \eta_k$$

  with $e_k = \sigma_k \eta_k$ and
  - **singular values** $\sigma_0 \geq \sigma_1 \geq \cdots \geq 0$
  - **right singular functions** $e_0, e_1, \dots$ forming an ONS in $H(K)$
  - **left singular functions** $\eta_0, \eta_1, \dots$ forming an ONS in L_2
- Kolmogorov width

$$d_m(H(K)) := \inf_{\substack{\ell_0, \dots, \ell_{m-1} \colon H(K) \to \mathbb{C} \\ \varphi_0, \dots, \varphi_{m-1} \in L_2}} \sup_{f \in H(K)} \Big\| f - \sum_{k=0}^{m-1} \ell_k(f) \varphi_k \Big\|_{L_2}$$

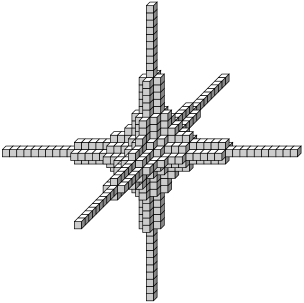
$$= \sup_{f \in H(K)} \Big\| f - \sum_{k=0}^{m-1} \langle f, \eta_k \rangle_{L_2} \eta_k \Big\|_{L_2} = \sigma_m$$

Sobolev spaces with dominating mixed smoothness on $\mathbb{T}^d$

- $H^s_{\text{mix}}(\mathbb{T}^d) = \{f \in L_2 : \|f\|_{H^s_{\text{mix}}} < \infty\}$ $(s > 1/2)$ with

$$\langle f, g\rangle_{H^s_{\text{mix}}} := \sum_{\boldsymbol{j}\in\{0,s\}^d} \langle D^{(\boldsymbol{j})} f, D^{(\boldsymbol{j})} g\rangle_{L_2}$$

- singular functions $\eta_{\boldsymbol{k}} = \exp(2\pi \mathrm{i}\langle \boldsymbol{k}, \cdot\rangle)$
- singular values $\sigma_{\boldsymbol{k}} = \prod_{j=1}^{d} (1 + (2\pi|k_j|)^{2s})^{-1/2}$
- Kolmogorov width $d_m(H^s_{\text{mix}}) = m^{-s}(\log m)^{(d-1)s}$



I_R^{hc}

Least squares

- assume $n \geq m$
- ansatz $f(\boldsymbol{x}) = \sum_{k=0}^{m-1} c_k \eta_k(\boldsymbol{x})$ with $\boldsymbol{c} = (c_0, \ldots, c_{m-1})^\mathsf{T}$ solving

$$\begin{pmatrix} \eta_0(\boldsymbol{x}^1) & \cdots & \eta_{m-1}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_0(\boldsymbol{x}^n) & \cdots & \eta_{m-1}(\boldsymbol{x}^n) \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_{m-1} \end{pmatrix} = \begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^n) \end{pmatrix}$$

Least squares

- assume $n \geq m$
- ansatz $f(\boldsymbol{x}) = \sum_{k=0}^{m-1} c_k \eta_k(\boldsymbol{x})$ with $\boldsymbol{c} = (c_0, \ldots, c_{m-1})^\mathsf{T}$ solving

$$\left\| \begin{pmatrix} \eta_0(\boldsymbol{x}^1) & \cdots & \eta_{m-1}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_0(\boldsymbol{x}^n) & \cdots & \eta_{m-1}(\boldsymbol{x}^n) \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_{m-1} \end{pmatrix} - \begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^n) \end{pmatrix} \right\|_{\boldsymbol{W}}^2 \to \min$$

for a weight matrix $\boldsymbol{W} = \mathrm{diag}(\omega_1, \ldots, \omega_n)$

Least squares

- assume $n \geq m$
- ansatz $f(\boldsymbol{x}) = \sum_{k=0}^{m-1} c_k \eta_k(\boldsymbol{x})$ with $\boldsymbol{c} = (c_0, \ldots, c_{m-1})^\mathsf{T}$ solving

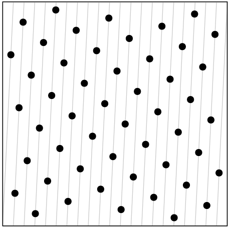
$$\Bigg\| \underbrace{\begin{pmatrix} \eta_0(\boldsymbol{x}^1) & \cdots & \eta_{m-1}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_0(\boldsymbol{x}^n) & \cdots & \eta_{m-1}(\boldsymbol{x}^n) \end{pmatrix}}_{=:\boldsymbol{L}} \begin{pmatrix} c_0 \\ \vdots \\ c_{m-1} \end{pmatrix} - \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^n) \end{pmatrix}}_{=:\boldsymbol{f}} \Bigg\|_{\boldsymbol{W}}^2 \to \min$$

  for a weight matrix $\boldsymbol{W} = \mathrm{diag}(\omega_1, \ldots, \omega_n)$
- the solution is given by

$$\boldsymbol{c} = (\boldsymbol{L}^* \boldsymbol{W} \boldsymbol{L})^{-1} \boldsymbol{L}^* \boldsymbol{W} \boldsymbol{f} \quad \text{and} \quad (S_{\boldsymbol{X}} f)(\boldsymbol{x}) = \sum_{k=0}^{m-1} c_k \eta_k(\boldsymbol{x})$$

Structured points: rank-1 lattice

- $D = \mathbb{T}^d$, $\{\eta_0, \dots, \eta_{m-1}\} = \{\exp(2\pi\mathrm{i}\langle \boldsymbol{k}, \cdot\rangle\}_{\boldsymbol{k}\in I}$
- $\boldsymbol{X} = \Lambda(\boldsymbol{z}, n) = \left\{\frac{i}{n}\boldsymbol{z} \bmod \boldsymbol{1}, i = 0, \dots, n-1\right\}$
- [Nuyens '07], [Kämmerer, Potts, Volkmer '15]:
  - given I, there exist algorithms that find $\boldsymbol{z}$ and n such that $\boldsymbol{L}^*\boldsymbol{W}\boldsymbol{L} = \boldsymbol{I}$
  - multiplication with $\boldsymbol{L}$ can be carried out with the LFFT in $\mathcal{O}(n \log n)$



$z = (1, 21)^\mathsf{T}, n = 55$

Theorem

[Byrenheid, Kämmerer, T. Ullrich, Volkmer '17]

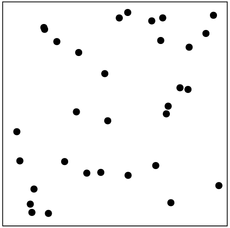
$$n^{-s/2} \lesssim \sup_{\|f\|_{H^s_{\text{mix}}} \le 1} \|f - S_{\boldsymbol{X}} f\|_{L_2} \lesssim n^{-s/2} (\log n)^{\frac{d-2}{2}s + \frac{d-1}{2}}$$

Random points

- $\boldsymbol{X} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n\}$ drawn randomly i.i.d. w.r.t. Lebesgue measure such that

$$n \sim m \log m$$

- no fast matrix-vector multiplication due to lack of structure

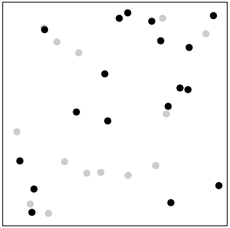


Theorem [M. Ullrich, Krieg '19], [Kämmerer, T. Ullrich, Volkmer '19]

$$n^{-s}(\log n)^{(d-1)s} \lesssim \sup_{\|f\|_{H^s_{\text{mix}}} \leq 1} \|f - S_{\boldsymbol{X}} f\|_{L_2} \lesssim n^{-s}(\log n)^{(d-1)s+s}$$

Subsampled random points

- $\boldsymbol{X} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n\}$ drawn randomly i.i.d. w.r.t. Lebesgue measure such that $n \sim m \log m$
- there exist $\boldsymbol{X}' = \{\boldsymbol{x}^{i_1}, \ldots, \boldsymbol{x}^{i_{n'}}\} \subset \boldsymbol{X}$ subsampled points suitable for reconstruction with
$$n' \in \mathcal{O}(|I|)$$
- no fast matrix-vector multiplication



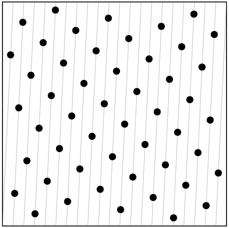
Theorem [Nagel, Schäfer, T. Ullrich '21], [B., Schäfer, T. Ullrich '22]

$$n^{-s}(\log n)^{(d-1)s} \lesssim \sup_{\|f\|_{H(K)} \le 1} \|f - S_{\boldsymbol{X}} f\|_{L_2} \lesssim n^{-s}(\log n)^{(d-1)s+1/2}$$

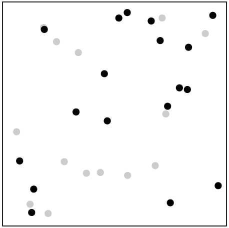
Theorem [Dolbeaut, Krieg, M. Ullrich '22]

$$\sup_{\|f\|_{H(K)} \le 1} \|f - S_{\boldsymbol{X}} f\|_{L_2} \sim n^{-s}(\log n)^{(d-1)s}$$

Goal: combine both advantages



structure for **fast algorithms**

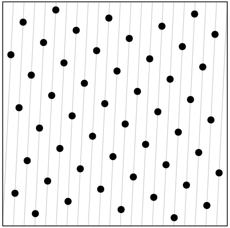


optimal rates

Approach (motivated by [Kunis, Rauhut '08])

1

$\boldsymbol{X}_{\mathsf{MZ}}$: good deterministic set of M points



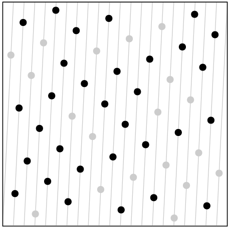
$\boldsymbol{X}_{\mathsf{MZ}} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\}$

$\boldsymbol{W}_{\mathsf{MZ}} \in [0, \infty)^{M \times M}$

random $\longrightarrow$ subsampling

2

$\boldsymbol{X}$: randomly subsample n points with $n \sim m \log m$



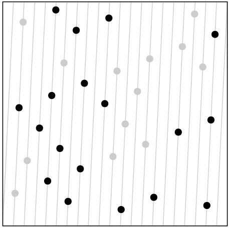
$\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J}$

$\boldsymbol{W} \in [0, \infty)^{|\boldsymbol{X}| \times |\boldsymbol{X}|}$

BSS $\longrightarrow$ subsampling

3

$\boldsymbol{X}'$: BSS subsample n' points with $n' \sim m$



$\boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i \in J'}$

$\boldsymbol{W}' \in [0,\infty)^{|\boldsymbol{X}'| \times |\boldsymbol{X}'|}$

Starting points: MZ inequalities and exact quadrature

- points $\boldsymbol{X}_{\mathsf{MZ}}$ and weights $\omega_1, \dots, \omega_M$ fulfill a L_2 **Marcinkiewicz-Zygmund inequality** for $V \subset L_2$, iff

$$A\|g\|_{L_2}^2 \leq \sum_{i=1}^{M} \omega_i |g(\boldsymbol{x}^i)|^2 \leq B\|g\|_{L_2}^2 \quad \text{for all} \quad g \in V$$

- the full system matrix is well-conditioned

$$A \leq \sigma_{\min}^2(\boldsymbol{W}^{1/2}\boldsymbol{L}) \leq \sigma_{\max}^2(\boldsymbol{W}^{1/2}\boldsymbol{L}) \leq B$$

Starting points: MZ inequalities and exact quadrature

- **MZ inequalities are widly available**: [Mhaskar, Narcowich, Ward '01], [Keiner, Kunis, Potts '07], [Filbir, Mhaskar '11], [Müller-Gronbach, Novak, Ritter '12], [Temlyakov '18], [Gröchenig '20], [Filbir, Hielscher, Jahn, T. Ullrich '22], ...

Lemma

The MZ inequality on V with $A = B$ is equivalent to the exact integration

$$\sum_{i=1}^{M} \omega_i g(\boldsymbol{x}^i)\overline{h(\boldsymbol{x}^i)} = \int_D g(\boldsymbol{x})\overline{h(\boldsymbol{x})}\,\mathrm{d}\nu(\boldsymbol{x}) \quad \text{for all} \quad g, h \in V\,.$$

- **exact integration is widly available:** [Nuyens, Cools '06], [Kämmerer, Potts, Volkmer '15], [Trefethen '19], ...

Incorporate fast algorithms

- big point set $\boldsymbol{X}_{\mathsf{MZ}} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\}$ with $\mathcal{O}(M \log M)$ algorithm
- small point set $\boldsymbol{X} = \{\boldsymbol{x}^{j_1}, \ldots, \boldsymbol{x}^{j_n}\} \subset \boldsymbol{X}_{\mathsf{MZ}}$
- we may use the algorithm for the big point set

$$\boldsymbol{L}_{\boldsymbol{X}} = \boldsymbol{P}\boldsymbol{L}_{\boldsymbol{X}_{\mathsf{MZ}}} \quad \text{where} \quad \boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ & \vdots & & & & \vdots & \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix} \begin{matrix} 1 \\ 2 \\ \vdots \\ n \end{matrix}$$

$$\begin{matrix} j_1 & & j_2 & & & j_n & \end{matrix}$$

- with $n \sim m \log m$ and $M \sim m^2$ we obtain the same complexity as for naive matrix-vector multiplication

$$\mathcal{O}(mn) = \mathcal{O}(M \log M)$$

Central theorem

Assumptions 1 of 2

Let

- $H(K)$ separable RKHS with $\sup_{\boldsymbol{x}\in D} K(\boldsymbol{x},\boldsymbol{x}) < \infty$ and $\int_D K(\boldsymbol{x},\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < \infty$,
- σ_k and η_k singular values and functions of $\mathrm{Id}_{K,\nu}\colon H(K) \hookrightarrow L_2(\nu)$
- $I = \{0,\dots,m-1\} \subset I_M$
- $\boldsymbol{X}_{\mathsf{MZ}}$ points and $\boldsymbol{W}_{\mathsf{MZ}}$ weights fulfilling an L_2-Marcinkiewicz-Zygmund inequality with constants A and B for $V = \mathrm{span}\{\eta_k\}_{k\in I_M}$,

Central theorem

Assumptions 2 of 2

- $n := \left\lceil \frac{72B}{A} |I| \log |I| \right\rceil$,
- $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J}$, $|J| = n$ points drawn i.i.d. from $\boldsymbol{X}_{\mathrm{MZ}}$ w.r.t. the **discrete density weights** $\varrho_i = \omega_i \varrho(\boldsymbol{x}^i)$ with

$$\varrho(\boldsymbol{x}^i) = \frac{\sum_{k \in I} |\eta_k(\boldsymbol{x}^i)|^2}{3 \sum_{j=1}^{M} \omega_j \sum_{k \in I} |\eta_k(\boldsymbol{x}^j)|^2} + \frac{\sum_{k \in I_M \setminus I} |e_k(\boldsymbol{x}^i)|^2}{3 \sum_{j=1}^{M} \omega_j \sum_{k \in I_M \setminus I} |e_k(\boldsymbol{x}^j)|^2} + \frac{\sum_{k \notin I_M} |e_k(\boldsymbol{x}^i)|^2}{3 \sum_{j=1}^{M} \omega_j \sum_{k \notin I_M} |e_k(\boldsymbol{x}^j)|^2},$$

- $b > 1 + \frac{10}{|I|}$.

- uses spectral properties of the embedding
- discrete version of [M. Ullrich, Krieg '19], [Kämmerer, T. Ullrich, Volkmer '19]
- $\varrho(\boldsymbol{x}^i)$ neglectable for BOS, i.e., $\sup_{x \in D} |\eta_k(x)| \leq B$ for all k

Central theorem

Theorem [B, Kämmerer, Potts, T. Ullrich '22]

Given the assumptions we construct $\boldsymbol{X}' \subset \boldsymbol{X}_{\mathsf{MZ}}$ with $|\boldsymbol{X}'| \le \lceil b|I|\rceil$ such that

$$\sup_{\|f\|_{H(K)}\le 1} \|f - S_I^{\boldsymbol{X}'} f\|_{L_2}^2 \le C_{A,B,b} \log|I| \Bigg(\sup_{k\notin I} \sigma_k^2 + \frac{1}{|I|} \sum_{k\in I_{\mathsf{MZ}}\setminus I} \sigma_k^2$$

$$+ \sup_{\|f\|_{H(K)}\le 1} \|f - S_{I_{\mathsf{MZ}}}^{\boldsymbol{X}_{\mathsf{MZ}}} f\|_{L_2}^2 + \frac{1}{|I|}\Big(\sum_{i=1}^{M} \omega_i\Big) \sup_{\|f\|_{H(K)}\le 1} \|f - P_{I_{\mathsf{MZ}}} f\|_{\ell_\infty(D)} \Bigg)$$

with probability larger than $1 - 6/|I|$ and $C_{A,B,b} = (\frac{B}{A})^2 (\frac{82\,b}{b-1})^3$.

Theorem for rank-1 lattices [B, Kämmerer, Potts, T. Ullrich '22]

Let
- $s > 1/2$,
- $I \subset I_{\mathsf{MZ}} \subset \mathbb{Z}^d$ hyperbolic cross frequency index sets, $|I| \geq 10$,
- $\boldsymbol{X}_{\mathsf{MZ}}$ reconstructing rank-1 lattice for I_{MZ} with M points,
- $b > 1 + \frac{10}{|I|}$.

Then we construct $\boldsymbol{X}' \subset \boldsymbol{X}_{\mathsf{MZ}}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ such that

$$\sup_{\|f\|_{H^s_{\mathsf{mix}}} \leq 1} \|f - S_I^{\boldsymbol{X}'} f\|_{L_2} \leq C_{d,s,b} \Big(|I|^{-s} (\log|I|)^{(d-1)s+1/2} + |I_{\mathsf{MZ}}|^{-s} (\log|I_{\mathsf{MZ}}|)^{(d-1)s+d/2} \Big)$$

with probability $1 - 6/|I|$.

Central theorem for rank-1 lattices

Theorem for rank-1 lattices [B, Kämmerer, Potts, T. Ullrich '22]

Let
- $s > 1/2$,
- $I \subset I_{\mathrm{MZ}} \subset \mathbb{Z}^d$ hyperbolic cross frequency index sets, $|I| \geq 10$,
- $\boldsymbol{X}_{\mathrm{MZ}}$ reconstructing rank-1 lattice for I_{MZ} with M points,
- $b > 1 + \frac{10}{|I|}$.

Then we construct $\boldsymbol{X}' \subset \boldsymbol{X}_{\mathrm{MZ}}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ such that

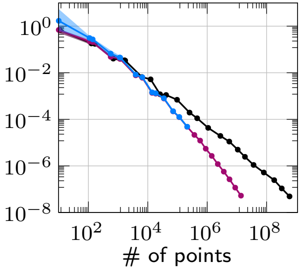
$$\sup_{\|f\|_{H^s_{\mathrm{mix}}} \leq 1} \|f - S_I^{\boldsymbol{X}'} f\|_{L_2} \leq C'_{d,s,b} \Big(|\boldsymbol{X}'|^{-s} (\log |\boldsymbol{X}'|)^{(d-1)s+1/2} + M^{-s/2} (\log M)^{\frac{(d-2)s+d}{2}} \Big)$$

with probability $1 - 6/|I|$.

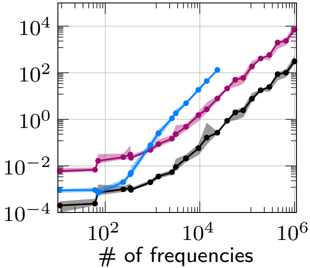
Numerical experiment for random subsampling

- $d = 5$, f tensorized bumb function in $H^s_{\text{mix}}(\mathbb{T}^5)$ for $s < 3/2$
- initial $\boldsymbol{X}_{\text{MZ}}$: fast probabilistic CBC rank-1 lattice construction [Kämmerer '20]

L_2-error



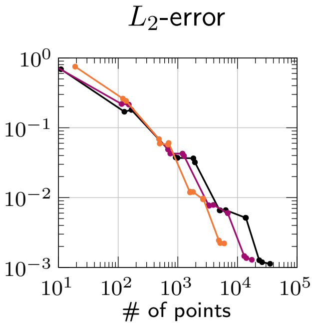
time in seconds

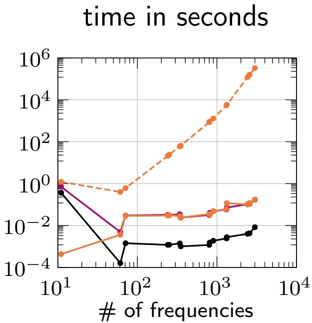


black: full rank-1 lattice
azure: continuously random
magenta: randomly subsampled

Numerical experiment for random + BSS subsampling

- $d = 5$, f tensorized bumb function in $H^s_{\text{mix}}(\mathbb{T}^5)$ for $s < 3/2$
- initial $\boldsymbol{X}_{\text{MZ}}$: fast probabilistic CBC rank-1 lattice construction [Kämmerer '20]

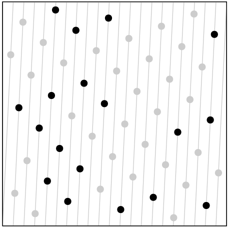




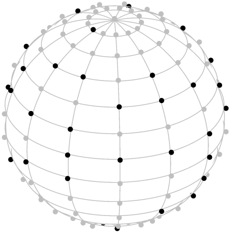
black: full rank-1 lattice
magenta: randomly subsampled
orange: BSS subsampled

Conclusion

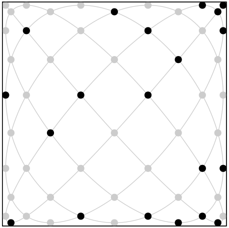
$\mathbb{T}^d$



$\mathbb{S}^d$



$[0,1]^d$



- we proposed **random + BSS** subsampling procedure for **MZ inequalities**
  - achieving the near optimal convergence rates ✓
  - making it possible to utilize fast algorithms ✓

- [B., Kämmerer, Potts, T. Ullrich] on **arXiv:2208.13597** "On the reconstruction of functions from values at subsampled quadrature points"