Butterfly Sparse Fast Fourier Transform

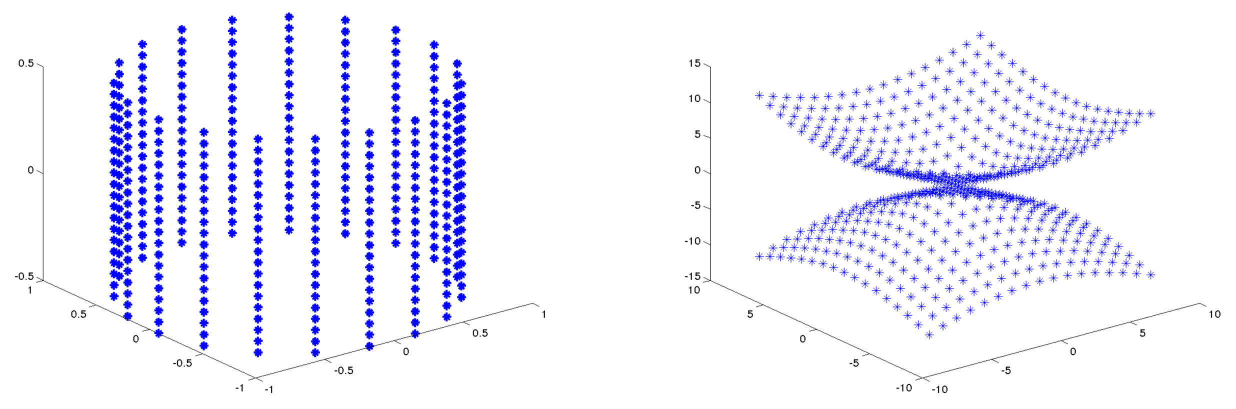
Stefan Kunis and Ines Melzer - School of Mathematics/Computer Science, Osnabrück University

1. Problem

given:

- dimension $d \in \mathbb{N}$
- spatial domain $X = [0, N]^d$, $N = 2^L$, $L \in \mathbb{N}$
- frequency domain $T = [0, N]^d$
- sampling nodes on $d-1$ dimensional smooth manifolds

$$\widetilde{X} := \{\mathbf{x}_i \in X : i = 1, \ldots, M_1\}, \quad \widetilde{T} := \{\boldsymbol{\xi}_j \in T : j = 1, \ldots, M_2\},$$



Example for sampling nodes on a cylinder or on a cone in the 3-dimensional case.

- Fourier coefficients

$$\hat{f}_j \in \mathbb{C}, \quad j = 1, \ldots, M_2$$

task: Compute

$$f(\mathbf{x}) = \sum_{j=1}^{M_2} \hat{f}_j e^{2\pi i \boldsymbol{\xi}_j \cdot \mathbf{x}/N}, \quad \mathbf{x} \in \widetilde{X}. \tag{1}$$

- $\mathcal{O}(M_1 M_2)$ arithmetic operations, **Faster computation?**

2. Lowrank approximation

Theorem 1. Let $N, p \in \mathbb{N}$, $p \geq 5$, and let $T, X \subset [0, N]^d$ fulfill the admissibility condition

$$\text{diam}(T)\text{diam}(X) \leq N.$$

Moreover, let the function $g(x) = \sum_{\xi \in T'} \hat{g}_k e^{2\pi i \boldsymbol{\xi} \mathbf{x}/N}$, $T' \subset T$, be approximated by using p^d points $T_p \subset T$ on a regular tensor product grid. The interpolation $I_p^{XT} g(\mathbf{x_s}) = g(\mathbf{x_s})$ at p^d nodes $\mathbf{x_s} \in X$ on the Chebyshev tensor product grid then yields

$$\|g - I_p^{XT} g\|_{C(X)} \leq \frac{(1 + C_p)(C_p^d - 1)}{C_p - 1} \cdot c_p \cdot \|\hat{\mathbf{g}}\|_1,$$

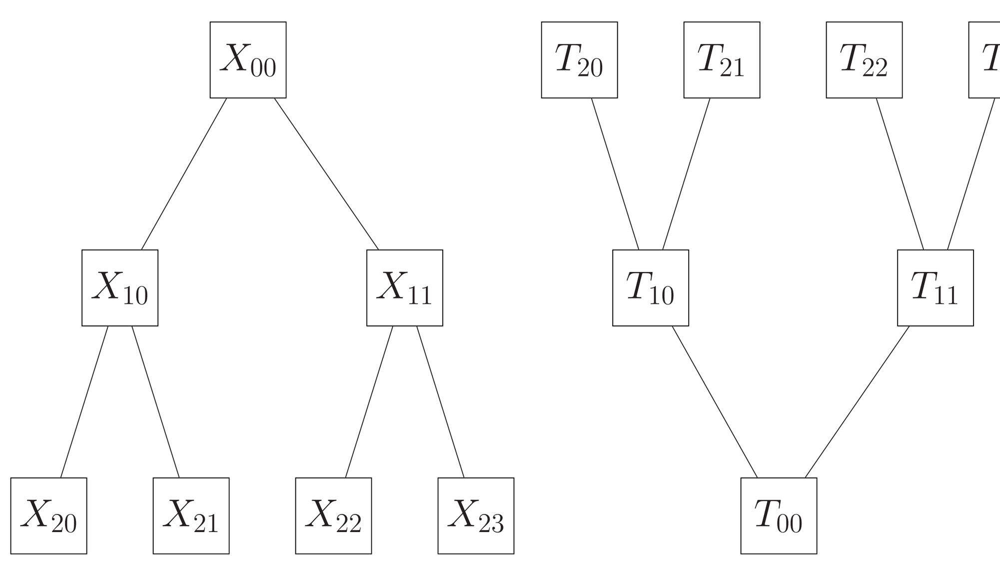
with $c_p := \frac{1}{\pi p}\left(\frac{\pi}{p-1}\right)^p$, $\quad C_p := \sqrt{K_p}\left(1 + \frac{2}{\pi}\log p\right)$, $\quad \frac{\pi^4}{16} \geq K_p \searrow 1$. □

3. Dyadic decomposition

The following ideas are presented for $d = 1$. For $N = 2^L$, $L \in \mathbb{N}$, the dyadic decompositions

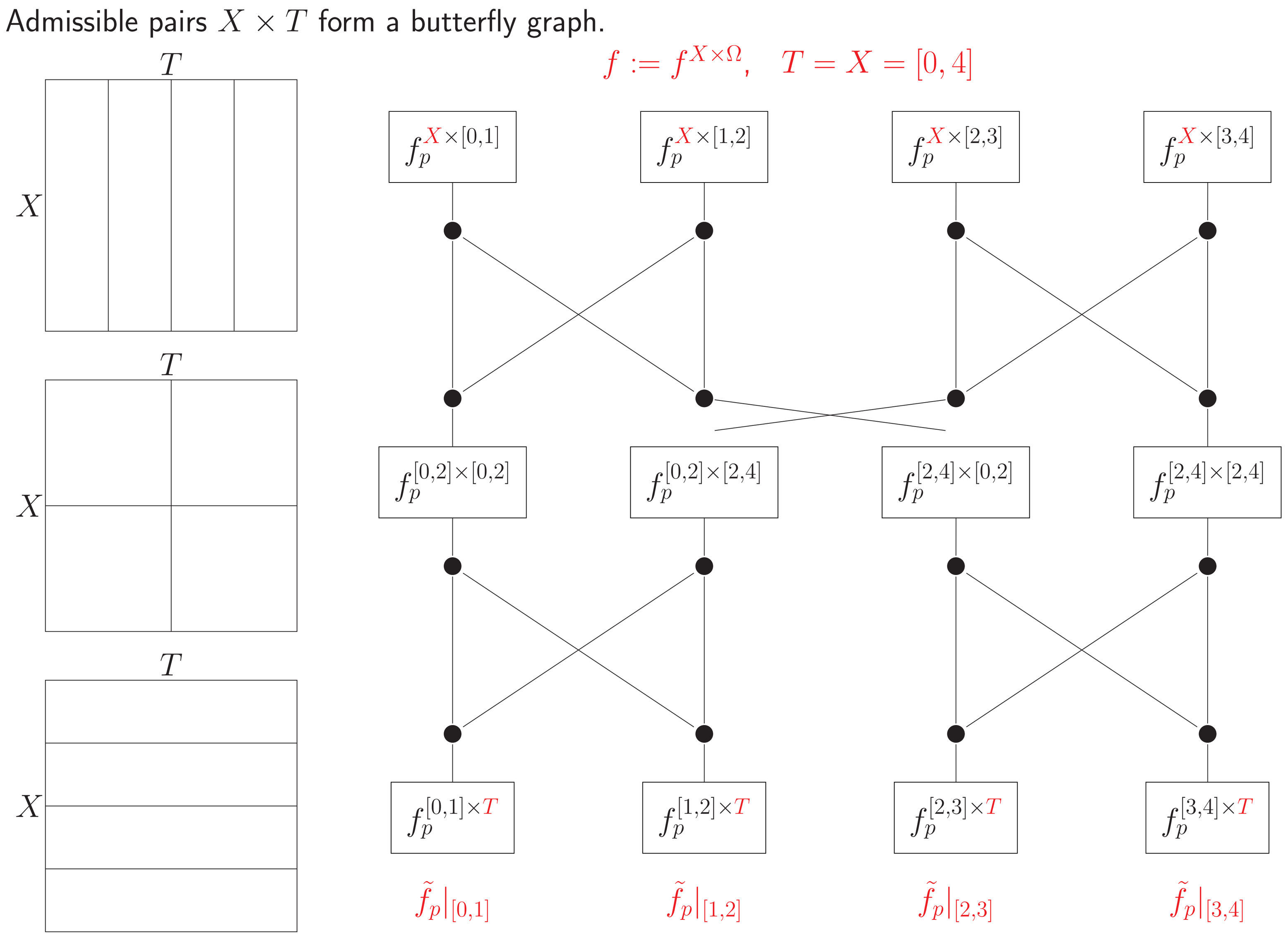
$$X_{l,m} = \left[N/2^l m, N/2^l (m+1)\right) \quad \text{for } m = 0, \ldots, 2^l - 1,$$
$$T_{L-l,n} = \left[N/2^{L-l} n, N/2^{L-l}(n+1)\right) \quad \text{for } n = 0, \ldots, 2^{L-l} - 1,$$

with levels $l = 0, \ldots, L$ and locations m, n, form two binary trees.



4. Butterfly Scheme

Admissible pairs $X \times T$ form a butterfly graph.



5. Algorithm

The butterfly scheme traverses the graph top down. Starting locally over T, we build approximations from its two predecessors, include more frequencies and restrict to smaller X-boxes.

Algorithm (BSFFT).

1. exact functions, local in frequency- and global in spatial domain

$$f^{T_{L,n}}(x) := \sum_{\xi_j \in T_{L,n} \cap \widetilde{T}} \hat{f}_j e^{2\pi i \xi_j x/N}, \quad n = 0, \ldots, 2^L - 1$$

level $\ell = 0$:

$$f_p^{X \times T_{L,n}} := I_p^{X T_{L,n}} f^{T_{L,n}}, \quad n = 0, \ldots, 2^L - 1$$

2. level $\ell = 1, \ldots, L$

$$f_p^{X_{\ell,m} T_{L-\ell,n}} := I_p^{X_{\ell,m} T_{L-\ell,n}} \left[f_p^{X_{\ell-1,\lfloor m/2 \rfloor} \times T_{L-\ell+1,2n}} + f_p^{X_{\ell-1,\lfloor m/2 \rfloor} \times T_{L-\ell+1,2n+1}} \right]$$

for $m = 0, \ldots, 2^\ell - 1$ and $n = 0, \ldots, 2^{L-\ell} - 1$

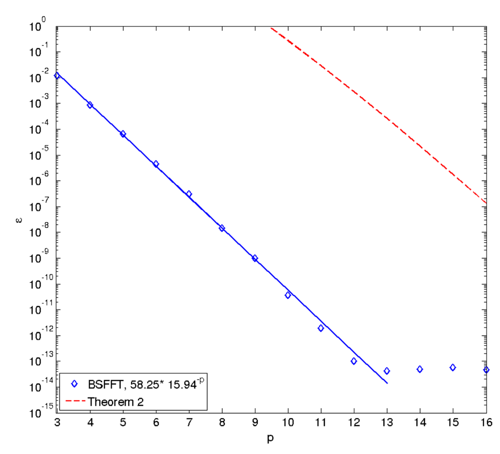
3. finally

$$\tilde{f}_p(x) = f_p^{X_{L,m} T}(x) \text{ for } x \in X_{L,m}, \quad m = 0, \ldots, 2^L - 1$$

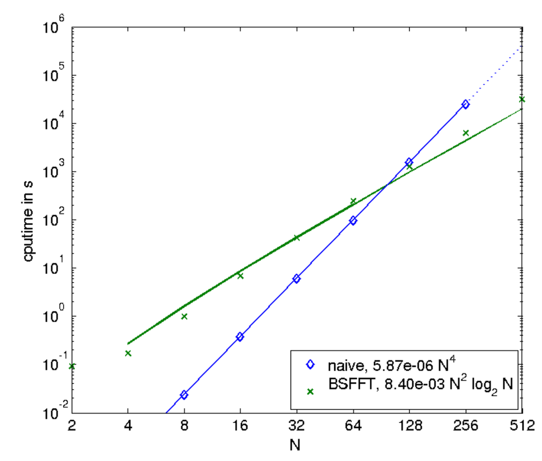
is an approximation to f in $X_{L,m}$

6. Error Analysis

Theorem 2. Let $L, p \in \mathbb{N}$, $N = 2^L$, $T, X \subset [0, N]^d$, $\varepsilon \in (0, 1]$, and $p \geq C \log L/\varepsilon$. The butterfly approximation $\tilde{f}_p$ fulfills $\|f - \tilde{f}_p\|_\infty \leq \varepsilon \|\hat{\mathbf{f}}\|_1$. In particular, the computation of (1) takes $\mathcal{O}(N^{d-1}(\log N)^{d+2})$ floating point operations for fixed target accuracy. □



: relative error ε for $d = 1$ in dependence on p



: cputime in s for $d = 3$ for fixed $p = 8$ in dependence on N

7. Outlook

Can we compute sums of the form

$$f(z) := \sum_{k=1}^{M_2} \hat{f}_k e^{2\pi i \xi_k z/N},$$

where $\xi_k \in T$ and $z \in X \times iY$, $Y \subset \mathbb{R}_+$ fast?

First idea. Factorize $e^{2\pi i \xi_k z/N} = e^{2\pi i \xi_k x/N} e^{-2\pi \xi_k y/N}$ and use a combination of the BSFFT and the Fast Laplace transform, [3]. The main idea of the Laplace transform is the low rank approximation of the exponential kernel $\kappa : Y \times \Omega \to \mathbb{R}$, $\kappa(y, \xi) = e^{-y\xi}$, by interpolation in Chebyshev nodes.

Theorem 3. Let $p \in \mathbb{N}$, $p \geq 2$, and $Y, \Omega \subset \mathbb{R}_+$ be admissible in the sense, that $\operatorname{diam} Y \leq \operatorname{dist}(0, Y)$ and $\operatorname{diam} \Omega \leq \operatorname{dist}(0, \Omega)$ is fullfilled. The usual interpolation of κ in Chebyshev-nodes y_s^Y in Y and ξ_r^Ω in Ω is defined by

$$(\mathcal{I}_p^Y \otimes \mathcal{I}_p^\Omega)\kappa(y, \xi) := \sum_{s=0}^{p-1} L_s^Y(y) \sum_{r=0}^{p-1} L_r^\Omega(\xi) \kappa(y_s^Y, \xi_r^\Omega),$$

where L_s^Y and L_r^Ω are the appropriate Lagrange polynomials. It yields

$$\|(\mathcal{I}_p^Y \otimes \mathcal{I}_p^\Omega)\kappa - \kappa\|_{C(Y \times \Omega)} \leq \frac{4^{-p}}{\sqrt{2\pi p}}\left(2 + \frac{2}{\pi}\log p\right). \quad \square$$

For admissible boxes Y and Ω the approximation of the sum

$$u(x_j, y_j) := \sum_{k=1}^{n} \hat{f}_k e^{2\pi i x_j \xi_k/N} \kappa(y_j, \xi_k) \approx \sum_{k=1}^{n} \hat{f}_k e^{2\pi i x_j \xi_k/N} (\mathcal{I}_p^Y \otimes \mathcal{I}_p^\Omega)\kappa(y_j, \xi_k)$$

leads to

$$u(x_j, y_j) \approx \sum_{s=0}^{p-1} L_s^Y(y_j) \sum_{k=1}^{n} \hat{g}_k e^{2\pi i x_j \xi_k/N}$$

with coefficients $\hat{g}_k := \hat{f}_k \sum_{r=0}^{p-1} L_r^\Omega(\xi_k)\kappa(y_s^Y, \xi_r^\Omega)$. The inner sum can be computed approximatively by the BSFFT.

8. References

[1] S. Kunis and I. Melzer.
BSFFT, Matlab toolbox for computing the butterfly sparse fast Fourier transform.
http://www.analysis.uos.de/software/, 2012.

[2] S. Kunis and I. Melzer.
A stable and accurate butterfly sparse Fourier transform.
SIAM J. Numer. Anal., 50(3):1777–1800, 2012.

[3] V. Rokhlin.
A fast algorithm for the discrete Laplace transformation.
J. Complexity, 4(1):12–32, 1988.

[4] L. Ying.
Sparse Fourier transform via butterfly algorithm.
SIAM J. Sci. Comput., 31(3):1678–1694, 2009.