Chapter 1
Computational Methods for the Fourier Analysis of Sparse High-Dimensional Functions

Lutz Kämmerer, Stefan Kunis, Ines Melzer, Daniel Potts, and Toni Volkmer

Abstract A straightforward discretisation of high-dimensional problems often leads to a curse of dimensions and thus the use of sparsity has become a popular tool. Efficient algorithms like the fast Fourier transform (FFT) have to be customised to these thinner discretisations and we focus on two major topics regarding the Fourier analysis of high-dimensional functions: We present stable and effective algorithms for the fast evaluation and reconstruction of trigonometric polynomials with frequencies supported on an index set $\mathcal{I} \subset \mathbb{Z}^d$.

1.1 Introduction

Let $d \in \mathbb{N}$ be the spatial dimension and $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d \simeq [0,1)^d$ denote the torus. We consider trigonometric polynomials $f : \mathbb{T}^d \to \mathbb{C}$ with Fourier coefficients $\hat{f}_{\mathbf{k}} \in \mathbb{C}$ supported on the frequency index set $\mathcal{I} \subset \mathbb{Z}^d$ of finite cardinality. The evaluation of the trigonometric polynomial

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{I}} \hat{f}_{\mathbf{k}} \, \mathrm{e}^{2\pi \mathrm{i} \mathbf{k} \cdot \mathbf{x}} \tag{1.1}$$

at a sampling set $\mathcal{X} \subset \mathbb{T}^d$ of finite cardinality can be written as the matrix-vector product

Stefan Kunis*, Ines Melzer
Institut für Mathematik, Universität Osnabrück, *and Institute of Computational Biology, Helmholtz Zentrum München, Germany
e-mail: {stefan.kunis,ines.melzer}@uos.de

Lutz Kämmerer and Daniel Potts and Toni Volkmer
Fakultät für Mathematik, Technische Universität Chemnitz, Germany
e-mail: {lutz.kaemmerer,daniel.potts,toni.volkmer}@mathematik.tu-chemnitz.de

$$\mathbf{f} = \mathbf{A}\,\hat{\mathbf{f}}, \quad \mathbf{f} = (f(\mathbf{x}))_{\mathbf{x}\in\mathcal{X}} \in \mathbb{C}^{|\mathcal{X}|},\ \hat{\mathbf{f}} = (\hat{f}_{\mathbf{k}})_{\mathbf{k}\in\mathcal{I}} \in \mathbb{C}^{|\mathcal{I}|}, \tag{1.2}$$

with the Fourier matrix $\mathbf{A} = \mathbf{A}(\mathcal{I},\mathcal{X}) = \left(\mathrm{e}^{2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\right)_{\mathbf{x}\in\mathcal{X},\mathbf{k}\in\mathcal{I}} \in \mathbb{C}^{|\mathcal{X}|\times|\mathcal{I}|}$.

We are interested in the following two problems:

1. Evaluation: given a support $\mathcal{I} \subset \mathbb{Z}^d$, Fourier coefficients $\hat{f}_{\mathbf{k}} \in \mathbb{C}$, $\mathbf{k} \in \mathcal{I}$, and sampling nodes $\mathcal{X} = \{\mathbf{x}_\ell \in \mathbb{T}^d : \ell = 0,\dots,L-1\}$, evaluate the trigonometric polynomial (1.1) efficiently, i.e., compute $\mathbf{f} = \mathbf{A}\hat{\mathbf{f}}$ by means of a fast algorithm,
2. Reconstruction: given a support of Fourier coefficients $\mathcal{I} \subset \mathbb{Z}^d$, construct a set of sampling nodes $\mathcal{X} \subset \mathbb{T}^d$ with small cardinality $L = |\mathcal{X}|$ which allows for the unique and stable reconstruction of all trigonometric polynomials (1.1) from their sampling values $f(\mathbf{x}_\ell)$. In particular, solve the system of linear equations $\mathbf{A}\hat{\mathbf{f}} \approx \mathbf{f}$.

As an extension to the reconstruction problem, the efficient approximate reconstruction of a smooth function from subspaces of the Wiener algebra by a trigonometric polynomial (1.1), which guarantees a good approximation to the function, was considered in [37].

1.2 Evaluation of trigonometric polynomials

Clearly, the straightforward evaluation of the trigonometric polynomial (1.1) in all sampling nodes $\mathcal{X} \subset \mathbb{T}^d$, or equivalently the matrix vector multiplication (1.2), takes $\mathcal{O}(|\mathcal{X}|\cdot|\mathcal{I}|)$ floating point operations. For special index sets faster algorithms have been constructed as detailed subsequently.

1.2.1 FFT and nonequispaced FFT

We consider trigonometric polynomials supported on the full grid, i.e., with Fourier coefficients $\hat{f}_{\mathbf{k}}$ are defined on the full d-dimensional set $\mathcal{I} := \hat{G}_n^d = \mathbb{Z}^d \cap \bigtimes_{j=1}^d (-2^{n-1}, 2^{n-1}]$ of refinement $n \in \mathbb{N}$ and bandwidth $N = 2^n$ with the cardinality $|\mathcal{I}| = N^d$. The evaluation of the trigonometric polynomial

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\hat{G}_n^d} \hat{f}_{\mathbf{k}}\,\mathrm{e}^{2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{x}} \tag{1.3}$$

at all sampling nodes of an equispaced grid $\mathbf{x} \in \mathcal{X} = (2^{-n}\hat{G}_n^d \bmod \mathbf{1})$, with the cardinality $|\mathcal{X}| = N^d$, requires only $\mathcal{O}(2^{nd}n) = \mathcal{O}(N^d \log N)$ by the famous fast Fourier transform (FFT). A well understood generalisation considers an arbitrary sampling set $\mathcal{X} = \{\mathbf{x}_\ell \in \mathbb{T}^d :\ \ell = 0,\dots,L-1\}$ and leads to the so-

called nonequispaced FFT which takes $\mathcal{O}(2^{nd}n + |\log\varepsilon|^d L) = \mathcal{O}(N^d \log N + |\log\varepsilon|^d L)$ floating point operations for a target accuracy $\varepsilon > 0$, see e.g. [16, 5, 59, 51, 39] and the references therein. In both cases, already the huge cardinality of the support $\hat{G}_n^d$ of the Fourier coefficients $\hat{f}_{\mathbf{k}}$ causes immense computational costs for high dimensions d even for moderate refinement n. Hence, we restrict the index set $\mathcal{I}$ to smaller sets.

1.2.2 Hyperbolic cross FFT

Functions of dominating mixed smoothness can be well approximated by trigonometric polynomials supported on reduced frequency index sets, so called dyadic hyperbolic crosses

$$\mathcal{I} = H_n^d := \bigcup_{\substack{\mathbf{j}\in\mathbb{N}_0^d \\ \|\mathbf{j}\|_1 = n}} \left(\mathbb{Z}^d \cap \bigtimes_{l=1}^{d} (-2^{j_l-1}, 2^{j_l-1}] \right)$$

of dimension d and refinement n, cf. [57]. Compared to the trigonometric polynomial in (1.3), we strongly reduce the number of used Fourier coefficients $|H_n^d| = \mathcal{O}(2^n n^{d-1}) \ll 2^{nd}$. A natural spatial discretisation of trigonometric polynomials supported on the dyadic hyperbolic cross H_n^d is given by the sparse grid

$$\mathcal{X} = S_n^d := \bigcup_{\substack{\mathbf{j}\in\mathbb{N}_0^d \\ \|\mathbf{j}\|_1 = n}} \bigtimes_{l=1}^{d} 2^{-j_l}(\mathbb{N}_0 \cap [0, 2^{j_l})).$$

The cardinalities of the sparse grid and the dyadic hyperbolic cross are $|S_n^d| = |H_n^d| = \mathcal{O}(2^n n^{d-1})$. Fig. 1.1a(left) shows an example for a two-dimensional dyadic hyperbolic cross and Fig. 1.1a(right) depicts the corresponding sparse grid of identical cardinality. Based on [3, 27] there exists a fast algorithm for evaluating the trigonometric polynomial with frequencies supported on the hyperbolic cross H_n^d at all $\mathbf{x} \in S_n^d$ in $\mathcal{O}(2^n n^d)$ floating point operations, called hyperbolic cross fast Fourier transform (HCFFT). A generalisation to sparser index set, i.e., to index sets for so called energy-norm based hyperbolic crosses, is presented in [22].

1.2.3 Lattice and generated set FFT

Using lattices as sampling set $\mathcal{X}$ is motivated from the numerical integration of functions of many variables by lattice rules, see [58, 47, 14] for an introduction. In contrast to general lattices which may be spanned by several vectors,

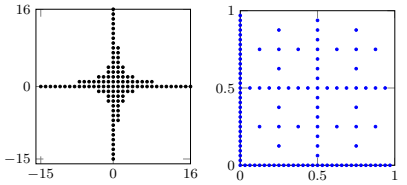

Fig. 1.1a Dyadic hyperbolic cross H_5^2 (left) and sparse grid S_5^2 (right).

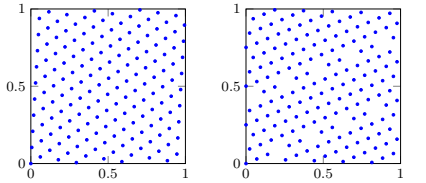

Fig. 1.1b Rank-1 lattice (left) and generated set $\Lambda(\mathbf{r}, M)$ (right), $M = 163$.

we only consider so-called rank-1 lattices and a generalisation of this concept called generated sets [32]. For a given number $L \in \mathbb{N}$ of sampling nodes and a generating vector $\mathbf{r} \in \mathbb{R}^d$, we define the generated set

$$\mathcal{X} = \Lambda(\mathbf{r}, L) := \{\mathbf{x}_\ell = \ell\mathbf{r} \bmod \mathbf{1},\ \ell = 0, \dots, L-1\} \subset \mathbb{T}^d.$$

For $\ell = 0, \dots, L-1$, the evaluation of a d-variate trigonometric polynomial supported on an arbitrary frequency index set $\mathcal{I}$ simplifies dramatically since

$$f(\mathbf{x}_\ell) = \sum_{\mathbf{k}\in\mathcal{I}} \hat{f}_\mathbf{k}\, \mathrm{e}^{2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{x}_\ell} = \sum_{\mathbf{k}\in\mathcal{I}} \hat{f}_\mathbf{k}\, \mathrm{e}^{2\pi\mathrm{i}\ell\mathbf{k}\cdot\mathbf{r}} = \sum_{y\in\mathcal{Y}} \hat{g}_y\, \mathrm{e}^{2\pi\mathrm{i}\ell y}, \tag{1.4}$$

with some set $\mathcal{Y} = \{\mathbf{k}\cdot\mathbf{r} \bmod 1 : \mathbf{k} \in \mathcal{I}\} \subset \mathbb{T}$ and the aliased coefficients

$$\hat{g}_y = \sum_{\mathbf{k}\cdot\mathbf{r}\equiv y \pmod 1} \hat{f}_\mathbf{k}. \tag{1.5}$$

Using a one-dimensional adjoint nonequispaced FFT [39], this takes $\mathcal{O}(L\log L + (d + |\log\varepsilon|)|\mathcal{I}|)$ floating point operations for a target accuracy $\varepsilon > 0$. Moreover, given $L \in \mathbb{N}$ and a generating vector $\mathbf{r} = \mathbf{z}/L$, $\mathbf{z} \in \mathbb{Z}^d$, the sampling scheme $\Lambda(\mathbf{r}, L)$ is called rank-1 lattice and the computational costs of the evaluation reduce to $\mathcal{O}(L\log L + d|\mathcal{I}|)$ by applying a one dimensional FFT. We stress on the fact that in both cases, the computational costs only depend on the number L of samples subsequent to the aliasing step (1.5) which takes $d|\mathcal{I}|$ floating point operations. Fig. 1.1b(left) and Fig. 1.1b(right) show an example for a two-dimensional rank-1 lattice and generated set, respectively.

1.2.4 Butterfly sparse FFT

Another generalisation of the classical FFT to nonequispaced nodes has been suggested in [1, 62, 40]. While the above mentioned nonequispaced FFT still relies on a equispaced FFT, the so-called butterfly scheme only relies on

local low rank approximations of the complex exponentials - in particular this locality allows for its application to sparse data. The idea of local low rank approximations can be traced back at least to [21, 63, 4, 26] for smooth kernel functions and to [45, 64, 48, 61, 13] for oscillatory kernels. In a linear algebra setting, it was pointed out in [17] that certain blocks of the Fourier matrix are approximately of low rank.

We consider real frequencies $\mathcal{I} \subset [0, 2^n)^d$ and nonequispaced evaluation nodes $\mathbf{x}_\ell \in \mathcal{X} \subset [0, 1)^d$ in

$$f(\mathbf{x}_\ell) = \sum_{\mathbf{k}\in\mathcal{I}} \hat{f}_\mathbf{k}\, \mathrm{e}^{2\pi \mathrm{i}\mathbf{k}\cdot\mathbf{x}_\ell}, \quad \ell = 0, \ldots, L-1. \tag{1.6}$$

For ease of notation, we outline the main idea for the one-dimensional case. We decompose both domains dyadically starting with the whole interval $[0, 2^n)$ and $[0, 1)$ as root, respectively, see also Fig. 1.2(left) and 1.2(middle). Each pair of a frequency interval in the $(n-j)$-th level and a space interval in the j-th level now fulfils the admissibility condition $\mathrm{diam}(\mathcal{I}')\mathrm{diam}(\mathcal{X}') \leq 1$. These pairs are depicted in Fig. 1.2(right), where an edge in this butterfly graph is set if and only if the associated pairs of intervals are connected in both trees. We note that the properly frequency shifted exponential function

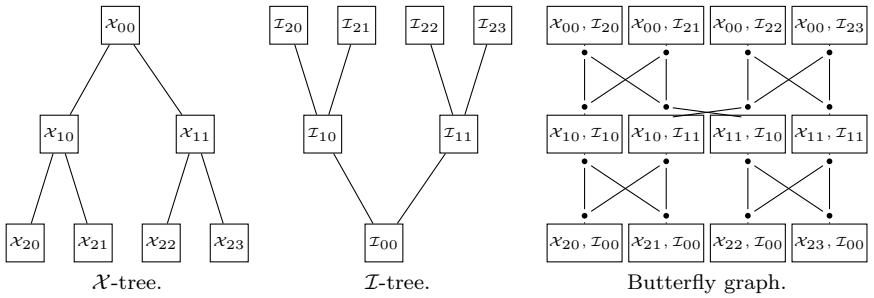

Fig. 1.2 Trees and butterfly graph for $N = 4$.

is a smooth function within the admissible region and can be well approximated by a trigonometric sum with equispaced frequencies interpolating in Chebyshev nodes, see [40, Thm. 2.6] for details.

The generalisation to spatial dimension $d \geq 2$ is straightforward by decomposing $\mathcal{I} \subset [0, 2^n)^d$ and $\mathcal{X} \subset [0, 1)^d$ dyadically in each coordinate, using a tensor product ansatz, and interpolate in a product grid. The butterfly scheme now traverses the butterfly graph top down. We start in the zeroth level, sum frequencies in the finest decomposition, and approximate on the whole spatial domain. In each subsequent level, we sum up two predecessors including more frequencies and approximate on each smaller spatial box. The final approximation is a function piecewise defined on the finest spatial decomposition. The butterfly scheme guarantees the following target accuracy.

Theorem 1.1. *([40, Thm. 3.1]). Let $d, n, p \in \mathbb{N}$, $p \geq 5$, $\mathcal{I} \subset [0, 2^n)^d$, $\mathcal{X} \subset [0, 1)^d$, and the trigonometric sum f as in (1.6), then the butterfly approximation g obeys the error estimate*

$$\|f - g\|_\infty \leq \frac{(C_p + 1)(C_p^{d(n+1)} - 1)}{C_p - 1} c_p \|\hat{\mathbf{f}}\|_1.$$

The constants are explicitly given by

$$K_p := \left(\frac{2\pi^2}{\left(1 - \cos \frac{2\pi}{p-1}\right)(p-1)^2} \right)^{p-1}, \quad K_p \leq \frac{\pi^4}{16}, \quad \lim_{p \to \infty} K_p = 1,$$
$$C_p := \sqrt{K_p} \left(1 + \tfrac{2}{\pi} \log p\right), \qquad c_p := \frac{1}{\pi p} \left(\frac{\pi}{p-1} \right)^p.$$

In particular, the butterfly scheme achieves relative error at most ε if the local expansion degree fulfils $p \geq \max\{10, 2|\log \varepsilon|, 2d(n+1)\}$.

In case $1 \leq t < d$ and $|\mathcal{X}| = |\mathcal{I}| = 2^{nt}$ well distributed sets on smooth t-dimensional manifolds, the dyadic decompositions of the sets remain sparse. Consequently, the butterfly graph, which represents the admissible pairs where computations are performed, remains sparse as well and the computation of (1.6) takes $\mathcal{O}(2^{nt} n (n + |\log \varepsilon|)^{d+1})$ floating point operations only.

1.3 Reconstruction of trigonometric polynomials

Beyond the fast evaluation of Fourier expansions, the sampling problem is concerned with the recovery of the Fourier coefficients $\hat{f}_{\mathbf{k}} \in \mathbb{C}$ from a sequence of function samples. This inverse transform constructs a trigonometric polynomial (1.1) such that for given data points $(\mathbf{x}_\ell,\, f_\ell) \in \mathbb{T}^d \times \mathbb{C}$, $\ell = 0, \ldots, L-1$, the approximate identity

$$f(\mathbf{x}_\ell) \approx f_\ell$$

is fulfilled. Thus, we aim to solve the linear system of equations $\mathbf{A}\hat{\mathbf{f}} \approx \mathbf{f}$ for the vector of Fourier coefficients $\hat{\mathbf{f}} \in \mathbb{C}^{|\mathcal{I}|}$. In contrast to the ordinary Fourier matrix, its generalized analogue $\mathbf{A}$ is in general neither unitary nor square. The meaningful variants of this reconstruction problem include

1. the weighted least squares approximation

$$\|\mathbf{f} - \mathbf{A}\hat{\mathbf{f}}\|_{\mathbf{W}}^2 = \sum_{\ell=0}^{L-1} w_\ell |f_\ell - f(\mathbf{x}_\ell)|^2 \xrightarrow{\hat{\mathbf{f}}} \min, \tag{1.7}$$

   for the over-determined case $|\mathcal{I}| < |\mathcal{X}|$, where the weights w_ℓ compensate for clusters in the sampling set,
2. the optimal interpolation problem

$$\|\hat{\mathbf{f}}\|^2_{\hat{\mathbf{W}}^{-1}} = \sum_{\mathbf{k}\in\mathcal{I}} \frac{|\hat{f}_{\mathbf{k}}|^2}{\hat{w}_{\mathbf{k}}} \xrightarrow{\hat{\mathbf{f}}} \min \quad \text{subject to} \quad \mathbf{A}\hat{\mathbf{f}} = \mathbf{f}, \tag{1.8}$$

for the under-determined case $|\mathcal{I}| > |\mathcal{X}|$, where the weights $\hat{w}_{\mathbf{k}}$ damp high-frequency components, and

3. the sparse recovery problem

$$\|\hat{\mathbf{f}}\|_0 = |\{\mathbf{k}\in\mathcal{I} : \hat{f}_{\mathbf{k}} \neq 0\}| \xrightarrow{\hat{\mathbf{f}}} \min \quad \text{subject to} \quad \mathbf{A}\hat{\mathbf{f}} = \mathbf{f}, \tag{1.9}$$

for the under-determined case $|\mathcal{I}| > |\mathcal{X}|$.

The main tool in iterative methods to solve these three problems is the use of fast matrix-vector multiplications with $\mathbf{A}$ and $\mathbf{A}^*$ and bounding involved condition numbers uniformly.

1.3.1 FFT and nonequispaced FFT

The reconstruction of the Fourier coefficients $\hat{f}_{\mathbf{k}}$, $\mathbf{k} \in \hat{G}_n^d$ from sampling values of an equispaced grid $\mathbf{x} \in \mathcal{X} = (2^{-n}\hat{G}_n^d \bmod \mathbf{1})$, see (1.3), can be realized by the inverse fast Fourier transform, since the Fourier matrix $\mathbf{F} := \mathbf{A}(2^{-n}\hat{G}_n^d, \hat{G}_n^d)$ has orthogonal columns, and takes $\mathcal{O}(N^d \log N)$ floating point operations. This is no longer true for the nonequispaced Fourier matrix given by

$$\mathbf{A} := \mathbf{A}(\mathcal{X}, \hat{G}_n^d) = \left(\mathrm{e}^{2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{x}_\ell}\right)_{\ell=0,\dots,L-1,\mathbf{k}\in\hat{G}_n^d}.$$

Here, we use an iterative algorithm since the fast matrix times vector multiplication with the matrix $\mathbf{A}$ and $\mathbf{A}^\top$ takes only $\mathcal{O}(2^{nd}n + |\log\varepsilon|^d L)$ floating point operations, see [39]. The conditioning of the reconstruction problems relies on the uniformity of $\mathcal{X}$, measured by the mesh norm and the separation distance

$$\delta := 2\max_{\mathbf{x}\in\mathbb{T}^d}\min_{j=0,\dots,L-1}\operatorname{dist}(\mathbf{x}_j,\mathbf{x}), \qquad q := \min_{j,l=0,\dots,L-1;j\neq l}\operatorname{dist}(\mathbf{x}_j,\mathbf{x}_l),$$

where $\operatorname{dist}(\mathbf{x},\mathbf{x}_0) := \min_{\mathbf{j}\in\mathbb{Z}^d}\|(\mathbf{x}+\mathbf{j})-\mathbf{x}_0\|_\infty$, respectively.

For the overdetermined case $N^d < L$, it has been proven in [24] that the reconstruction problem (1.7) has a unique solution if $N < (\frac{\pi}{\log 2}\, d\, \delta)^{-1}$. The solution is computed iteratively by means of the conjugate gradient method in [18, 2, 23], where the multilevel Toeplitz structure of $\mathbf{A}^\top\mathbf{W}\mathbf{A}$ is used for fast matrix vector multiplications. Slightly more stable with respect to rounding errors is the CGNR method, cf. [6, pp. 288], which iterates the original residual $\mathbf{r}_l = \mathbf{y} - \mathbf{A}\hat{\mathbf{f}}_l$ instead of the residual $\mathbf{A}^\top\mathbf{W}\mathbf{r}_l$ of the normal equations. Further analysis of the numerical stability of the least squares approximation (1.7) relies on so-called Marcinkiewicz-Zygmund inequalities

which establish norm equivalences between a trigonometric polynomial and its samples, see e.g. [60, 44, 19, 38] and references therein for specific variants.

For the underdetermined case $N^d > L$, the optimal interpolation problem (1.8) has been shown to be stable in [41] if the sampling set is well separated with respect to the polynomial degree and the weights $\hat{w}_{\mathbf{k}}$ are constructed by means of a so-called smoothness-decay principle. In particular, we proved that the nonequispaced Fourier matrix $\mathbf{A}$ has full rank L for every polynomial degree $N > 2\,d\,q^{-1}$ and proposed to solve problem (1.8) by a version of the conjugate gradient method in combination with the nonequispaced FFT to efficiently perform each iteration step.

1.3.2 Hyperbolic cross FFT

The inverse HCFFT is not an orthogonal transform, but a fast algorithm is developed by reverting all steps of the HCFFT, see [3, 27], which makes this spatial discretisation most attractive in terms of efficiency. The inverse HCFFT requires also only $\mathcal{O}(2^n n^d)$ floating point operations. However, we proved in [35] that this transform is mildly ill conditioned, i.e.,

$$\begin{aligned} c_d 2^{\frac{n}{2}} n^{\frac{2d-3}{2}} \leq \operatorname{cond}_2 \mathbf{A}(H_n^d, S_n^d) \leq C_d 2^{\frac{n}{2}} n^{2d-2}, \quad n \to \infty, \\ c_n d^{2n} \leq \operatorname{cond}_2 \mathbf{A}(H_n^d, S_n^d) \leq C_n d^{2n}, \quad d \to \infty. \end{aligned}$$

In particular, we loose more than 5 decimal digits of accuracy already for $d = 10$ and $n = 5$ in the worst case.

1.3.3 Lattice and generated set FFT

As pointed out in Section 1.2.3, the evaluation of trigonometric polynomials with frequencies supported on an arbitrary index set $\mathcal{I}$, i.e., the mapping from the index set $\mathcal{I}$ in frequency domain to the rank-1 lattice in spatial domain reduces to a single one-dimensional FFT and thus can be computed very efficiently and stable. For the inverse transform, mapping the samples of a trigonometric polynomial to its Fourier coefficients on a specific frequency index set, we discuss the recently presented necessary and sufficient conditions on rank-1 lattices allowing a stable reconstruction of trigonometric polynomials supported on hyperbolic crosses and the generalisation to arbitrary index sets in the frequency domain. We suggest approaches for determining suitable rank-1 lattices using a component-by-component strategy [33, 12]. In conjunction with numerical found lattices, we show that this new method outperforms the classical hyperbolic cross FFT for realistic problem sizes [36].

The use of generated sets, a generalisation of rank-1 lattices, as spatial discretisations offers an additional suitable possibility for sampling sparse trigonometric polynomials. The fast computation of trigonometric polynomials on generated sets can be realized using the nonequispaced fast Fourier transform (NFFT), cf. [39]. A simple sufficient condition on a generated set $\Lambda(\mathbf{r}, L)$ allows the fast, unique and stable reconstruction of the frequencies of a d-dimensional trigonometric polynomial from its samples along $\Lambda(\mathbf{r}, L)$. In contrast to searching for suitable rank-1 lattices, we can use continuous optimization methods in order to determine generated sets that are suitable for reconstruction, see [32].

Reconstruction using rank-1 lattices. In order to state constructive existence results for reconstructing rank-1 lattices, i.e., rank-1 lattices which allow the unique reconstruction of trigonometric polynomials supported on a fixed frequency index set $\mathcal{I}$, we define the difference set

$$\mathcal{D}(\mathcal{I}) := \{\mathbf{k} - \mathbf{l} \colon \mathbf{k}, \mathbf{l} \in \mathcal{I}\}$$

of the frequency index set $\mathcal{I}$. As a consequence of [34, Cor. 1] we formulate the following

Theorem 1.2. *Let $\mathcal{I} \subset \{\mathbf{k} \in \mathbb{Z}^d \colon \mathbf{k} - \mathbf{a} \in [0, |\mathcal{I}|-1]^d\}$ for a fixed $\mathbf{a} \in \mathbb{Z}^d$ being a frequency index set of finite cardinality. Then there exists a reconstructing rank-1 lattice of prime cardinality L,*

$$|\mathcal{I}| \leq L \leq |\mathcal{D}(\mathcal{I})| \leq |\mathcal{I}|^2, \tag{1.10}$$

such that all trigonometric polynomials f with frequencies supported on $\mathcal{I}$ can be reconstructed from the sampling values $(f(\mathbf{x}))_{\mathbf{x} \in \Lambda(\mathbf{r}, L)}$. Moreover, the corresponding generating vector $\mathbf{r} \in L^{-1}\mathbb{Z}^d$ can be determined using a component–by–component strategy and the reconstruction of the Fourier coefficients can be realized by a single one-dimensional FFT of length L, and thus takes $\mathcal{O}(L \log L + d|\mathcal{I}|)$ floating point operations.

Proof. The result follows from [34, Cor. 1], Bertrand's postulate, and equations (1.4) and (1.5). □

We stress on the fact, that [34, Cor. 1] is a more general result on arbitrary frequency index sets $\mathcal{I}$. Some simple additional assumptions on L allow to replace the condition $\mathcal{I} \subset \{\mathbf{k} \in \mathbb{Z}^d \colon \mathbf{k} - \mathbf{a} \in [0, |\mathcal{I}|-1]^d\}$ by $\mathcal{I} \subset \mathbb{Z}^d$, $|\mathcal{I}| < \infty$.

In fact, the cardinality of the difference set $\mathcal{D}(\mathcal{I})$ is the theoretical upper bound in (1.10) for the number of samples needed to reconstruct trigonometric polynomials with frequencies supported on the index set $\mathcal{I}$ using a rank-1 lattice. This cardinality depends mainly on the structure of $\mathcal{I}$.

Example 1.1. Let $\mathcal{I} = I_{p,N}^d := \{\mathbf{k} \in \mathbb{Z}^d \colon \|\mathbf{k}\|_p \leq N\}$, $N \in \mathbb{N}$, be the ℓ_p-ball, $0 < p \leq \infty$, of size N, see Fig. 1.3. The cardinality of $I_{p,N}^d$ is bounded by $c_{p,d} N^d \leq |I_{p,N}^d| \leq C_d N^d$ and $c_{p,d} N^d \leq \mathcal{D}(I_{p,N}^d) \leq C_d 2^d N^d$, $c_{p,d}, C_d \in \mathbb{R}$,

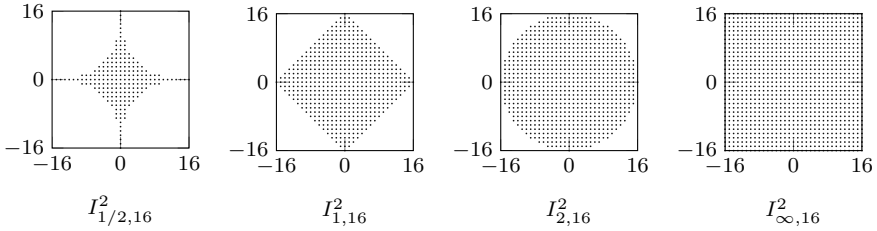

Fig. 1.3 Two-dimensional frequency index sets $I^2_{p,16}$ for $p \in \{\frac{1}{2}, 1, 2, \infty\}$.

$0 < c_{p,d} \leq C_d$. Consequently, we can find a reconstructing rank-1 lattice of size $L \leq \tilde{C}_{p,d}|I^d_{p,N}|$, $\tilde{C}_{p,d} > 0$, using a component-by-component strategy.

On the other hand, we obtain for the limit $p \to 0$ the frequency index set $\mathcal{I} := \{\mathbf{k} \in \mathbb{Z}^d \colon \|\mathbf{k}\|_1 = \|\mathbf{k}\|_\infty \leq N\}$, $N \in \mathbb{N}$, which is supported on the coordinate axis. We have $|\mathcal{I}| = 2dN+1$ and $(2N+1)^2 \leq |\mathcal{D}(\mathcal{I})| \leq (2dN+1)^2$. Hence, we estimate $\tilde{c}_d|\mathcal{I}|^2 \leq |\mathcal{D}(\mathcal{I})|$, $\tilde{c}_d \in \mathbb{R}$, $0 < \tilde{c}_d$, and the theoretical upper bound on L is quadratic in $|\mathcal{I}|$ for fixed dimension d. In fact, reconstructing rank-1 lattices for these specific frequency index sets need at least a number of $L \in \Omega(N^2)$ nodes, cf. [36, Thm. 3.5]. □

Example 1.2. More useful frequency index sets in higher dimensions $d > 2$ are so-called (energy-norm based) hyperbolic crosses, cf. [3, 7, 8]. In particular, we consider frequency index sets $\mathcal{I}$ of the form

$$\mathcal{I}_N^{d,T} := \left\{ \mathbf{k} \in \mathbb{Z}^d \colon \max(1, \|\mathbf{k}\|_1)^{\frac{T}{T-1}} \prod_{s=1}^{d} \max(1, |k_s|)^{\frac{1}{1-T}} \leq N \right\},$$

with parameter $T \in [0, 1)$ and $N \in \mathbb{N}$, see Fig. 1.4 for illustration. The frequency index set $\mathcal{I}_N^{d,0}$, i.e., $T = 0$, is in fact a symmetric hyperbolic cross and frequency index sets $\mathcal{I}_N^{d,T}$, $T \in (0, 1)$, are called energy-norm based hyperbolic crosses. The cardinality of $\mathcal{I}_N^{d,T}$ can be estimated, cf. [37, Lem. 2.6], by

$$\begin{aligned} c_{d,0} N \log^{d-1} N \leq |\mathcal{I}_N^{d,T}| \leq C_{d,0} N \log^{d-1} N, \quad \text{for } T = 0, \\ c_{d,T} N \leq |\mathcal{I}_N^{d,T}| \leq C_{d,T} N, \quad \text{for } T \in (0, 1), \end{aligned}$$

where $c_{d,T}, C_{d,T} \in \mathbb{R}$, $0 < c_{d,T} \leq C_{d,T}$. Since the axis cross is a subset of the considered frequency index sets, i.e., $\{\mathbf{k} \in \mathbb{Z}^d \colon \|\mathbf{k}\|_1 = \|\mathbf{k}\|_\infty \leq N\} \subset \mathcal{I}_N^{d,T}$, $T \in [0, 1)$, we obtain $(2N+1)^2 \leq |\mathcal{D}(\mathcal{I}_N^{d,T})|$. On the other hand, we obtain upper bounds of the cardinality of the difference set $\mathcal{D}(\mathcal{I}_N^{d,T})$

$$\begin{aligned} |\mathcal{D}(\mathcal{I}_N^{d,T})| \leq \tilde{C}_{d,0} N^2 \log^{d-2} N, \quad \text{for } T = 0, \text{ cf. [33, Thm. 4.8]}, \\ |\mathcal{D}(\mathcal{I}_N^{d,T})| \leq |\mathcal{I}_N^{d,T}|^2 \leq C_{d,T}^2 N^2, \quad \text{for } T \in (0, 1). \end{aligned}$$

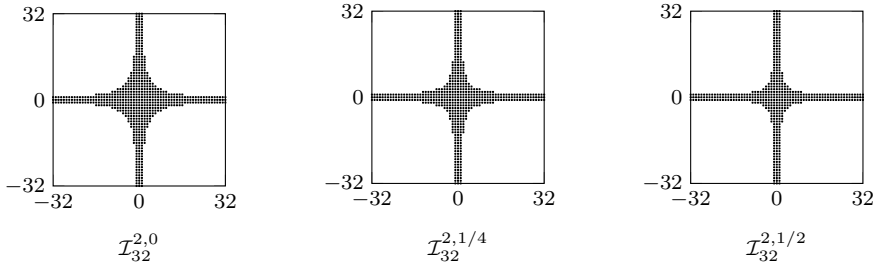

Fig. 1.4 Two-dimensional frequency index sets $\mathcal{I}_{32}^{2,T}$ for $T \in \{0, \frac{1}{4}, \frac{1}{2}\}$.

Consequently, Theorem 1.2 offers a constructive strategy in order to find reconstructing rank-1 lattices for $\mathcal{I}_N^{d,T}$ of cardinality $L \leq |\mathcal{D}(\mathcal{I}_N^{d,T})|$. We would like to stress that, at least for $T \in (0,1)$, we are able to construct rank-1 lattices of optimal order in N, cf. [33, Lem. 2.1, 2.3, and Cor. 2.4].

For instance, Fig. 1.1b(left) shows a reconstructing rank-1 lattice for the symmetric hyperbolic cross $\mathcal{I}_8^{2,0}$ and Fig. 1.1b(right) shows an example for a generated set, which allows the exact reconstruction of trigonometric polynomials with frequencies supported on $\mathcal{I}_8^{2,0}$. The condition number of the Fourier matrix $\mathbf{A}(\mathcal{I}, \mathcal{X})$ is always one when $\mathcal{X}$ is a reconstructing rank-1 lattice for $\mathcal{I}$, since the columns of the Fourier matrix $\mathbf{A}(\mathcal{I}, \mathcal{X})$ are orthogonal. When the frequency index set $\mathcal{I} = \mathcal{I}_8^{2,0}$ and $\mathcal{X}$ is the specific generated set in Fig. 1.1b(right), then the condition number of the Fourier matrix $\mathbf{A}(\mathcal{I}, \mathcal{X})$ is approximately 2.19. □

Reconstruction using generated sets. Up to now, we discussed reconstructing rank-1 lattices. We generalized this concept to so-called generated sets, cf. Section 1.2.3 and determined sufficient and necessary conditions on generated sets $\Lambda(\mathbf{r}, L)$ guaranteeing a full rank and stable Fourier matrix $\mathbf{A}(\mathcal{I}, \Lambda(\mathbf{r}, L))$ in [32]. In general, the set $\mathcal{Y} = \{\mathbf{k} \cdot \mathbf{r} \bmod 1 : \mathbf{k} \in \mathcal{I}\} \subset \mathbb{T}$ is of our main interest, where $\mathbf{r} \in \mathbb{R}^d$ is the generating vector of the generated set $\Lambda(\mathbf{r}, L)$. We determined the necessary condition $|\mathcal{Y}| = |\mathcal{I}|$ in order to obtain a Fourier matrix $\mathbf{A}(\mathcal{I}, \Lambda(\mathbf{r}, L))$ of full column rank.

Theorem 1.3. *Let $\mathcal{I} \subset \mathbb{Z}^d$ be an arbitrary d-dimensional index set of finite cardinality $|\mathcal{I}|$. Then, the exact reconstruction of a trigonometric polynomial with frequencies supported on $\mathcal{I}$ is possible from only $|\mathcal{I}|$ samples using a suitable generated set.*

Proof. Let $\mathbf{r} \in \mathbb{R}^d$ be a vector such that

$$\mathbf{k} \cdot \mathbf{r} \bmod 1 \neq \mathbf{k}' \cdot \mathbf{r} \bmod 1 \quad \text{for all } \mathbf{k}, \mathbf{k}' \in \mathcal{I},\ \mathbf{k} \neq \mathbf{k}'. \tag{1.11}$$

For instance, Theorem 1.2 guarantees the existence of a reconstructing rank-1 lattice $\Lambda(\mathbf{r}, L)$ for the index set $\mathcal{I}$, where $\mathbf{r} \in L^{-1}\mathbb{Z}^d$ fulfills property (1.11). The corresponding Fourier matrix $\mathbf{A} := (\mathrm{e}^{2\pi \mathrm{i} \mathbf{k} \cdot \mathbf{x}_\ell})_{\ell=0,\ldots,L-1;\, \mathbf{k} \in \mathcal{I}} =$

$(\mathrm{e}^{(2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{r})^\ell})_{\ell=0,\ldots,L-1;\,\mathbf{k}\in\mathcal{I}}$ is a transposed Vandermonde matrix of (full column) rank $|\mathcal{I}|$. If we use only the first $|\mathcal{I}|$ rows of the matrix $\mathbf{A}$ and denote this matrix by $\widetilde{\mathbf{A}}$, the matrix $\widetilde{\mathbf{A}} := (\mathrm{e}^{(2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{r})^\ell})_{\ell=0,\ldots,|\mathcal{I}|-1;\,\mathbf{k}\in\mathcal{I}} = (\mathrm{e}^{(2\pi\mathrm{i}y_j)^\ell})_{\ell=0,\ldots,|\mathcal{I}|-1;\,j=0,\ldots,|\mathcal{I}|-1}$ is a transposed Vandermonde matrix of size $|\mathcal{I}|\times|\mathcal{I}|$, where $y_j := \mathbf{k}_j\cdot\mathbf{r} \bmod 1$ and $\mathcal{I} = \{\mathbf{k}_0,\ldots,\mathbf{k}_{|I|-1}\}$ in the specified order. Furthermore, the determinant of the transposed Vandermonde matrix $\widetilde{\mathbf{A}}$, cf. [31, Sec. 6.1], is $\det\widetilde{\mathbf{A}} = \prod_{1\le k<j\le|I|-1}(\mathrm{e}^{2\pi\mathrm{i}y_j} - \mathrm{e}^{2\pi\mathrm{i}y_k}) \neq 0$, since we have $\mathrm{e}^{2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \neq \mathrm{e}^{2\pi\mathrm{i}\mathbf{k}'\cdot\mathbf{r}}$ for all $\mathbf{k},\mathbf{k}'\in\mathcal{I}$, $\mathbf{k}\neq\mathbf{k}'$, due to property (1.11). This means the transposed Vandermonde matrix $\widetilde{\mathbf{A}}$ has full rank $|\mathcal{I}|$ and is invertible. □

Theorem 1.3 states that $L = |\mathcal{I}|$ many samples are sufficient to exactly reconstruct a trigonometric polynomial with frequencies supported on the index set $\mathcal{I}$. In general, we obtain a large condition number for the Fourier matrix $\widetilde{\mathbf{A}} := (\mathrm{e}^{(2\pi\mathrm{i}\mathbf{k}\cdot\mathbf{r})^\ell})_{\ell=0,\ldots,|\mathcal{I}|-1;\,\mathbf{k}\in\mathcal{I}}$. Using $L > |\mathcal{I}|$ samples, we also obtain matrices $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L))$ of full column rank, since the first $|\mathcal{I}|$ rows of the matrix $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L))$ are linear independent. In practice, growing oversampling, i.e., increasing $L > |\mathcal{I}|$, decreases at least an estimator of the condition number of $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L))$, as published in [32]. In this context, for each generating vector $\mathbf{r}\in\mathbb{R}^d$ bringing $|\mathcal{Y}| = |\mathcal{I}|$ and constant $C > 1$ we determined a generated set of size L_C such that the Fourier matrix $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L_C))$ has a condition number of at most C, cf. [32, Cor. 1]. We discuss a nonlinear optimization strategy in [32] in order to determine generated sets $\Lambda(\mathbf{r},L)$ of relatively small cardinality bringing a Fourier matrix $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L))$ with small condition number.

The reconstruction of trigonometric polynomials with frequencies supported on an fixed index set $\mathcal{I}$ from samples along a generated set can be realized solving the normal equation, which can be done in a fast way using the one-dimensional NFFT and a conjugate gradient (CG) method. One step of the CG method needs one NFFT of length L and one adjoint NFFT of length L. Consequently, one CG step has a complexity of $\mathcal{O}(L\log L+(d+|\log\epsilon|)|I|)$, cf. Section 1.2.3. The convergence of the CG method depends on the condition number of the Fourier matrix $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L))$. Hence, generated sets $\Lambda(\mathbf{r},L)$ with small condition numbers of the Fourier matrices $\mathbf{A}(\mathcal{I},\Lambda(\mathbf{r},L))$ guarantee a fast approximative computation of the reconstruction of trigonometric polynomials with frequencies supported on the index set $\mathcal{I}$.

1.3.4 Random sampling and sparse recovery

Stable deterministic sampling schemes with a minimal number of nodes are constructed above. For arbitrary index sets of frequencies $\mathcal{I}\subset\mathbb{Z}^d$, we showed that orthogonality of the Fourier matrix necessarily implies $|\mathcal{X}| \geq |\mathcal{D}(\mathcal{I})|$ which scales (almost) quadratically in $|\mathcal{I}|$ for several interesting cases. In

contrast, injectivity of the Fourier matrix can be guaranteed for a linear scaling and numerical results also support that a small oversampling factor suffices for stable reconstruction generically. Subsequently, we discuss known results for randomly chosen sampling nodes. Let $d \in \mathbb{N}$, arbitrary frequencies $\mathcal{I} \subset \mathbb{Z}^d$ be given, and sampling nodes $\mathcal{X}$ are drawn independently from the uniform distribution over the spatial domain $\mathbb{T}^d$, then [25] implies

$$\operatorname{cond}_2 \mathbf{A}(\mathcal{I}, \mathcal{X}) \leq \sqrt{\frac{1+\gamma}{1-\gamma}},\ \gamma \in (0,1), \quad \text{if} \quad |\mathcal{X}| \geq \frac{C}{\gamma^2} |\mathcal{I}| \log \frac{|\mathcal{I}|}{\eta},$$

with probability $1-\eta$, where $C > 0$ is some universal constant independent of the spatial dimension d. A partial derandomization can be obtained by randomly subsampling a fixed rank-1 lattice as constructed in Theorem 1.2.

Moreover, random sampling has been applied sucessfully in compressed sensing [15, 9, 20] to solve the sparse recovery problem (1.9), where both the support $\mathcal{I} \subset \mathcal{I}_0 \subset \mathbb{Z}^d$ as well as the Fourier coefficients $\hat{f}_{\mathbf{k}} \in \mathbb{C}$, $\mathbf{k} \in \mathcal{I}$, of the expansion (1.1) are sought. Provided a so-called restricted isometry condition is met, the sparse recovery problem can be solved efficiently, cf. [10, 54, 55, 56, 46, 42], and with probability at least $1-\eta$ this is true if

$$|\mathcal{X}| \geq C |\mathcal{I}| \log^4 |\mathcal{I}_0| \log \frac{1}{\eta}.$$

Well studied algorithmic approaches to actually solve the sparse recovery problem are then ℓ^1-minimisation [11], orthogonal matching pursuit [43], and their successors. Optimal variants of these algorithms have the same arithmetic complexity as one matrix vector multiplication with $\mathbf{A}(\mathcal{I}_0, \mathcal{X})$, which is however worse than the recent developments [29, 28].

Prony type methods. In contrast to compressed sensing approaches, Prony type methods aim to recover the finite and real support $\mathcal{I}$ within the bounded interval $[-\frac{N}{2}, \frac{N}{2}]$ as well as the Fourier coefficients in the nonharmonic Fourier series

$$f(x) = \sum_{k \in \mathcal{I}} \hat{f}_k \mathrm{e}^{2\pi \mathrm{i} k x},$$

from equally spaced samples $f(\frac{\ell}{N})$, $\ell = 0, \ldots, L-1$, cf. [52, 50, 49]. If the number of samples fulfils a Nyquist type relation

$$|\mathcal{X}| \geq C N q_{\mathcal{I}}^{-1}$$

with respect to the nonharmonic bandwidth N and to the separation distance $q_{\mathcal{I}} := \min\{|k-k'| : k, k' \in \mathcal{I}, k \neq k'\}$, then a newly developed variant of the Prony method solves this reconstruction problem in a stable way, see e.g. [53]. The arithmetic complexity $\mathcal{O}(|\mathcal{I}|^3)$ has been improved for integer frequencies in [30] using ideas from [29, 28].

Acknowledgements We gratefully acknowledge support by the German Research Foundation (DFG) within the Priority Program 1324, project PO 711/10-2 and KU 2557/1-2. Moreover, Ines Melzer and Stefan Kunis gratefully acknowledge their support by the Helmholtz Association within the young investigator group VH-NG-526.

References

1. A. A. Aydıner, W. C. Chew, J. Song, and T. J. Cui. A sparse data fast Fourier transform (SDFFT). *IEEE Trans. Antennas and Propagation*, 51(11):3161–3170, 2003.
2. R. F. Bass and K. Gröchenig. Random sampling of multivariate trigonometric polynomials. *SIAM J. Math. Anal.*, 36:773 – 795, 2004.
3. G. Baszenski and F.-J. Delvos. A discrete Fourier transform scheme for Boolean sums of trigonometric operators. In C. K. Chui, W. Schempp, and K. Zeller, editors, *Multivariate Approximation Theory IV*, ISNM 90, pages 15 – 24. Birkhäuser, Basel, 1989.
4. M. Bebendorf. *Hierarchical Matrices*, volume 63 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, 2008.
5. G. Beylkin. On the fast Fourier transform of functions with singularities. *Appl. Comput. Harmon. Anal.*, 2:363 – 381, 1995.
6. Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, PA, USA, 1996.
7. H.-J. Bungartz and M. Griebel. A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. *J. Complexity*, 15:167 – 199, 1999.
8. H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147 – 269, 2004.
9. E. J. Candès. Compressive sampling. In *International Congress of Mathematicians. Vol. III*, pages 1433 – 1452. Eur. Math. Soc., Zürich, 2006.
10. E. J. Candès and T. Tao. Decoding by linear programming. IEEE *Trans. Inform. Theory*, 51:4203 – 4215, 2005.
11. S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20:33 – 61, 1998.
12. R. Cools, F. Y. Kuo, and D. Nuyens. Constructing lattice rules based on weighted degree of exactness and worst case error. *Computing*, 87:63 – 89, 2010.
13. L. Demanet, M. Ferrara, N. Maxwell, J. Poulson, and L. Ying. A butterfly algorithm for synthetic aperture radar imaging. *SIAM J. Imaging Sci.*, 5(1):203–243, 2012.
14. J. Dick, F. Y. Kuo, and I. H. Sloan. High-dimensional integration: The quasi-monte carlo way. *Acta Numer.*, 22:133 – 288, 2013.
15. D. L. Donoho. Compressed sensing. IEEE *Trans. Inform. Theory*, 52:1289 – 1306, 2006.
16. A. Dutt and V. Rokhlin. Fast Fourier transforms for nonequispaced data II. *Appl. Comput. Harmon. Anal.*, 2:85 – 100, 1995.
17. A. Edelman, P. McCorquodale, and S. Toledo. The future fast Fourier transform? *SIAM Journal on Scientific Computing*, 20:1094–1114, 1999.
18. H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numer. Math.*, 69:423 – 440, 1995.
19. F. Filbir and W. Themistoclakis. Polynomial approximation on the sphere using scattered data. *Math. Nachr.*, 281:650 – 668, 2008.
20. S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2013.
21. L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73:325 – 348, 1987.
22. M. Griebel and J. Hamaekers. Fast discrete Fourier transform on generalized sparse grids. Univ. Bonn, INS Preprint No. 1305, 2013.

23. D. Grishin and T. Strohmer. Fast multi-dimensional scattered data approximation with Neumann boundary conditions. *Lin. Alg. Appl.*, 391:99 – 123, 2004.
24. K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comput.*, 59:181 – 194, 1992.
25. K. Gröchenig, B. Pötscher, and H. Rauhut. Learning trigonometric polynomials from random samples and exponential inequalities for eigenvalues of random matrices. *Preprint arXiv:math/0701781*, 2007.
26. W. Hackbusch. *Hierarchische Matrizen. Algorithmen und Analysis*. Springer-Verlag, 2009.
27. K. Hallatschek. Fouriertransformation auf dünnen Gittern mit hierarchischen Basen. *Numer. Math.*, 63:83 – 97, 1992.
28. H. Hassanieh, P. Indyk, D. Katabi, and E. Price. Nearly optimal sparse Fourier transform. In *STOC*, 2012.
29. H. Hassanieh, P. Indyk, D. Katabi, and E. Price. Simple and practical algorithm for sparse Fourier transform. In *SODA*, pages 1183–1194, 2012.
30. S. Heider, S. Kunis, D. Potts, and M. Veit. A sparse prony fft. *10th International Conference on Sampling Theory and Applications*, 2013.
31. R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1991.
32. L. Kämmerer. Reconstructing multivariate trigonometric polynomials by sampling along generated sets. *Monte Carlo and Quasi-Monte Carlo Methods 2012*, 2012.
33. L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along rank-1 lattices. *SIAM J. Numer. Anal.*, 51:2773 – 2796, 2013.
34. L. Kämmerer. Reconstructing multivariate trigonometric polynomials from samples along rank-1 lattices. *submitted*, 2013.
35. L. Kämmerer and S. Kunis. On the stability of the hyperbolic cross discrete Fourier transform. *Numer. Math.*, 117:581 – 600, 2011.
36. L. Kämmerer, S. Kunis, and D. Potts. Interpolation lattices for hyperbolic cross trigonometric polynomials. *J. Complexity*, 28:76 – 92, 2012.
37. L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate functions by trigonometric polynomials based on rank-1 lattice sampling. *Preprint 145, DFG Priority Program 1324*, 2013.
38. J. Keiner, S. Kunis, and D. Potts. Fast summation of radial functions on the sphere. *Computing*, 78:1 – 15, 2006.
39. J. Keiner, S. Kunis, and D. Potts. Using NFFT3 - a software library for various nonequispaced fast Fourier transforms. *ACM Trans. Math. Software*, 36:Article 19, 1 – 30, 2009.
40. S. Kunis and I. Melzer. A stable and accurate butterfly sparse Fourier transform. *SIAM J. Numer. Anal.*, 50(3):1777–1800, 2012.
41. S. Kunis and D. Potts. Stability results for scattered data interpolation by trigonometric polynomials. *SIAM J. Sci. Comput.*, 29:1403 – 1419, 2007.
42. S. Kunis and H. Rauhut. Random sampling of sparse trigonometric polynomials II, Orthogonal matching pursuit versus basis pursuit. *Found. Comput. Math.*, 8:737 – 763, 2008.
43. S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. IEEE *Trans. Sig. Proc.*, 41:3397 – 3415, 1993.
44. H. N. Mhaskar, F. J. Narcowich, and J. D. Ward. Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature. *Math. Comput.*, 70:1113 – 1130, 2001. Corrigendum on the positivity of the quadrature weights in 71:453 – 454, 2002.
45. E. Michielssen and A. Boag. A multilevel matrix decomposition algorithm for analyzing scattering from large structures. *IEEE Trans. Antennas and Propagation*, 44:1086 – 1093, 1996.
46. D. Needell and R. Vershynin. Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Found. Comput. Math.*, 9:317 – 334, 2009.

47. E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems Volume II: Standard Information for Functionals*. Eur. Math. Society, EMS Tracts in Mathematics Vol 12, 2010.
48. M. O'Neil, F. Woolfe, and V. Rokhlin. An algorithm for the rapid evaluation of special function transforms. *Appl. Comput. Harmon. Anal.*, 28:203 – 226, 2010.
49. T. Peter and G. Plonka. A generalized prony method for reconstruction of sparse sums of eigenfunctions of linear operators. *Inverse Problems*, 29:025001, 2013.
50. T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM J. Sci. Comput.*, 33:314 – 334, 2011.
51. D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247 – 270, Boston, MA, USA, 2001. Birkhäuser.
52. D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Process.*, 90:1631 – 1642, 2010.
53. D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra Appl.*, 439:1024 – 1039, 2013.
54. H. Rauhut. Random sampling of sparse trigonometric polynomials. *Appl. Comput. Harmon. Anal.*, 22:16 – 42, 2007.
55. H. Rauhut. On the impossibility of uniform sparse reconstruction using greedy methods. *Sampl. Theory Signal Image Process.*, 7:197 – 215, 2008.
56. H. Rauhut. Stability results for random sampling of sparse trigonometric polynomials. IEEE *Trans. Inform. Theory*, 54:5661 – 5670, 2008.
57. W. Sickel and T. Ullrich. The Smolyak algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness. *East J. Approx.*, 13:387 – 425, 2007.
58. I. H. Sloan and S. Joe. *Lattice methods for multiple integration*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1994.
59. G. Steidl. A note on fast Fourier transforms for nonequispaced grids. *Adv. Comput. Math.*, 9:337 – 353, 1998.
60. V. N. Temlyakov. Approximation of functions with bounded mixed derivative. *Proc. Steklov Inst. Math.*, 1989. A translation of Trudy Mat. Inst. Steklov 178 (1986).
61. M. Tygert. Fast algorithms for spherical harmonic expansions, III. *J. Comput. Phys.*, 229(18):6181–6192, 2010.
62. L. Ying. Sparse Fourier transform via butterfly algorithm. *SIAM J. Sci. Comput.*, 31:1678 – 1694, 2009.
63. L. Ying, G. Biros, and D. Zorin. A kernel-independent adaptive fast multipole method in two and three dimensions. *J. Comput. Physics*, 196:591 – 626, 2004.
64. L. Ying and S. Fomel. Fast computation of partial Fourier transforms. *Multiscale Model. Simul.*, 8(1):110–124, 2009.