Constructive subsampling of finite frames with applications in optimal function recovery

Felix Bartel   Martin Schäfer   Tino Ullrich

15th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing

21st July 2022

Mathematik!
TU Chemnitz

Motivation

- **Given:**
  - samples $f(\boldsymbol{x}^1), \dots, f(\boldsymbol{x}^M)$ and
  - a finite dimensional function space $V = \operatorname{span}\{\eta_0, \dots, \eta_{m-1}\}$.
- **Goal:** Find approximation g of f in V based on the samples with small L_2-error $\int_D |f(\boldsymbol{x}) - g(\boldsymbol{x})|^2 \, \mathrm{d}\mu(\boldsymbol{x})$.

Motivation

- **Given:**
  - samples $f(\boldsymbol{x}^1), \ldots, f(\boldsymbol{x}^M)$ and
  - a finite dimensional function space $V = \operatorname{span}\{\eta_0, \ldots, \eta_{m-1}\}$.
- **Goal:** Find approximation g of f in V based on the samples with small L_2-error $\int_D |f(\boldsymbol{x}) - g(\boldsymbol{x})|^2 \,\mathrm{d}\mu(\boldsymbol{x})$.
- **Ansatz:** $g(\boldsymbol{x}) = \sum_{k=0}^{m-1} c_k \eta_k(\boldsymbol{x})$ with c_k minimizing

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

Motivation

- **Given:**
  - samples $f(\boldsymbol{x}^1), \dots, f(\boldsymbol{x}^M)$ and
  - a finite dimensional function space $V = \operatorname{span}\{\eta_0, \dots, \eta_{m-1}\}$.
- **Goal:** Find approximation g of f in V based on the samples with small L_2-error $\int_D |f(\boldsymbol{x}) - g(\boldsymbol{x})|^2 \, \mathrm{d}\mu(\boldsymbol{x})$.
- **Ansatz:** $g(\boldsymbol{x}) = \sum_{k=0}^{m-1} c_k \eta_k(\boldsymbol{x})$ with c_k minimizing

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

for a weight matrix $\boldsymbol{W} = \operatorname{diag}(\omega_1, \dots, \omega_M)$.

Motivation

Definition

Points $\boldsymbol{x}^1, \dots, \boldsymbol{x}^M$ and weights $\omega_1, \dots, \omega_M$ fulfill a **L_2 Marcinkiewicz-Zygmund inequality** for a function space V, iff

$$A\|f\|_{L_2}^2 \le \sum_{i=1}^{M} \omega_i |f(\boldsymbol{x}^i)|^2 \le B\|f\|_{L_2}^2 \quad \forall f \in V\,.$$

- widly available: [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], [Hielscher, Jahn, T. Ullrich '21], ...

Motivation

Definition

Points $\boldsymbol{x}^1, \dots, \boldsymbol{x}^M$ and weights $\omega_1, \dots, \omega_M$ fulfill a L_2 **Marcinkiewicz-Zygmund inequality** for a function space V, iff

$$A\|f\|_{L_2}^2 \le \sum_{i=1}^{M} \omega_i |f(\boldsymbol{x}^i)|^2 \le B\|f\|_{L_2}^2 \quad \forall f \in V\,.$$

Lemma (equivalence to frame condition)

Let
- $\eta_0, \dots, \eta_{m-1}$ be an ONB of V and
- $\boldsymbol{y}^i = (\eta_0(\boldsymbol{x}^i), \dots, \eta_{m-1}(\boldsymbol{x}^i))^\mathsf{T} \in \mathbb{C}^m$.

Then the above is equivalent to

$$A\|\boldsymbol{a}\|_2^2 \le \sum_{i=1}^{M} |\langle \boldsymbol{a}, \sqrt{\omega_i}\boldsymbol{y}^i\rangle|^2 \le B\|\boldsymbol{a}\|_2^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m\,.$$

Problem: Subsampling of frames

Given:

- a **frame**, i.e., vectors $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M \in \mathbb{C}^m$ and constants $0 < A \leq B < \infty$ s.t.

$$A\|\boldsymbol{a}\|_2^2 \leq \sum_{i=1}^{M} |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \leq B\|\boldsymbol{a}\|_2^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m ,$$

- target oversampling factor $b > 1$.

Problem: Subsampling of frames

Given:

- a **frame**, i.e., vectors $\boldsymbol{y}^1, \dots, \boldsymbol{y}^M \in \mathbb{C}^m$ and constants $0 < A \le B < \infty$ s.t.

$$A\|\boldsymbol{a}\|_2^2 \le \sum_{i=1}^{M} |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \le B\|\boldsymbol{a}\|_2^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m ,$$

- target oversampling factor $b > 1$.

Goal: find an index set $J \subset [M]$ with $|J| \le \lceil bm \rceil$ such that

$$A'\|\boldsymbol{a}\|_2^2 \le \sum_{i \in J} |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m .$$

Previous results

Theorem (existence) [Nagel, Schäfer, T. Ullrich, '21]

Let $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M \in \mathbb{C}^m$ be a frame with $\|\boldsymbol{y}^i\|_2^2 \le \frac{m}{M}$ and $b > \frac{1642}{A}$. Then there exists $J \subset [M]$ of size $|J| \le bm$, with

$$12\|\boldsymbol{a}\|_2^2 \le \frac{1}{m} \sum_{i \in J} |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \le 1642 \frac{B}{A} \|\boldsymbol{a}\|_2^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m .$$

- ☺ The desired subframe exists!
- ☹ The approach is non-constructive as it is based on the Kadison-Singer theorem [Marcus, Spielman, Srivastava '15].
- ☹ The oversampling factor b cannot be choosen close to one.

Previous results

Theorem (weighted construction) [Batson, Spielman, Srivastava, '12]

Let $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M \in \mathbb{C}^m$ form a real 1-tight frame and $b > 1$. Then the BSS algorithm computes $J \subset [M]$ with $|J| \le \lceil bm \rceil$ and $s_i \ge 0$, s.t.

$$\|\boldsymbol{a}\|_2^2 \le \sum_{i \in J} s_i |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \le \underbrace{\frac{(\sqrt{b}+1)^2}{(\sqrt{b}-1)^2}}_{=:\gamma} \|\boldsymbol{a}\|_2^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m .$$

- ☺ The approach is constructive.
- ☹ It only works for real 1-tight frames.
- ☹ We introduce further weights s_i.

Q: How to extend the BSS Algorithm to non-tight frames?

BSS for non-tight frames

Lemma (construction)

For every matrix $\boldsymbol{Y} \in \mathbb{C}^{M \times m}$ with $M \geq m$ there is a matrix $\tilde{\boldsymbol{Y}} \in \mathbb{C}^{M \times m}$ with rows $\tilde{\boldsymbol{y}}^1, \ldots, \tilde{\boldsymbol{y}}^M \in \mathbb{C}^m$ such that

$$\operatorname{range}(\tilde{\boldsymbol{Y}}) \supset \operatorname{range}(\boldsymbol{Y}) \quad \text{and} \quad \tilde{\boldsymbol{Y}}^* \tilde{\boldsymbol{Y}} = \boldsymbol{I}\,.$$

BSS for non-tight frames

Lemma (construction)

For every matrix $\boldsymbol{Y} \in \mathbb{C}^{M\times m}$ with $M \geq m$ there is a matrix $\tilde{\boldsymbol{Y}} \in \mathbb{C}^{M\times m}$ with rows $\tilde{\boldsymbol{y}}^1, \ldots, \tilde{\boldsymbol{y}}^M \in \mathbb{C}^m$ such that

$$\mathrm{range}(\tilde{\boldsymbol{Y}}) \supset \mathrm{range}(\boldsymbol{Y}) \quad \text{and} \quad \tilde{\boldsymbol{Y}}^*\tilde{\boldsymbol{Y}} = \boldsymbol{I}\,.$$

Applying BSS to $\tilde{\boldsymbol{y}}^1, \ldots, \tilde{\boldsymbol{y}}^M$, which form a tight frame, gives

$$\begin{aligned}
& \|\tilde{\boldsymbol{Y}}\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{S}^{\frac{1}{2}}(\tilde{\boldsymbol{Y}}\boldsymbol{a})|_J\|_2^2 \leq \gamma\|\tilde{\boldsymbol{Y}}\boldsymbol{a}\|_2^2 \quad \text{for all } \boldsymbol{a} \in \mathbb{C}^m \\
\Leftrightarrow \quad & \|\boldsymbol{u}\|_2^2 \leq \|\boldsymbol{S}^{\frac{1}{2}}(\boldsymbol{u})|_J\|_2^2 \leq \gamma\|\boldsymbol{u}\|_2^2 \quad \text{for all } \boldsymbol{u} \in \mathrm{range}(\tilde{\boldsymbol{Y}}) \supset \mathrm{range}(\boldsymbol{Y}) \\
\Leftrightarrow \quad & \|\boldsymbol{Y}\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{S}^{\frac{1}{2}}(\boldsymbol{Y}\boldsymbol{a})|_J\|_2^2 \leq \gamma\|\boldsymbol{Y}\boldsymbol{a}\|_2^2 \quad \text{for all } \boldsymbol{a} \in \mathbb{C}^m\,.
\end{aligned}$$

BSS for non-tight frames

Lemma (BSS$^\perp$)

Let $\tilde{\boldsymbol{y}}^1, \ldots, \tilde{\boldsymbol{y}}^M \in \mathbb{C}^m$ be the vectors associated to $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M$ from the previous Lemma, $b > 1$. Applying BSS to them gives

$$\sum_{i=1}^{M} |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \leq \sum_{i \in J} s_i |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \leq \gamma \sum_{i=1}^{M} |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m .$$

In particular, if $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M$ form a frame with constants A and B, we have

$$A \|\boldsymbol{a}\|_2^2 \leq \sum_{i \in J} s_i |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \leq \gamma B \|\boldsymbol{a}\|_2^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m .$$

Q: How to get rid of the weights s_i in the BSS Algorithm?

$$A\|\boldsymbol{a}\|_2^2 \leq \sum_{i\in J} s_i |\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle|^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m$$

Lemma (non-weighted BSS under extra condition)

Let • $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^M \in \mathbb{C}^m$ be a frame with bounds $0 < A \le B < \infty$,

• $\|\boldsymbol{y}^i\|_2^2 \ge \beta m/M$, $\beta > 0$.

Then, for $b > 1$, there is a subset $J \subset \{1, \ldots, M\}$ with $|J| \le \lceil bm \rceil$ such that

$$\frac{A}{M}\|\boldsymbol{a}\|_2^2 \le \frac{B\gamma}{\beta}\frac{1}{m}\sum_{i\in J}\left|\left\langle \boldsymbol{a}, \boldsymbol{y}^i\right\rangle\right|^2 \quad \text{for all} \quad \boldsymbol{a} \in \mathbb{C}^m .$$

Lemma (non-weighted BSS under extra condition)

Let • $\boldsymbol{y}^1, \dots, \boldsymbol{y}^M \in \mathbb{C}^m$ be a frame with bounds $0 < A \le B < \infty$,

•$\|\boldsymbol{y}^i\|_2^2 \ge \beta m/M$, $\beta > 0$.

Then, for $b > 1$, there is a subset $J \subset \{1, \dots, M\}$ with $|J| \le \lceil bm \rceil$ such that

$$\frac{A}{M}\|\boldsymbol{a}\|_2^2 \le \frac{B\gamma}{\beta}\frac{1}{m}\sum_{i\in J}\left|\langle \boldsymbol{a}, \boldsymbol{y}^i\rangle\right|^2 \quad \text{for all} \quad \boldsymbol{a} \in \mathbb{C}^m .$$

Proof idea: Use BSS$^\perp$. Setting $\boldsymbol{a} = \boldsymbol{y}^j$ for $j \in J$, we obtain

$$s_j\|\boldsymbol{y}^j\|_2^4 \le \sum_{i\in J} s_i|\langle \boldsymbol{y}^j, \boldsymbol{y}^i\rangle|^2 \le \gamma\sum_{i=1}^{M}|\langle \boldsymbol{y}^j, \boldsymbol{y}^i\rangle|^2 \le \gamma B\|\boldsymbol{y}^j\|_2^2 .$$

■

Non-weighted BSS

Lemma (construction)

Let $\boldsymbol{Y} \in \mathbb{C}^{M\times m}$ be a matrix and $\alpha > 0$ such that $M/\lceil \alpha m \rceil \in \mathbb{N}$. Then there is a matrix $\tilde{\boldsymbol{Y}} \in \mathbb{C}^{M\times m'}$ with rows $\tilde{\boldsymbol{y}}^1, \dots, \tilde{\boldsymbol{y}}^M \in \mathbb{C}^{m'}$, $m' \in \{\lceil \alpha m \rceil, \dots, \lceil (1+\alpha)m \rceil\}$ such that

$$\operatorname{range}(\tilde{\boldsymbol{Y}}) \supset \operatorname{range}(\boldsymbol{Y}), \quad \tilde{\boldsymbol{Y}}^*\tilde{\boldsymbol{Y}} = \boldsymbol{I}, \quad \text{and} \quad \|\tilde{\boldsymbol{y}}^i\|_2^2 \geq \frac{\lceil \alpha m \rceil}{M}.$$

Non-weighted BSS

Lemma (construction)

Let $\boldsymbol{Y} \in \mathbb{C}^{M\times m}$ be a matrix and $\alpha > 0$ such that $M/\lceil \alpha m \rceil \in \mathbb{N}$. Then there is a matrix $\tilde{\boldsymbol{Y}} \in \mathbb{C}^{M\times m'}$ with rows $\tilde{\boldsymbol{y}}^1, \dots, \tilde{\boldsymbol{y}}^M \in \mathbb{C}^{m'}$, $m' \in \{\lceil \alpha m \rceil, \dots, \lceil (1+\alpha)m \rceil\}$ such that

$$\operatorname{range}(\tilde{\boldsymbol{Y}}) \supset \operatorname{range}(\boldsymbol{Y}), \quad \tilde{\boldsymbol{Y}}^*\tilde{\boldsymbol{Y}} = \boldsymbol{I}, \quad \text{and} \quad \|\tilde{\boldsymbol{y}}^i\|_2^2 \geq \frac{\lceil \alpha m \rceil}{M}.$$

Proof idea: Define

$$\boldsymbol{d}^1 := \sqrt{\frac{\lceil \alpha m \rceil}{M}} \begin{bmatrix} \boldsymbol{1} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix}, \ \boldsymbol{d}^2 := \sqrt{\frac{\lceil \alpha m \rceil}{M}} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{1} \\ \vdots \\ \boldsymbol{0} \end{bmatrix}, \dots, \ \boldsymbol{d}^{\lceil \alpha m \rceil} := \sqrt{\frac{\lceil \alpha m \rceil}{M}} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{1} \end{bmatrix},$$

where each $\boldsymbol{0}$ and $\boldsymbol{1}$ vector is of size $M/\lceil \alpha m \rceil \times 1$.
Gram-Schmidt the colums of $\boldsymbol{Y} = [\boldsymbol{c}^1 | \dots | \boldsymbol{c}^m]$ onto that

$$\tilde{\boldsymbol{Y}} := \left[\boldsymbol{d}^1 \middle| \cdots \middle| \boldsymbol{d}^{\lceil \alpha m \rceil} \middle| \tilde{\boldsymbol{c}}^1 \middle| \cdots \middle| \tilde{\boldsymbol{c}}^l \right] = \begin{bmatrix} (\tilde{\boldsymbol{y}}^1)^* \\ \hline \vdots \\ \hline (\tilde{\boldsymbol{y}}^M)^* \end{bmatrix}.$$

■

Non-weighted BSS

Theorem [B, Schäfer, T. Ullrich '22]

Let $\boldsymbol{y}^1, \dots, \boldsymbol{y}^M \in \mathbb{C}^m$ be vectors with $m \in \mathbb{N}_{\geq 10}$ and $b \geq 1 + \frac{10}{m}$.
We then obtain indices $J \subset [M]$ with $|J| \leq \lceil bm \rceil$ such that

$$\frac{1}{M} \sum_{i=1}^{M} \left|\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle\right|^2 \leq \frac{432 b^3}{(b-1)^3} \frac{1}{m} \sum_{i \in J} \left|\langle \boldsymbol{a}, \boldsymbol{y}^i \rangle\right|^2 \quad \forall \boldsymbol{a} \in \mathbb{C}^m .$$

Application

Corollary (subsampling of MZ-inequalities)

Let points $\boldsymbol{X}_M$ and weights $\omega_1, \ldots, \omega_M$ fulfill a L_2 Marcinkiewicz-Zygmund inequality for a function space V and $b \geq 1 + \frac{10}{m}$. Then we obtain indices $J \subset [M]$ with $|J| \leq \lceil bm \rceil$ such that

$$A \frac{(b-1)^3}{432 b^3} \|f\|_{L_2}^2 \leq \frac{M}{m} \sum_{i \in J} \omega_i |f(\boldsymbol{x}^i)|^2 \quad \forall f \in V \, .$$

Application

Theorem (nearly optimal function recovery) [B, Schäfer, T. Ullrich '22]

Let

- $H(K)$ be a RKHS with a kernel of finite trace,
- $\mathrm{Id}\colon H(K) \hookrightarrow L_2$ compact embedding with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$,
- V be the subspace spanned by the first m left singular functions
- $1 + \frac{m}{10} \leq b$.

Then there is a point set $\boldsymbol{X}_n$ with $|\boldsymbol{X}_n| \leq \lceil bm \rceil$ and weights w_m such that

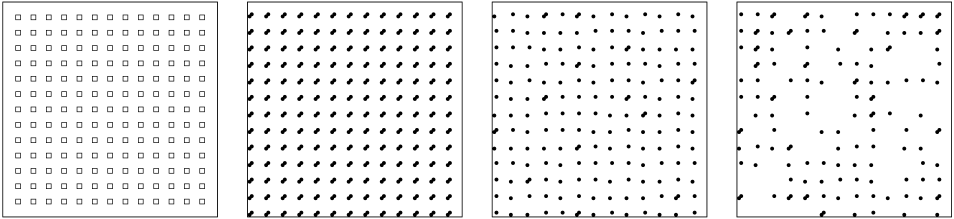
$$\sup_{\|f\|_{H(K)} \leq 1} \|f - S_{V,w_m}^{\boldsymbol{X}_n} f\|_{L_2}^2 \leq \frac{C}{(b-1)^3} \log\Big(\frac{m}{p}\Big)\Big(\sigma_{m+1}^2 + \frac{7}{m}\sum_{k=m+1}^{\infty} \sigma_k^2\Big)$$

with probability exceeding $1 - \frac{3}{2}p$, the weighted least squares operator $S_{V,w_m}^{\boldsymbol{X}_n}$, and $C = 23\,717 b^3$.

Numerics - sanity check

- Fourier matrix $\boldsymbol{Y} = [\exp(2\pi \mathrm{i}\langle \boldsymbol{k}, \boldsymbol{x}^i\rangle)]_{\boldsymbol{k}\in I, i=1,\dots,M}$.
- frequencies I on a grid
- initial points $\boldsymbol{x}^1, \dots, \boldsymbol{x}^M$ on two grids

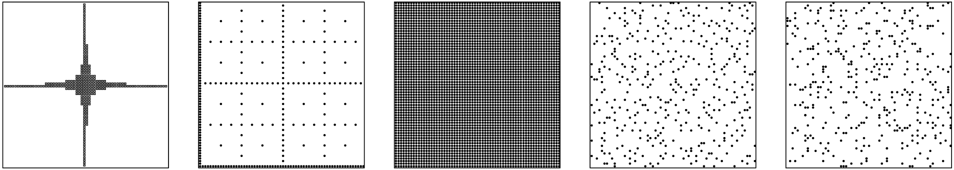
frequencies I	initial nodes	BSS subsampling	random subsampling



$m = 169$	$M = 338$ $(b = 2)$	$n = 182$ $(b \approx 1.1)$	$n = 182$ $(b \approx 1.1)$
	$A = 1$	$A = 0.49471$	$A = 0$
	$B = 1$	$B = 1.81720$	$B = 5.38234$

Numerics - hyperbolic cross ($d = 2$)

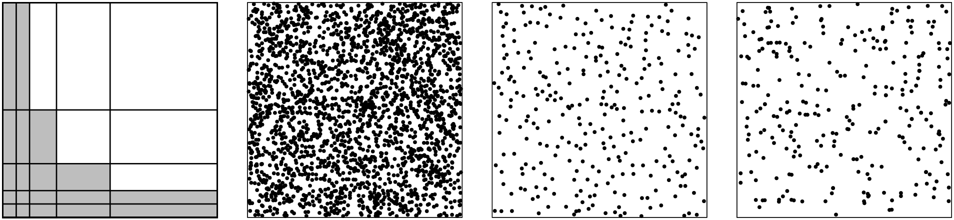
- Fourier matrix $\boldsymbol{Y} = [\exp(2\pi \mathrm{i} \langle \boldsymbol{k}, \boldsymbol{x}^i \rangle)]_{\boldsymbol{k} \in I, i=1,\dots,M}$.
- frequencies I on a hyperbolic cross
- initial points $\boldsymbol{x}^1, \dots, \boldsymbol{x}^M$ on a grids



frequencies I	sparse grid	initial nodes	BSS subsampling	random subsampling
$m = 256$	$M = 256$ $(b = 1)$	$M = 4225$ $(b \approx 16.5)$	$n = 384$ $(b \approx 1.5)$	$n = 384$ $(b \approx 1.5)$
	$A = 0.04336$	$A = 1$	$A = 0.06672$	$A = 0.02379$
	$B = 16$	$B = 1$	$B = 2.58239$	$B = 3.18402$

Numerics - wavelets

- Wavelet matrix $\boldsymbol{Y}$ with Chui-Wang Wavelets $\varphi_{\boldsymbol{j},\boldsymbol{k}}$
- initial points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M$ choosen randomly



I	initial nodes	BSS subsampling	random subsampling
$m = 192$	$M = 2400$ $(b \approx 12.5)$	$n = 288$ $(b \approx 1.5)$	$n = 288$ $(b \approx 1.5)$
	$A = 0.01708$	$A = 0.00321$	$A = 0.00030$
	$B = 1.01502$	$B = 1.05583$	$B = 1.13434$

Conclusion

- ☺ We extended the BSS algorithm to non-tight frames.
- ☺ We eliminated the weights s_i of the BSS algorithm sacrificing the upper frame bound.
- ☺ We have a constructive approach for finding points achieving **nearly** the optimal approximation error in L_2.

Julia-package: `www.github.com/felixbartel/BSSsubsampling.jl`