Sampling recovery from random partial quadrature nodes

Felix Bartel   Lutz Kämmerer   Daniel Potts   Tino Ullrich

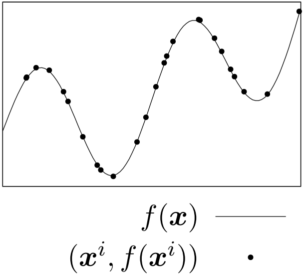
Sigmundsburg Seminar on Analysis and Theoretical Numerics

1st of September 2021

Mathematik!
TU Chemnitz

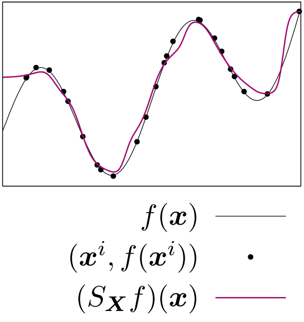
Recovery of functions from RKHS

- **Given:**
  - reproducing kernel Hilbert space $\mathrm{H}(K) \hookrightarrow \mathrm{L}_2$
  - samples $(f(\boldsymbol{x}^1), \dots, f(\boldsymbol{x}^n))^{\mathsf{T}} \in \mathbb{C}^n$ from $f \in \mathrm{H}(K)$



- **Given:**
  - reproducing kernel Hilbert space $\mathrm{H}(K) \hookrightarrow \mathrm{L}_2$
  - samples $(f(\boldsymbol{x}^1), \ldots, f(\boldsymbol{x}^n))^\mathsf{T} \in \mathbb{C}^n$ from $f \in \mathrm{H}(K)$
- **Goal:**
  - recover every $f \in \mathrm{H}(K)$ from the samples in $\boldsymbol{X} = (\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n)$
  - control **worst-case error** for sampling recovery operator $S_{\boldsymbol{X}} \colon \mathrm{H}(K) \to \mathrm{L}_2$

$$\sup_{\|f\|_{\mathrm{H}(K)} \le 1} \|f - S_{\boldsymbol{X}} f\|_{\mathrm{L}_2}$$



Reproducing kernel Hilbert space

- assume the finite trace condition $\int_D K(\boldsymbol{x},\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{x}) < \infty$
- embedding $\mathrm{Id}_{K,\nu}\colon \mathrm{H}(K) \to \mathrm{L}_2$ has the representation

$$\mathrm{Id}_{K,\nu}(f) = \sum_{k=1}^{\infty} \sigma_k \langle f, e_k\rangle_{\mathrm{H}(K)} \eta_k$$

  with

  - **singular values** $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ of $\mathrm{Id}_{K,\nu}$
  - **right singular functions** $\{e_1, e_2, \dots\}$ forming an ONS in $\mathrm{H}(K)$
  - **left singular functions** $\{\eta_1, \eta_2, \dots\}$ forming an ONS in L_2

- since $\mathrm{Id}_{K,\nu}$ is the identity on functions

$$e_k = \sigma_k \cdot \eta_k$$

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

- for $s > 1/2$ we define $\mathrm{H}^s_{\text{mix}}(\mathbb{T}^d) = \{f \in \mathrm{L}_2 : \|f\|_{\mathrm{H}^s_{\text{mix}}} < \infty\}$ with

$$\langle f, g\rangle_{\mathrm{H}^s_{\text{mix}}} := \sum_{\boldsymbol{j}\in\{0,s\}^d} \langle D^{(\boldsymbol{j})} f, D^{(\boldsymbol{j})} g\rangle_{\mathrm{L}_2}$$

- kernel:

$$K^1_s(x,y) = \sum_{k\in\mathbb{Z}} \frac{\exp(2\pi \mathrm{i} k(y-x))}{w_s(k)^2} \quad \text{with} \quad w_s(k) = (1 + (2\pi|k|)^{2s})^{1/2}$$

$$K^d_s(\boldsymbol{x},\boldsymbol{y}) = K^1_s(x_1,y_1) \otimes \cdots \otimes K^1_s(x_d,y_d)$$

- singular numbers: $\sigma_n = (1/w_s(\boldsymbol{k}_n))_n$ (non-increasing rearrangement)
- singular functions: $e_n(\boldsymbol{x}) = \sigma_n \eta_n(\boldsymbol{x}) = \sigma_n \exp(2\pi \mathrm{i} \boldsymbol{k}_n \cdot \boldsymbol{x})$

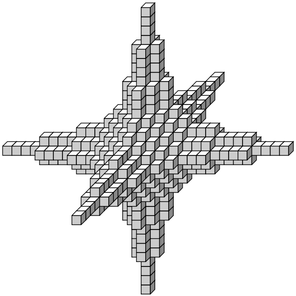
Example: Sobolev spaces with mixed smoothness on $\mathbb{T}^d$

- first $\eta_1, \dots, \eta_{m-1}$ are most important
- frequencies belong to hyperbolic cross

$$I_R^{hc} = \left\{ \boldsymbol{k} \in \mathbb{Z}^d : \prod_{j=1}^{d} (1 + (2\pi|k_j|)^{2s})^{1/2} \leq R \right\}$$

- approximation

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k} \in I_R^{hc}} c_{\boldsymbol{k}} \exp(2\pi \mathrm{i} \boldsymbol{k} \cdot \boldsymbol{x})$$



I_R^{hc}

Least squares

- assume $n > m$
- seek for $\boldsymbol{c} = (c_1, \dots, c_{m-1})^\mathsf{T}$ solving

$$\begin{pmatrix} \eta_1(\boldsymbol{x}^1) & \cdots & \eta_{m-1}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_1(\boldsymbol{x}^n) & \cdots & \eta_{m-1}(\boldsymbol{x}^n) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{m-1} \end{pmatrix} = \begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^n) \end{pmatrix}$$

Least squares

- assume $n > m$
- seek for $\boldsymbol{c} = (c_1, \ldots, c_{m-1})^{\mathsf{T}}$ solving

$$\left\| \begin{pmatrix} \eta_1(\boldsymbol{x}^1) & \cdots & \eta_{m-1}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_1(\boldsymbol{x}^n) & \cdots & \eta_{m-1}(\boldsymbol{x}^n) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{m-1} \end{pmatrix} - \begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^n) \end{pmatrix} \right\|_{\boldsymbol{W}}^2 \to \min$$

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

Least squares

- assume $n > m$
- seek for $\boldsymbol{c} = (c_1, \ldots, c_{m-1})^\mathsf{T}$ solving

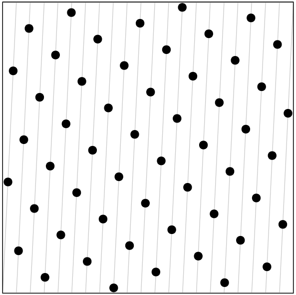
$$\left\| \underbrace{\begin{pmatrix} \eta_1(\boldsymbol{x}^1) & \cdots & \eta_{m-1}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_1(\boldsymbol{x}^n) & \cdots & \eta_{m-1}(\boldsymbol{x}^n) \end{pmatrix}}_{=:\boldsymbol{L}} \begin{pmatrix} c_1 \\ \vdots \\ c_{m-1} \end{pmatrix} - \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^n) \end{pmatrix}}_{=:\boldsymbol{f}} \right\|_{\boldsymbol{W}}^2 \to \min$$

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

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

Deterministic nodes: rank-1 lattices

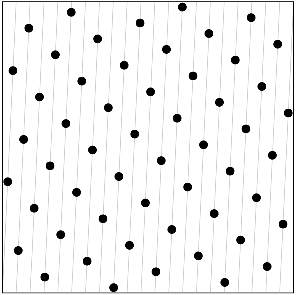
- $\boldsymbol{X} = \Lambda(\boldsymbol{z}, M)$ with $\boldsymbol{x}^i = \frac{1}{M}(i\boldsymbol{z} \bmod M\mathbf{1})$ for $i = 0, \ldots, M-1$
- [Kämmerer, Potts, Volkmer '15]:
  - given I, there exist algorithms that find $\boldsymbol{z}$ and M such that $\boldsymbol{L}^*\boldsymbol{W}\boldsymbol{L} = \boldsymbol{I}$
  - multiplication with $\boldsymbol{L}$ can be carried out with the LFFT in $\mathcal{O}(M \log M)$



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

Deterministic nodes: rank-1 lattices

- $\boldsymbol{X} = \Lambda(\boldsymbol{z}, M)$ with $\boldsymbol{x}^i = \frac{1}{M}(i\boldsymbol{z} \bmod M\mathbf{1})$ for $i = 0, \ldots, M-1$
- [Kämmerer, Potts, Volkmer '15]:
  - given I, there exist algorithms that find $\boldsymbol{z}$ and M such that $\boldsymbol{L}^*\boldsymbol{W}\boldsymbol{L} = \boldsymbol{I}$
  - multiplication with $\boldsymbol{L}$ can be carried out with the LFFT in $\mathcal{O}(M \log M)$



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

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

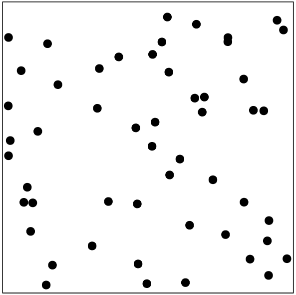
$$M^{-s/2} \lesssim \sup_{\|f\|_{\mathrm{H}(K)} \le 1} \|f - S_{\boldsymbol{X},\boldsymbol{W}} f\|_{\mathrm{L}_2} \lesssim M^{-s/2} (\log M)^{\frac{d-2}{2}s + \frac{d-1}{2}}$$

Random nodes on $\mathbb{T}^d$

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

$$|I| \lesssim \frac{n}{\log n}$$

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



random nodes

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

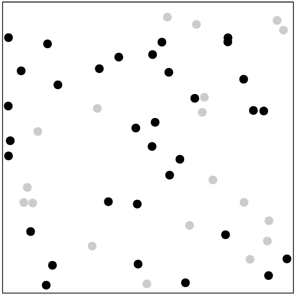
$$\sup_{\|f\|_{\mathrm{H}(K)} \leq 1} \|f - S_{\boldsymbol{X},\boldsymbol{W}} f\|_{\mathrm{L}_2} \lesssim n^{-s} (\log n)^{(d-1)s+s}$$

Weaver subsampled random nodes

- $\boldsymbol{X} = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^n\}$ drawn randomly i.i.d. w.r.t. Lebesgue measure such that $|I| \leq n/(10r\log n)$
- there exist $\boldsymbol{X}' = \{\boldsymbol{x}^{i_1}, \dots, \boldsymbol{x}^{i_{n'}}\} \subset \boldsymbol{X}$ subsampled nodes suitable for reconstruction with

$$n \in \mathcal{O}(|I|)$$

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



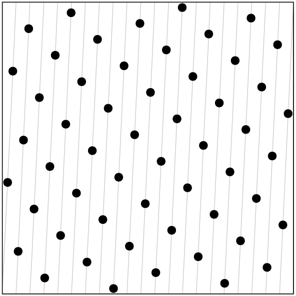
subsampled random nodes

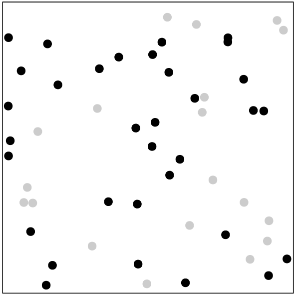
[Nagel, Schäfer, T. Ullrich '21]:

$$\sup_{\|f\|_{\mathrm{H}(K)}\leq 1} \|f - S_{\boldsymbol{X},\boldsymbol{W}} f\|_{\mathrm{L}_2} \lesssim n^{-s}(\log n)^{(d-1)s+1/2}$$

Goal

We want to combine both advantages:





- structure for fast algorithms
- awesome rates

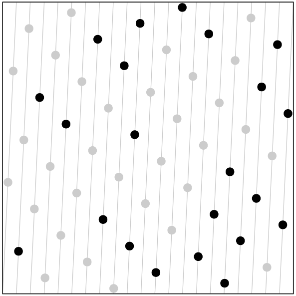
Approach (motivated by [Kunis, Rauhut '08])

1. $\boldsymbol{X}_M$: good deterministic set of M nodes
2. $\boldsymbol{X}$: randomly subsample n nodes from $\boldsymbol{X}_M$ such that

$$|I| \lesssim \frac{n}{\log n}$$

3. $\boldsymbol{X}'$: Weaver subsample n' nodes from $\boldsymbol{X}$ such that

$$n' \in \mathcal{O}(|I|)$$



Starting node set

- nodes $\boldsymbol{X}_M$ and weights $\omega_1, \dots, \omega_M$ fulfill a L_2 **Marcinkiewicz-Zygmund inequality** iff

$$c_{\mathsf{MZ}} \|f\|_{\mathrm{L}_2}^2 \leq \sum_{i=1}^{M} \omega_i |f(\boldsymbol{x}^i)|^2 \leq C_{\mathsf{MZ}} \|f\|_{\mathrm{L}_2}^2$$

  for $f \in \mathrm{span}\{\eta_1, \dots, \eta_{m-1}\}$.
- the full system matrix is well-conditioned

$$\mathrm{spec}(\boldsymbol{W}^{1/2} \boldsymbol{L}_{\boldsymbol{X}_M}) \in [\sqrt{c_{\mathsf{MZ}}}, \sqrt{C_{\mathsf{MZ}}}]$$

Starting node set

- 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], [Hielscher, Jahn, Ullrich '21], ...

Starting node set

- 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], [Hielscher, Jahn, Ullrich '21], ...

Lemma

The MZ-inequality for $c_{\mathsf{MZ}} = C_{\mathsf{MZ}}$ is equivalent to the exact integration

$$\sum_{i=1}^{M} \omega_i g(\boldsymbol{x}^i)\overline{h(\boldsymbol{x}^i)} = \int_D g(\boldsymbol{x})\overline{h(\boldsymbol{x})}\,\mathrm{d}\nu(\boldsymbol{x})$$

for $g, h \in \operatorname{span}\{\eta_1, \ldots, \eta_{m-1}\}$.

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

Fast algorithm

- big node set $\boldsymbol{X}_M = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\}$ with $\mathcal{O}(M \log M)$ algorithm
- small node set $\boldsymbol{X} = \{\boldsymbol{x}^{j_1}, \dots, \boldsymbol{x}^{j_n}\}$
- we may use the algorithm for the big node set

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

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

Fast algorithm

- big node set $\boldsymbol{X}_M = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\}$ with $\mathcal{O}(M \log M)$ algorithm
- small node set $\boldsymbol{X} = \{\boldsymbol{x}^{j_1}, \dots, \boldsymbol{x}^{j_n}\}$
- we may use the algorithm for the big node set

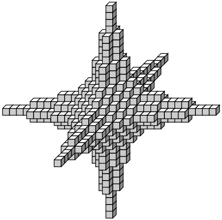
$$\boldsymbol{L}_{\boldsymbol{X}} = \boldsymbol{P}\boldsymbol{L}_{\boldsymbol{X}_M} \quad \text{where} \quad \boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ & \vdots & & & & \vdots & \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix} \begin{matrix} 1 \\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

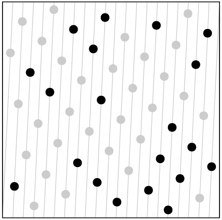
$$\qquad\qquad j_1 \qquad j_2 \qquad\qquad j_n$$

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

$$\mathcal{O}(|I| \cdot n) = \mathcal{O}(M \log M)$$

- $d = 2$, $\boldsymbol{X}_M$ Fibonacci lattice
- hyperbolic cross frequencies





$\lvert I \rvert$	M	n	store matrix	column-wise	$1d$ FFT
6 089	196 418	67 729	17 + 0.176 s 7 GB	16 s	0.047 s
13 581	832 040	162 991	99 + 1.377 s 35 GB	93 s	0.119 s
29 977	3 524 578	385 579	490 + 10 s 185 GB	480 s	0.744 s

Assumptions 1 of 2

Let

- $\mathrm{H}(K)$ RKHS with $\sup_{\boldsymbol{x}\in D} K(\boldsymbol{x},\boldsymbol{x}) < \infty$,
- $I \subset I_M$ frequency index sets,
- $\boldsymbol{X}_M$ nodes and $\boldsymbol{W}$ weights fulfilling an L_2-Marcinkiewicz-Zygmund inequality with constants c_{MZ} and C_{MZ} for $V = \mathrm{span}\{\eta_k\}_{k\in I_M}$,
- n smallest such that

$$|I| \le \frac{c_{\mathsf{MZ}} n}{30 C_{\mathsf{MZ}} r \log n},$$

Assumptions 2 of 2

- $\boldsymbol{X} = (\boldsymbol{x}^i)_{i\in J}$, $|J| = n$ nodes drawn i.i.d. from $\boldsymbol{X}_M$ w.r.t. the **discrete density weights** $\varrho_i = \omega_i \varrho(\boldsymbol{x}^i)$ for $1 \le i \le M$ with

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

- uses spectral properties of the embedding
- $\varrho(\boldsymbol{x}^i)$ neglectable for BOS
- discrete version of [M. Ullrich, Krieg '19], [Kämmerer, T. Ullrich, Volkmer '19]

Central theorem

Theorem

Given the assumptions we have with probability larger than $1-5n^{1-r}$

$$\sup_{\|f\|_{\mathrm{H}(K)}\le 1}\left\|f-S^{I}_{\boldsymbol{X},\tilde{\boldsymbol{W}}}f\right\|_2^2 \le \frac{6C_{\mathsf{MZ}}}{c_{\mathsf{MZ}}}\Bigg(\sup_{k\notin I}\sigma_k^2+\frac{1}{|I|}\sum_{k\in I_M\setminus I}\sigma_k^2$$
$$+\frac{1}{|I|}\sup_{\|f\|_{\mathrm{H}(K)}\le 1}\|f-P_{I_M}f\|_\infty^2+\sup_{\|f\|_{\mathrm{H}(K)}\le 1}\left\|f-S^{I_M}_{\boldsymbol{X}_M,\boldsymbol{W}}f\right\|_{\mathrm{L}_2}^2\Bigg).$$

Theorem

Given the same assumptions but with

$$n' \in \mathcal{O}(|I|)$$

we have the existence of $\boldsymbol{X}' = \{\boldsymbol{x}^{i_1}, \dots, \boldsymbol{x}^{i_{n'}}\}$ such that

$$\sup_{\|f\|_{\mathrm{H}(K)} \leq 1} \left\| f - S^{I}_{\boldsymbol{X}', \tilde{\boldsymbol{W}}} f \right\|_2^2 \lesssim \log|I| \Bigg(\sup_{k \notin I} \sigma_k^2 + \frac{1}{|I|} \sum_{k \in I_M \setminus I} \sigma_k^2$$
$$+ \frac{1}{|I|} \sup_{\|f\|_{\mathrm{H}(K)} \leq 1} \|f - P_{I_M} f\|_\infty^2 + \sup_{\|f\|_{\mathrm{H}(K)} \leq 1} \left\| f - S^{I_M}_{\boldsymbol{X}_M, \boldsymbol{W}} f \right\|_{\mathrm{L}_2}^2 \Bigg).$$

Theorem on $\mathbb{T}^d$ for $\mathrm{H}^s_{\text{mix}}$

Theorem

Let

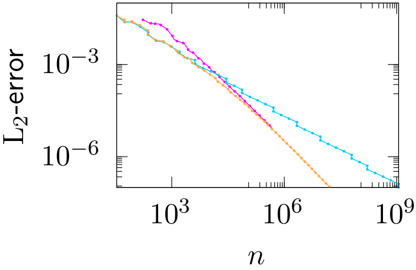
- $I \subset I_M$ frequency index sets,
- $\boldsymbol{X}_M$ and $\boldsymbol{W}$ exact for trigonometric polynomials supported on $\mathcal{D}(I_M)$
- n smallest such that $|I| \leq \frac{n}{30r \log n}$,
- $\boldsymbol{X} = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^n\}$ drawn w.r.t. the discrete density weights $\boldsymbol{W}$.

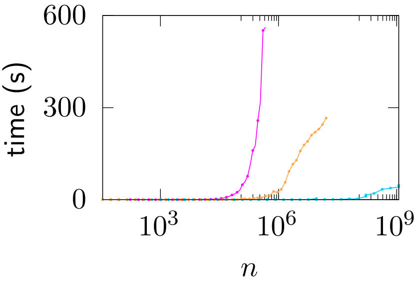
Then we have with probability larger than $1 - 5n^{-r}$

$$\sup_{\|f\|_{\mathrm{H}^s_{\text{mix}}} \leq 1} \left\| f - S^I_{\boldsymbol{X}, \tilde{\boldsymbol{W}}} f \right\|_2^2 \leq \sup_{\boldsymbol{k} \notin I} \frac{5}{w(\boldsymbol{k})^2} + \frac{6}{|I|} \sum_{\boldsymbol{k} \notin I} \frac{1}{w(\boldsymbol{k})^2} + 4 \sup_{\|f\|_{\mathrm{H}(K)} \leq 1} \left\| f - S^{I_M}_{\boldsymbol{X}_M, \boldsymbol{W}} f \right\|_{\mathrm{L}_2}^2 .$$

Numerics on $\mathbb{T}^d$

- $d = 2$, $\boldsymbol{X}_M$ Fibonacci lattice
- hyperbolic cross frequencies





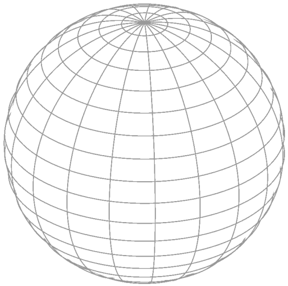
magenta: random nodes
cyan: full rank-1 lattice
orange: subsampled rank-1 lattice

The two-dimensional sphere $\mathbb{S}^2$

- **domain:** $D = \mathbb{S}^2 = \{\boldsymbol{\xi} \in \mathbb{R}^3 : \|\boldsymbol{\xi}\|_2 = 1\}$
- **basis functions** for the spherical harmonics

$$Y_{k,\ell}(\theta, \varphi) = \sqrt{\frac{2k+1}{4\pi}} P_{|\ell|}^{k}(\cos\theta) \exp(\mathrm{i}\ell\varphi)$$

for $k = 0, 1, \ldots$ and $-k \leq \ell \leq k$



Theorem

Let

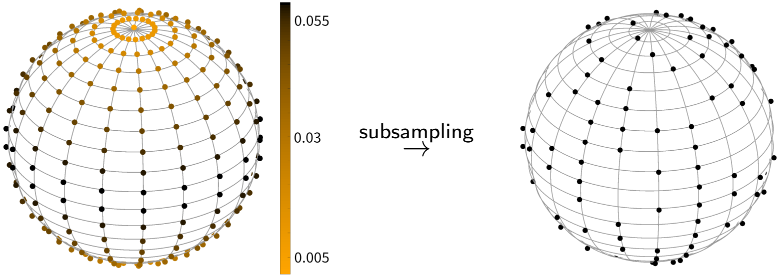
- $m \leq m_M$ polynomial degrees
- $\boldsymbol{X}_M$ and $\boldsymbol{W}$ exact for polynomials upto degre $2m_M$
- n smallest such that $(m+1)^2 \leq \frac{n}{30r\log n}$
- $\boldsymbol{X} = \{\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^n\}$ drawn w.r.t. the discrete density weights $\boldsymbol{W}$.

Then we have with probability larger than $1 - 5n^{-r}$

$$\sup_{\|f\|_{\mathrm{H}^s_{\text{mix}}} \leq 1} \left\| f - S^I_{\boldsymbol{X},\tilde{\boldsymbol{W}}} f \right\|_2^2 \leq \sup_{\boldsymbol{k}\notin I} \frac{5}{w(\boldsymbol{k})^2} + \frac{6}{(m+1)^2} \sum_{\boldsymbol{k}\notin I} \frac{2k+1}{4\pi w(\boldsymbol{k})^2}$$
$$+ 4 \sup_{\|f\|_{\mathrm{H}(K)} \leq 1} \left\| f - S^{I_M}_{\boldsymbol{X}_M,\boldsymbol{W}} f \right\|_{\mathrm{L}_2}^2 .$$

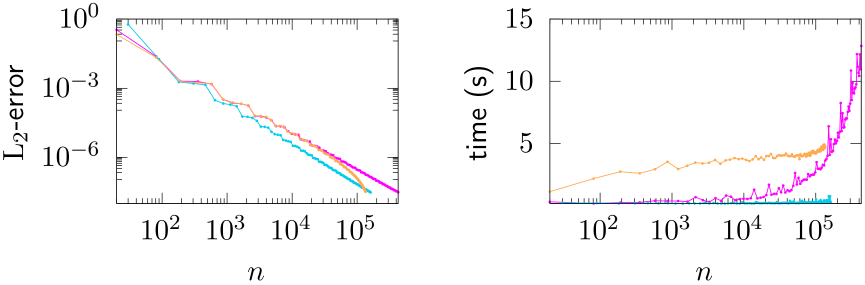
Example: quadrature nodes on $\mathbb{S}^2$

- $\boldsymbol{X}$ tensor product of equispaced nodes and Chebyshev nodes
- $\boldsymbol{W}$ Chebyshev weights



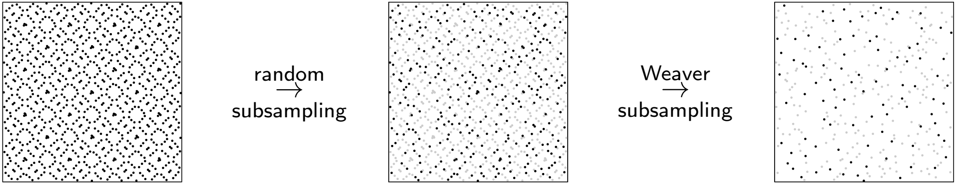
- exact upto polynomial degree m with $n = 2m(m+2)$ nodes
- $2d$ FFT for matrix-vector product

- $\boldsymbol{X}_M$ tensor product of equispaced nodes and Chebyshev nodes



magenta: random nodes
cyan: full tensor product grid
orange: subsampled tensor product grid

Conclusion



- we proposed a two step subsampling procedure
    - making it possible to utilize fast algorithms ✓
    - achieving the optimal convergence rates known so far ✓