Sparse Harmonic Transforms: Best s-Term Approximation Guarantees for Bounded Orthonormal Product Bases in Sublinear-Time

{ Bosu Choi[a] , Mark Iwen[b], Toni Volkmer[c] } | *Contact: bosuchoi@gmail.com*

a) (former) Oden Institute for Computational Engineering and Sciences, UT Austin (b) Department of Mathematics and Department of Computational Mathematics, Science and Engineering, Michigan State Univ. (c) Department of Mathematics, Chemnitz University of Technology, Germany

Motivation

- Numerical problems often require fast function learning algorithms as well as integration and interpolation from finite function evalutations.
- Application: Uncertainty Quantification (UQ)
- Consider a parametric PDE and a quantity of interest (QoI) as a functional G on solution u:

$$A(\boldsymbol{y})u = g$$
$$f(\boldsymbol{y}) = Gu(\boldsymbol{y}).$$

- Goal: for solution $u = u(\boldsymbol{y})$ of the PDE, approximate QoI f as a function of $\boldsymbol{y} \in \mathcal{D} \subset \mathbb{R}^D$ with large D.
- Observation: QoI $f(\boldsymbol{y})$ is approximately sparse in appropriate (truncated) product basis. (e.g., Fourier, Chebyshev, Legendre, etc.)

Problem Setting

- Consider a function $f \in L^2(\mathcal{D}, \mu)$:

$$f(\boldsymbol{x}) := \underbrace{\sum_{\boldsymbol{n}\in\mathcal{I}_{N,d}} c_{\boldsymbol{n}} T_{\boldsymbol{n}}(\boldsymbol{x})}_{=:\tilde{f}} + \sum_{\boldsymbol{n}\in\mathbb{N}^D\setminus\mathcal{I}_{N,d}} c_{\boldsymbol{n}} T_{\boldsymbol{n}}(\boldsymbol{x})$$

  where $\mathcal{I}_{N,d} := \{\boldsymbol{n} \in [N]^D \mid \|\boldsymbol{n}\|_0 \leq d\}$, $d \leq D$, and $[N] := [0, 1, \cdots, N-1]$.
- $\mathcal{D} := \times_{j\in[D]} \mathcal{D}_j \subset \mathbb{R}^D$ is a space with measure $\boldsymbol{\mu} := \otimes_{j\in[D]} \mu_j$.
- Find the best s-term approximation $\tilde{f}_s^{\text{opt}}$ of $\tilde{f}$ where $\tilde{f}$ is a projection of f onto $\mathcal{B}_{N,d} := \left\{ T_{\boldsymbol{n}}(\boldsymbol{x}) := \prod_{j\in[D]} T_{j;n_j}(x_j) \mid \boldsymbol{n} \in \mathcal{I}_{N,d} \right\}$.
- $\mathcal{B}_{N,d}$ is a set of bounded orthonormal product bases (BOPB) with $K := \max_{\boldsymbol{n}\in\mathcal{I}_{N,d}} \|T_{\boldsymbol{n}}\|_\infty$.
- Goal : recover $\tilde{f}_s^{\text{opt}}$ rapidly and sample-efficiently.

Main Theorem (Sublinearized CoSaMP)

There exist a finite set of grid points $\mathcal{G} \subset \mathcal{D}$, an algorithm $\mathcal{H} : \mathbb{C}^{|\mathcal{G}|} \to (\mathcal{I}_{N,d} \times \mathbb{C})^s$, and an absolute universal constant $C \in \mathbb{R}^+$ s.t. the function $a : \mathcal{D} \to \mathbb{C}$ defined by $a(\boldsymbol{x}) := \sum_{(\boldsymbol{n},a_{\boldsymbol{n}})\in\mathcal{H}(f(\mathcal{G}))} a_{\boldsymbol{n}} T_{\boldsymbol{n}}(\boldsymbol{x})$ satisfies

$$\|f - a\|_{L^2(\mathcal{D},\mu)} \leq \left\|f - \tilde{f}\right\|_{L^2(\mathcal{D},\mu)} + C\left(\sqrt{s}\left\|\tilde{\boldsymbol{c}} - \tilde{\boldsymbol{c}}_s^{\text{opt}}\right\|_2 + \left\|\tilde{\boldsymbol{c}} - \tilde{\boldsymbol{c}}_s^{\text{opt}}\right\|_1 + \gamma\sqrt{s}\right) + \eta$$

for all $f = \sum_{\boldsymbol{n}\in\mathbb{N}^D} c_{\boldsymbol{n}} T_{\boldsymbol{n}}$ with $\left\|f - \tilde{f}\right\|_\infty =: \gamma < \infty$. The grid $\mathcal{G}$ has

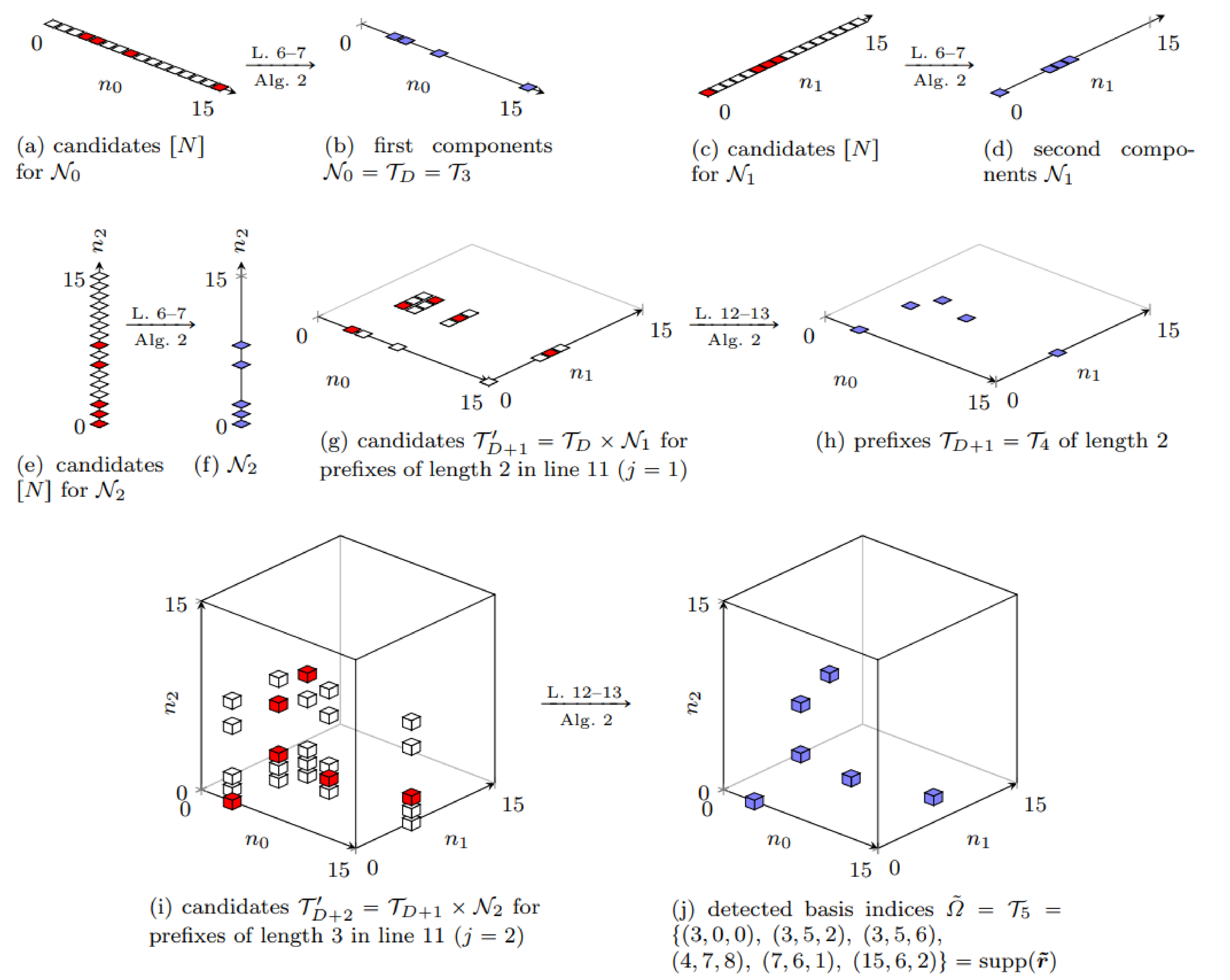
$$|\mathcal{G}| = \mathcal{O}\left(s^3 D K^4 d^4 \log^4(DN/d) \log^2(s) \log^2(D)\right)$$

and the algorithm $\mathcal{H}$ has runtime complexity

$$\mathcal{O}\left((s^5 + s^3 N) D^2 K^4 d^4 \log^4(DN/d) \log^2(s) \log^2(D) \log\left(\left\|\tilde{c}_s^{\text{opt}}\right\|_2 / \eta\right)\right)$$

Sublinearized Support Identification

- The following example illustrates the dimension incremental method for the sublinearized support identification to accelerate CoSaMP.
- $\tilde{\boldsymbol{r}} := \tilde{\boldsymbol{c}} - \boldsymbol{a}$ where $\boldsymbol{a} = \boldsymbol{a}^{t-1}$, approximation in $(t-1)$-th iteration.
- Residual function $h(\boldsymbol{x}) := \sum \tilde{r}_{\boldsymbol{n}} T_{\boldsymbol{n}}(\boldsymbol{x})$ are updated in each loop.



(a) candidates $[N]$ for $\mathcal{N}_0$ (b) first components $\mathcal{N}_0 = \mathcal{T}_D = \mathcal{T}_3$ (c) candidates $[N]$ for $\mathcal{N}_1$ (d) second components $\mathcal{N}_1$ (e) candidates $[N]$ for $\mathcal{N}_2$ (f) $\mathcal{N}_2$ (g) candidates $\mathcal{T}'_{D+1} = \mathcal{T}_D \times \mathcal{N}_1$ for prefixes of length 2 in line 11 ($j = 1$) (h) prefixes $\mathcal{T}_{D+1} = \mathcal{T}_4$ of length 2 (i) candidates $\mathcal{T}'_{D+2} = \mathcal{T}_{D+1} \times \mathcal{N}_2$ for prefixes of length 3 in line 11 ($j = 2$) (j) detected basis indices $\tilde{\Omega} = \mathcal{T}_5 = \{(3,0,0), (3,5,2), (3,5,6), (4,7,8), (7,6,1), (15,6,2)\} = \text{supp}(\tilde{\boldsymbol{r}})$

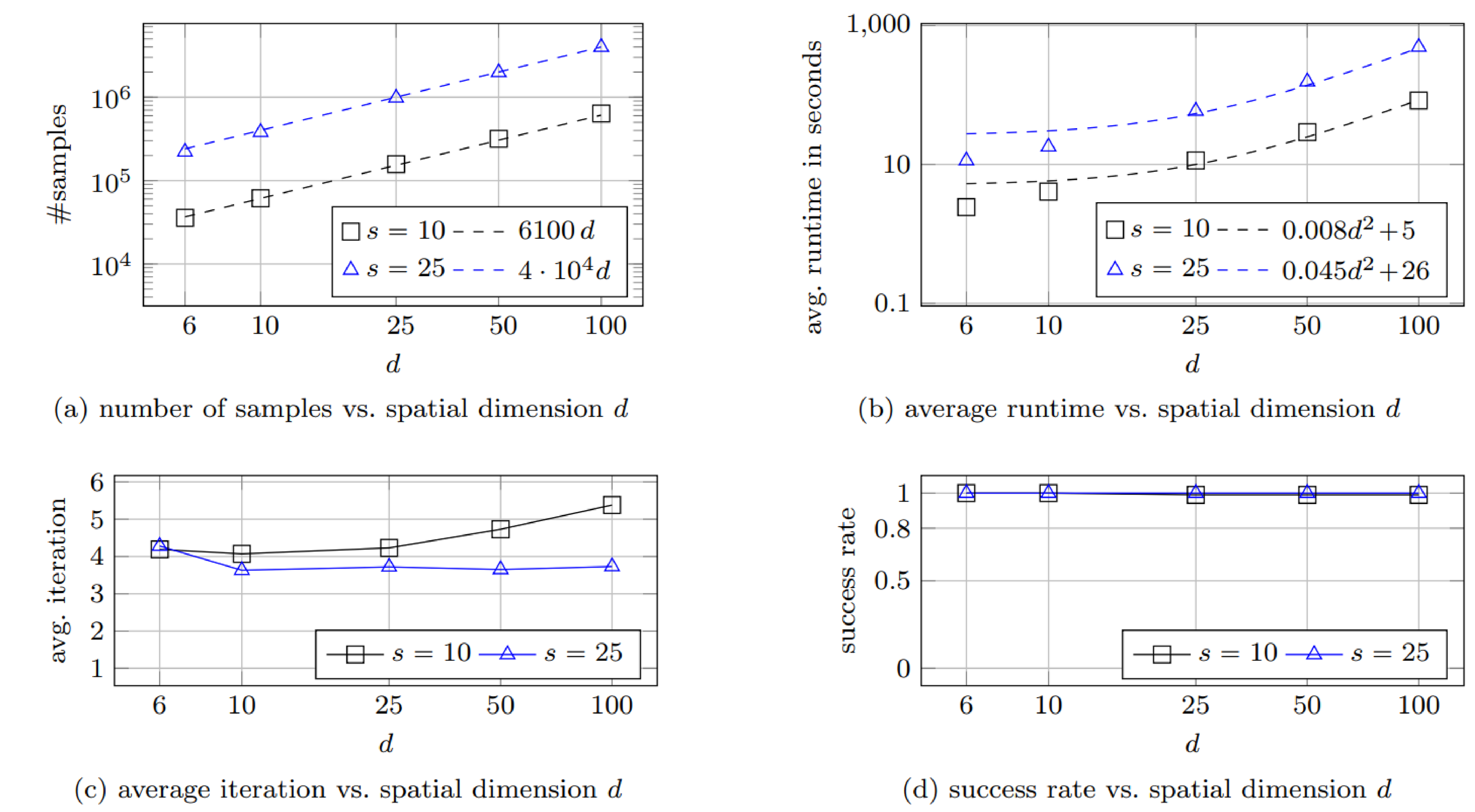
Fig. 1: Illustration of the major steps of Algorithm 2 for a simple 3-dimensional example with $s = 6$, $N = 16$, $d = D = 3$, and $\text{supp}(\tilde{\boldsymbol{r}}) = \{(3,0,0), (3,5,2), (3,5,6), (4,7,8), (7,6,1), (15,6,2)\} \subset \mathcal{I}_{16,3}$.

Pseudocode of New Support Identification

1: **procedure SuppID** (Alg. 2 in [5])
2: **Parameters:** N, D, s
3: **Input:** $\boldsymbol{v}_{\text{SID}} \in \mathbb{C}^{m_1 m_2(2D-1)}$ **containing** h**'s function evaluations split into** $2D-1$ **blocks**
4: **Output: support set** $\tilde{\Omega}$ **with** $|\tilde{\Omega}| \leq s$
5: **for** $j = 0 \to D-1$ **do**
6: Estimate $\frac{1}{m_2}\sum_{k\in[m_2]}\left|\frac{1}{m_1}\sum_{\ell\in[m_1]}(\boldsymbol{v}_{\text{SID},j})_{\ell,k}\overline{T_{j;n}(w_\ell^j)}\right|^2$ for each $n \in [N]$
7: $\mathcal{N}_j \leftarrow \{n \in [N] \mid \min(s, N) \text{ largest energy estimates}\}$
8: **end for**
9: $\mathcal{T}_D \leftarrow \mathcal{N}_0$
10: **for** $j = 1 \to D-1$ **do**
11: $\mathcal{T}'_{D+j} \leftarrow \mathcal{T}_{D+j-1} \times \mathcal{N}_j$:
12: Estimate for each $\boldsymbol{n} \in \mathcal{T}'_{D+j}$ $\frac{1}{m_2}\sum_{k\in[m_2]}\left|\frac{1}{m_1}\sum_{\ell\in[m_1]}(\boldsymbol{v}_{\text{SID},D-1+j})_{\ell,k}\overline{T_{j;n}(\boldsymbol{w}_\ell^{D-1+j})}\right|^2$
13: $\mathcal{T}_{D,j} \leftarrow \{\boldsymbol{n} \in \mathcal{T}'_{D+j} \mid \min(s, |\mathcal{T}'_{D+j}|) \text{ largest energy est.}\}$
14: **end for**
15: Return $\tilde{\Omega} \leftarrow \mathcal{T}_{2D-1}$
16: **end procedure**

Numerical Result

· Mixed BOPB for exactly sparse case using noisy samples



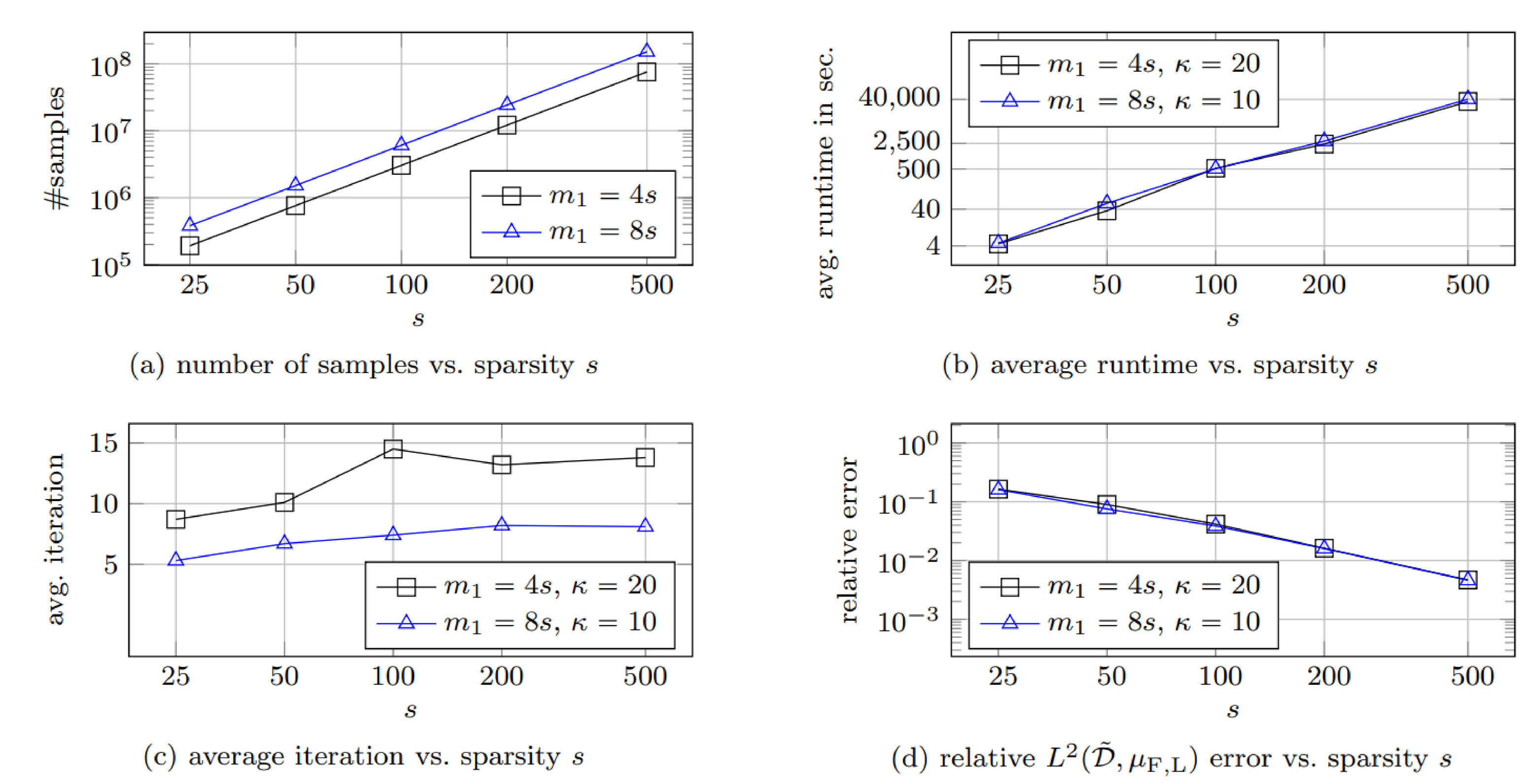
(a) number of samples vs. spatial dimension d (b) average runtime vs. spatial dimension d (c) average iteration vs. spatial dimension d (d) success rate vs. spatial dimension d

Fig. 2: Number of samples, runtime, number of iterations, success rates vs. spatial dimension $d = D \in \{6, 10, 25, 50, 100\}$ for mixed bases (3 Chebyshev, 3 Legendre, $d-6$ Fourier), $N = 200$, sparsity $s \in \{10, 25\}$, $\text{SNR}_{\text{db}} = 10$, $m_1 = 8s$, $m_2 = 4s$.

· Mixed BOPB for approx. sparse case with the following test function:

$$f(\boldsymbol{x}) = B_3(x_1)B_3(x_3)N_4(x_9) + B_5(x_4)B_5(x_5)N_2(x_2)N_2(x_7) + B_3(x_8)N_2(x_6)N_2(x_{10})$$

N_m: periodic B-spline of order m
B_m: (nonperiodic) shifted and dialated B-spline of order m



(a) number of samples vs. sparsity s (b) average runtime vs. sparsity s (c) average iteration vs. sparsity s (d) relative $L^2(\tilde{\mathcal{D}}, \mu_{\text{F,L}})$ error vs. sparsity s

Fig. 3: Number of samples, runtime, number of iterations, relative $L^2(\tilde{\mathcal{D}}, \mu_{\text{F,L}})$ error vs. sparsity $s \in \{25, 50, 100, 200\}$ for mixed Fourier+Legendre basis and test function.

References

1 D. Needell and J. A. Tropp. "*CoSaMP : Iterative signal recovery from incomplete and inaccurate samples*", Applied and Computational Harmonic Analysis, 2009
2 R. Holger and C. Schwab. "*Compressive sensing Petrov-Galerkin approximation of high-dimensional parametric operator equations*", Mathematics of Computation, 2017
3 B. Choi, A. Christlieb, and Y. Wang. "*High-dimensional Sparse Fourier Algorithms*", Numerical Algorithms, 2020
4 B. Choi, A. Christlieb, and Y. Wang. "*Multiscale High-dimensional Sparse Fourier Algorithms for Noisy Data*", Mathematics, Computation and Geometry of Data, 2020, to appear
5 B. Choi, M. Iwen, T. Volkmer. "Sparse Harmonic Transforms II: Best s-Term Approximation Guarantees for Bounded Orthonormal Product Bases in Sublinear-Time", 2020, arXiv:1909.09564
6 B. Choi, M. Iwen, F. Krahmer. "Sparse Harmonic Transforms: A New Class of Sublinear-time Algorithms for Learning Functions of Many Variables", Foundations of Computational Mathematics, 2020